Young Children's Human Figure Drawings:

An Investigation using the Goodenough-Harris Drawing Test

And the Rasch Model for Measurement

Thesis submitted by

Claire MALEY BEd (ECE) (Hons)

in September 2009

for the degree of Doctor of Philosophy

in the School of Education James Cook University

Statement of Access

I, the undersigned, author of this work, understand that James Cook University will make this thesis available for use within the university library and, via the Australian Digital Theses network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and; I do not wish to place any further restriction on access to this work.

____________________________ ____________________

Signature Date

Declaration on Sources

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Signature   Date

Contributions of Others

Financial support

I was awarded an Australian Postgraduate Award (APA) Scholarship which commenced in February 2006 and terminated in July 2009. Also, the School of Education awarded me a 'top up' scholarship from February 2006 until January 2009.

I also applied for, and received, several Graduate Research School (GRS) Grants which enabled me to travel to and present at four conferences (three of these were international symposia) as well as attend the associated workshops.

A portion of the School of Education Student Stipend was redeemed to cover the costs of consumables used in this research.

Editorial Support and Research Assistance

Professor Trevor Bond, my Principal Supervisor, provided feedback and advice on all aspects of the research design and this thesis. Although, all data were collected, scored and analysed by myself. Throughout my PhD candidature, Trevor and I co-authored six paper presentations. In particular, Trevor and I co-authored a refereed paper presentation for AARE in 2007; this paper was later published in the AARE conference proceedings. I sincerely thank Trevor for offering his time and energy to me.

Doctor Helen Boon, my Co-Supervisor, gave feedback and advice on the more completed chapters, and offered her understanding of the doctoral research process. Helen's supervision enhanced my research journey and I thank her for this.

Declaration on Ethics

The research presented and reported in this thesis was conducted within the guidelines for research ethics outlined in the *National Statement on Ethics Conduct in Research Involving Human Participants* (1999), the *Joint NHMRC/AVCC Statement and Guidelines on Research Practice* (1997), the *James Cook University Policy on Experimental Ethics, Standard Practices, and Guidelines* (2001), and the *James Cook University Statement and Guidelines on Research Practice* (2001). The proposed research methodology received clearance from the James Cook University Experimental Ethics review Committee (approval number: H2450).

Signature Date

Acknowledgements

This research is immensely significant to me. Upon commencing my Bachelor of Education degree, I never imagined that I would one day complete a doctorate. I simply enrolled in the four year course, tried my hardest, and was deeply excited about one day working with young children. I had only ever expected to be an early childhood teacher, in my classroom facilitating my young students in becoming lifelong learners and lovers of life. Thus, being offered the opportunity to complete a PhD – and get paid for it! – was a delightful surprise.

Many people at James Cook University supported and assisted me in completing this work. I sincerely thank my Principal Supervisor Professor Trevor Bond and my Co-Supervisor Doctor Helen Boon for their guidance, feedback and support. Additionally, I thank Associate Professor Malcolm Vick for his encouragement, interest and advice. He and Trevor were the people who suggested I apply for an APA scholarship after completing Honours; I am sincerely appreciative of their faith in me. Also, my fellow Post-Graduate friends were incredibly understanding of the arduous research process – I thank them for this.

I thank the Director of the Townsville Catholic Education Office, Doctor Cathy Day, and the Principal of Ryan Catholic College, Andrea O'Brien, for allowing this research to take place within their wonderful school. I also thank the students and families of Ryan Catholic College for providing me with hundreds of lovely human figure drawings.

I am continually inspired by, and in awe of, the numerous wonderful women in my life. In particular, I thank: my sister Louise, Corinne, Kia, Jo, Andrea P, Stacey, Jacine, Larelle, Sally, Margaret, Andrea O, and Paula.

I thank my Dad and my sister Louise for their love, understanding, and confidence in me. They have endured three years of an absent daughter and sister, and I am genuinely grateful for their patience.

I am incredibly indebted to my partner Malcolm. He was the inspiration, enthusiasm, strength, support and love which saw me through seven and a half years of non-stop study. He was the one who encouraged me to attend university in the first place, and upon my invitation into the Honours program at JCU, the first to state that I would one day become Dr. Claire! I thank him for *everything.*

Finally, I thank my Mum for loving me and providing me with such an enchanting childhood. Although she passed away when I was ten years old those precious years with her shaped the strong, confident and capable woman that I am today. I love her and miss her so much more than words could ever express. Mum, I dedicate this research to you - thank you for showing me how precious early childhood is.

Abstract

This study investigates young children's human figure drawings using the Goodenough-Harris Drawing Test (GHDT) and the Rasch model for measurement. Applying Constructivist and Latent Trait theoretical perspectives the aim, in particular, was to examine the psychometric properties of the test and the human figure drawings in general, and to investigate whether a more culturally, socially and educationally relevant prototype Human Figure Drawing Continuum (HFDC) could be constructed.

The Goodenough-Harris Drawing Test (GHDT) is a non-verbal assessment of young children's levels of intellectual maturity which are inferred from the detail and concepts included in human figure drawings. Originally known as the Goodenough Draw-a-Man Test (GDAMT; Goodenough, 1926a), Harris revised and extended the test to include the Draw-a-Woman (DAW), Self-Portrait-Man (SPM; for male children), and Self-Portrait-Woman (SPW; for female children) sub-tests. However, these sub-tests were added to the GDAMT despite there being no empirical evidence indicating that a single drawing of a man was actually insufficient for the task of inferring children's levels of intellectual development. Similarly, there is currently no empirical evidence that verifies the effectiveness of the tripled data collection load which gathers three human figure drawings from young children (that is, a DAM, DAW and a SPM or SPW drawing). Whilst Goodenough (1926a) and Harris (1963) established superficially the validity and reliability of the GHDT, the test remains unexamined from a modern test theory perspective.

The sample in this study comprised 107 children from years prep to five attending a P-12 school in North Queensland. Data collection consisted of administering the GHDT to the children either individually or in small groups across three phases, each approximately six months apart. The drawings were examined and scored in accordance with the GHDT scoring guides. Results from the Rasch analyses were then used to produce drawing development graphs and common linking plots which displayed visually what the Rasch model detected statistically.

The results indicated that the GHDT was apt for Rasch analysis and the measurement scales produced did not breach the unidimensionality requirement or the other Rasch model expectations. Furthermore, the measurement scales adequately summarised the children's human figure drawings with very few item and case performances detected as misfitting. The mean person measures indicated that the self-portrait (SPM and SPW) sub-tests were better targeted to the sample of children than the DAM or DAW sub-tests. However, the person mean raw scores indicated that children received less credit for their self-portraits than they did for their drawings of men or women. This suggests that the scoring guides, designed to evaluate drawings of men and women, were significantly less sensitive to drawings of children.

Examination of the development of the children's DAM and DAW drawings over the approximate twelve month data collection period indicated that the drawings developed in alignment with Piaget's theory (1956, 1971). However, the Rasch analysis results indicated that drawings of women developed more erratically than those of men.

Interestingly, Rasch output indicated that Harris's inclusion of the DAW and SPM/SPW components contributed little additional information beyond what can be inferred through the original DAM sub-test. Moreover, the common person linking plots revealed that some boys were *dis*advantaged by Harris's inclusion of the DAW sub-test. A common person linking plot including the measures produced by the Rasch analysis of the 50 common items across the DAM and DAW sub-tests suggested that the items exclusive to drawings of men and women were somewhat redundant. Moreover, common person linking plots which compared the researcher's prototype Human Figure Drawing Continuum (HFDC) and each of the GHDT sub-tests indicated that the 45 item HFDC was just as effective in evaluating the young children's human figure drawings as the 217 item GHDT (i.e. DAM, DAW and a SPM/SPW sub-test). Furthermore, the HFDC is considered to be more user-friendly, faster to administer and score, and it empowers young children to self-select the type of human figure drawing they would like to complete (i.e. drawings of men, women, boys or girls). Facilitating young children in selecting their own type of human figure drawing aligns more closely to the foundations of early childhood education in Australia which is child-centred, flexible, and play-based.

In addition to the findings described above, a modified drawing booklet and scoring guide were devised for use with the draft 45 item HFDC. Both the drawing booklet and scoring guide are based upon that used for the GHDT, however, modifications were made to bring these in line with current educational, societal and cultural expectations.

Overall, this research confirmed that young children's human figure drawings can be immensely useful to teachers, psychologists, parents and other interested parties; furthermore, close observation of these drawings can reveal much about the nature of young children's development.

Table of Contents

List of Tables

List of Figures

Chapter One

Introduction

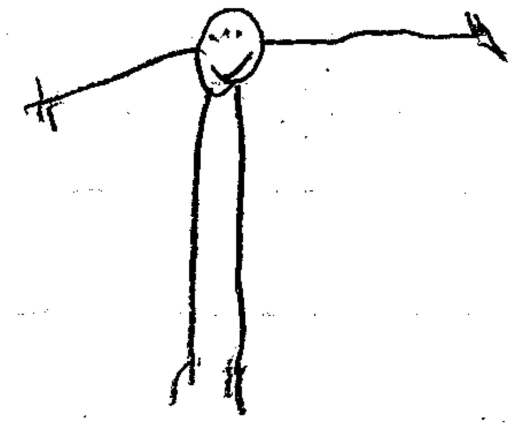
It has been claimed that young children's drawings are more often about "what they know rather than what they see" (Goodenough & Tyler, 1959, p. 316; Crook, 1985; Davis, 1985; Freeman & Cox, 1985; Goodenough, 1926a; Harris, 1963; Ingram, 1985; Light, 1985; Luquet, 1913; Piaget & Inhelder, 1956, 1971; Willats, 1977, 1981). Indeed, long before children can read and write – and understand the functions of reading and writing – they draw. Of all the things that children generally draw, such as animals, trees, flowers, houses, shapes and so on, they most often draw the human figure.

Young children typically enjoy drawing (Cox, 1997, 2005; Goodenough, 1926a; Harris, 1963), and the representations produced are considered delightful and often amusing by adults. At the same time, however, adults are amazed – and sometimes bewildered – by the unusual world depicted in young children's drawings. Seemingly incomprehensible scribbles are labeled as 'Mummy' and 'Daddy'. Later on, basic circles are transformed into smiling suns and flowers, and when the very first human figures emerge they have extraordinarily large heads with arms and legs protruding from the sides and, apparently, no bodies at all. Some parents wonder if this is normal for young children or if there is something wrong developmentally (Cox, 1997). Cox (1997) states that remarkable development can be observed through close examination of young children's drawings.

Many early childhood educators routinely request their young students to draw self-portraits throughout the year. To avoid influencing the nature of these drawings minimal instruction is given. Usually children are asked simply to draw themselves using a pencil and a clean sheet of paper. Generally, these drawings are then filed in individual portfolios or folders, and used by the teaching staff, together with a suite of other information, to make recommendations and judgements about the children's learning and development. Drawings produced earlier in the year are typically less detailed than those produced towards the end of the year. Indeed, most young children's early human figure drawings include little more than a circle representing a head with, perhaps, some arms and legs sprouting from the sides (see Figure 1.1). These first attempts at portraying the human figure are unmistakable. Known as a tadpole or cephalopod, these drawings are possibly the most recognisable of all children's early creations, "so unreal and yet so unmistakeably human – an uncanny reduction to essentials" (Di Leo, 1973, p. 14). Gradually, however, these early sketches become more detailed as hands, fingers, feet, toes, a torso, ears, necks and so on are included.

Figure 1.1

Example of a Cephalopod Drawing Made by a Young Child



Note. Self-portrait drawn by a male, aged 5 years. The arms and legs are attached directly to the head and the nose is between the eyes.

The uniqueness of early childhood is said to be characterized by numerous seemingly illogical behaviours. Not only do young children produce unusual drawings; they often attribute life-like qualities to inanimate objects (i.e. blaming the "Silly step!" for causing them to fall, known as physiognomic perception) (McInerney & McInerney, 2006); believe that the rain comes only to interfere with their playtime (phenomenalistic causality) (McInerney & McInerney, 2006); and think that they have more meat to eat for dinner simply because it has cut into little pieces as opposed to being left whole, amongst many other claims of this ilk. One of the first to investigate this phase of development, and its remarkable characteristics, was psychologist Jean Piaget (1896-1980). His theories have had a profound impact on the way early childhood is viewed, and offered insights into the mental characteristics of young children that are said to be a contributor to their delightful expressions and behaviours (McInerney & McInerney, 2006). Indeed, Piaget's theories are what underpin the epistemology of Constructivism (McInerney & McInerney, 2006).

Adherents to Piagetian theory consider children to be active investigators who construct their own knowledge about the world, and how to interact effectively with it, through the process of adaptation (Bidell & Fisher, 1992; Inhelder & de Caprona, 1987; Piaget, 1953, 1971). The young child is conceptualized as "a being with a distinctive mental structure" (Ginsburg & Opper, 1988, p. 219) which is inherently different to the adult's. Piagetian theory describes children's intellectual development as progressing through an invariant sequence of age-related (not age-dependent) stages: sensorimotor (birth to approximately 2 years); preoperational (approximately 2 to 7 years); concrete operational (approximately 7 to 11 years); and formal operational (approximately 11 years onwards). Each cognitive developmental stage is characterized by the construction of different psychological structures, each new one of which enables "a different type of interaction between the individual and the environment" (Ginsburg & Opper, 1988, p. 23). For early childhood educators, the transition between the preoperational and concrete operational stages is held to be significant.

Piaget's preoperational stage of cognitive development aligns with the time that young children begin to make representations, as children begin to understand objects not only physically, but also symbolically as well (McInerney & McInerney, 2006; Piaget, 1953; Wallach, 1969). Young children begin to experiment with symbolic play and demonstrate their developing understandings of the world through the use of dress-ups, talking on pretend telephones, and using hair brushes as microphones, amongst many other similar behaviours. Piaget believed that children at this stage of cognitive development, although competent and articulate, were less able to use logical mental operations to solve problems or to interpret effectively experiences in the world. The term *preoperational* refers to the idea that the child's internalized schemes have not yet been integrated into an organized, operational system (Flavell, Miller & Miller, 2002; McInerney & McInerney, 2006; Piaget, 1953; 1971). The preoperational stage is primarily a transitional phase "not marked by a stable equilibrium" (Baldwin, 1980, p. 154). This mental disequilibrium is, therefore, reflected in young children's sometimes illogical explanations, actions, and drawings.

Keen to understand further the cognitive developmental aspects of children, Piaget and Inhelder (1956, 1971) examined young children's representations, or drawings, as an extension of their earlier work on cognitive development. In brief, the theory of Piaget and Inhelder (1956) suggests that the 'borderlines' between the child's perception of an object, the child's mental representation of that object, and the child's actual pictorial representation of that object could be problematic for young children when drawing. Only after children have reached a particular operational stage are they then said to be able to transition through these borderlines with continuity and, consequently, produce more logical representations. To quote:

> [T]he transition from perception to mental representation – in other words, to notions which are no longer perceptual but imaginal – [implies] reconstruction of the relationships already grasped at the perceptual level, with functional continuity preserved between the new construction and the earlier perceptual one. (Piaget and Inhelder, 1956, p. 44)

In explaining these borderlines, Piaget and Inhelder (1971) state that "it is possible that at the moment of the gestural and more especially of the graphic copy some sort of anticipatory execution schemes are involved, coming between the perception and the reproduction of the [object]" (p. 15). They went on to say that, "… the graphic image, or reproduction by drawing, is an imitation of the object aiming at reducing its essential characteristics, and starting with and having a fore-image, or executional anticipation" (Piaget & Inhelder, 1971, p.20).

With regard to young children's human figure drawings, Piagetian theory suggests that whilst young children are aware that – as humans – they have multiple physical features, many of these features do not yet transition through to the child's fore-image when the child is producing a human figure representation. Some of the physical features of humans seem to be suppressed in the executional anticipation. The human figure is, consequently, reduced to what the child considers to be the essential characteristics when represented in a drawing – thus resulting in numerous variations of the unmistakable tadpole figure.

Cox (1997) shares a similar understanding of the challenges young children face when producing human figure representations:

> Preschool children certainly know about bodies and tummies and they also know that arms don't stick out of people's heads. It is not that they are ignorant; neither is there any evidence that they imagine real people to look like the ones in their pictures. The problem is to do with the actual process of drawing. Producing a drawing is not like taking a photograph of an object where there is a direct transfer of the image on to the film; it is a much more complicated affair … [t]he young child is … limited in the extent to which she can recall or 'capture' in her mind the features of the figure in her drawing; given this limitation it is likely to be the most defining features that she is able to think of. (p. 12)

With reflection on experience, however, children's cognitive abilities develop and are evidenced in their increasingly more logical thoughts, explanations, and behaviours (McInerney & McInerney, 2006). As young children's thinking transitions towards the concrete operational stage of thought their perception becomes progressively decentred; they are more able to integrate information from a range of sources, and reorganize knowledge into more sophisticated systems. As a result, these older and more experienced children become increasingly able to see others' points of view, consider multiple aspects of a problem at the one time, and produce more logical explanations and drawings. Children transitioning to the concrete operational stage of thought tend to produce drawings that are more logical, detailed, and 'coherent' than those produced by younger children whose thinking is considered to be in the preoperational stage of thought (Cox, 1997; Cox & Parkin, 1986; Piaget & Inhelder, 1956).

Whilst Piaget's ideas are well-known in educational psychology circles, many parents and lay people are unfamiliar with his theories. Accordingly, when young children display behaviours which could be considered characteristic of the preoperational stage of thought, many people dismiss it as one of the bizarre things about early childhood. Indeed, many people attempt to teach young children that it was not the "silly step's" fault that they tripped, or that drawings of human's should include a body, neck, and the like; however, others would assert that it is to no avail. Piaget's theory of cognitive development suggests that it is not until a child develops more organized mental structures that the benefit of this teaching, and other experiences, will be realized (Piaget, 1971).

Others have also confirmed that young children's drawing development, in particular, seems impervious to special tuition or assistance (Barrett & Light, 1976; Clark, 1902; Cox, 1992, 1993; Freeman, 1972, 1980; Freeman & Cox, 1985; Light & Simmons, 1983; Piaget & Inhelder, 1956, 1969). Young children appear to be more concerned with representing *their* actual understanding of an object, rather than a realistic view of the object, when completing drawings (Davis, 1985). Indeed, Vygotsky (1978) and others (Dyson, 1982; Kendrick & McKay, 2002; Ring, 2001) have long considered young children's drawings to be personal graphic speech that externally conceptualizes internal thought or understanding. Similarly, Bruner (1964) views young children's drawings as a form of iconic representation that reflects only the essential, or personally meaningful, features of the experience or object depicted.

This notion of a type of communication underpinning young children's drawings is one of the most fascinating aspects of early childhood development, and something that is valued highly by early childhood professionals. At a time when young children are still developing their vocabularies, and have not yet mastered reading or writing, their drawings – which are said to be more aligned with intellectual development than the amount of drawing experience or teaching children have had – can offer significant insight into their levels of intellectual development.

Florence Goodenough was amongst the first to take advantage of this unique 'communication' aspect of young children's drawings. As a primary school teacher interested in intelligence and conventional IQ tests (Goodenough, 1949), Goodenough examined the potential of young children's drawings as 'evidence' of their knowledge, intellectual maturity, or conceptual understanding[1] of the world. Her extensive research on young children's human figure drawings resulted in the world's first human figure drawing test, the Goodenough Draw-a-Man Test (Goodenough, 1926a). The test was considered useful for assessing the intellectual maturity of young children who were otherwise unsuitable for other forms of testing or assessment (i.e. due to age, disability, hearing or speech impairment, etc.). Goodenough's findings confirmed what many others had already suggested: that young children's drawings can reveal much about the nature of early childhood intellectual development.

It is easy to consider Goodenough as the single most influential researcher on young children's drawing development (Harris, 1963). Her work spawned innumerable

[1] The Collins English Dictionary states that the term *intellect* means "… the capacity for understanding, thinking, and reasoning, as distinct from feeling or wishing" (1998, p. 789). Therefore, the terms intellectual development and intellectual maturity are used in this thesis in association with the terms understanding, knowledge, and conceptualisation.

spin-off tests, assessments, and theories, none of which have had success in expanding the field beyond that already established by Goodenough herself (Di Leo, 1970, 1973; Harris, 1963). Even the revision and extension of the test by Dale Harris – Goodenough's doctoral student – revealed that little could be done to enhance the work already undertaken by Goodenough. The test is now known as the Goodenough-Harris Drawing Test (GHDT).

Undertaken almost 100 years ago, Goodenough's work encompasses the theme of this thesis: that school teachers, parents, psychologists and other interested parties can infer much about child development through the consideration of young children's human figure drawings.

Background to this Research

This researcher encountered the phenomenon of young children's human figure drawings several years ago whilst an under-graduate on second year practicum. It was the middle of the school year, and the School-Based Teacher Educator (SBTE) had suggested a morning activity that included the researcher observing the preschool children drawing. The small group of children sat eagerly at the table, applied the black felt-tip markers to the sheets of white A4 sized paper, and produced many fascinating representations. After a short time, the SBTE then asked the children to draw themselves, each on a blank sheet of A4 paper, and suggested that this researcher observe attentively. For, perhaps, the first time this researcher realized that young children *do* draw differently from older children.

Later that day the SBTE was filing these self-portraits in each child's portfolio, commenting on them as she went. She was comparing and contrasting these latest drawings with ones that the children had created during the first week of preschool at the beginning of the year. The difference between the two sets of drawings, to this researcher, was impressive. These drawings were *prima facie* evidence of some significant child development occurring within a relatively short amount of time. The SBTE explained, “Yes, it’s very important to collect these self-portraits from the children. All preschool teachers do it. Just look at the progress. We’ll get the children to do another one in a few months too. We use them, together with other information, to evaluate the children’s progress. There’s a whole theory on it. You should look it up.” As strange as it might sound, this was the first time that the idea of gathering young children’s drawings for anything other than classroom decoration or caregiver ‘keepsakes’ had been mentioned to this researcher.

Conversations with other early childhood professionals and fellow education under-graduate students indicated that human figure drawings *were* routinely collected in early childhood settings. It seemed that this ‘protocol’ was well understood by practicing early childhood educators, however, the benefits of collecting human figure drawings were not formally discussed in the four year early childhood education degree undertaken by this researcher. A brief search for literature on the topic of young children’s human figure drawings and the links to development revealed a quantity of research by the likes of Goodenough (1926a), Lowenfeld (1939, 1957), and Kellogg, (1967, 1970); however, little had been published recently.

Several years after this second year practicum this researcher considered investigating young children’s human figure drawings for a Doctorate. Discussion with the supervisory team indicated that this might be a worthwhile project. The research could include the renowned Goodenough-Harris Drawing Test – the world’s first human figure drawing test. However, there was a significant question: what new understandings could we possibly uncover about something that had been studied by many over more than 100 years? The Principal Supervisor knew an opportunity when he saw it – and so the idea to apply modern test theory to the world's most famous and widely-used human figure drawing test was born. Having already applied the Rasch model to another early childhood development topic for an Honours thesis (see Maley, 2005), this researcher could see the potential that modern test theory held, for not only advancing the body of knowledge on young children's human figure drawings or developing further the actual drawing instrument, but for all the early childhood teachers, psychologists, mums and dads, and under-graduate education students who value the phenomenon that is young children's human figure drawings.

The Goodenough-Harris Drawing Test

Florence Goodenough (1886-1959), an American teacher, was amongst the first to study young children's drawings, and – like this researcher – was interested principally in those of the human figure. Her study confirmed that the numerous and various changes that can be observed in children's drawings from age to age, as well as within any one age range, link more directly to general intelligence than to any special artistic talent or skill (for children under the age of about ten or twelve years) (Goodenough, 1926a; Goodenough & Tyler, 1959; Harris, 1963; Selfe, 1985). Furthermore, Goodenough demonstrated that the progressive stages of children's drawings have links to the progressive nature of intellectual development. Children who are older and / or more intellectually developed omit essential elements of a drawing (e.g., torso, ears, fingers) less often, display a superior sense of proportion (e.g., head to torso ratio), and have better ideas about the relationships of different parts of a drawing (Goodenough, 1926a; Goodenough & Tyler, 1959; Harris, 1963). For Goodenough, little or no emphasis was placed on how well children draw, or the creativity, neatness and artistic nature included. Rather, she was more concerned with what information was included in the drawings.

As a result of her research, Goodenough (1926a) developed the Draw-a-Man Test (GDAMT). It was the first formal human figure drawing test used to infer a child's level of cognitive (intellectual) development based on the detail included in the drawing. Considered the most basic in conception and general convenience, the GDAMT was acclaimed for its usefulness in assessing accurately the intellectual maturity of children (Harris, 1963).

The GDAMT was revised and extended by Dale B. Harris (1915-2007), and is now known as the Goodenough-Harris Drawing Test (GHDT; Harris, 1963). In addition to revising and extending the original Draw-a-Man test, Harris also included the Draw-a-Woman and experimental Self-Portrait tests. Thus, the GHDT now contains four sub-tests: the revised Draw-a-Man (DAM) sub-test with 73 item scoring points; the new Draw-a-Woman (DAW) sub-test with 71 item scoring points (sharing 49 items with the Draw-a-Man scoring scale); the Self-Portrait-Man sub-test (SPM) which is scored using the DAM scoring guide; and the Self-Portrait-Woman sub-test (SPW) which is scored using the DAW scoring guide. It might be held that these additional sub-tests were added by Harris in response to more recent societal views for equality of outcomes for male and female test takers. Goodenough's original selection of a man as the test subject apparently had little to do with sexism, politics, or views at that time; rather, she deduced that a man could be considered to have less variation in appearance than that of a woman or child (Goodenough, 1926a). The limited variation in appearance of a drawn man allowed Goodenough to focus on the more subtle details of the drawings which she claimed indicated intellectual development. Her extensive observation of thousands of children's human figure drawings over numerous years generally supported these notions.

Goodenough's idea that children's internal understandings might be inferred through external signs, such as their drawings, and counted to reveal insights about developmental levels is not unprecedented. Indeed, the idea of applying mathematical models to human attributes is the very notion that underpins Latent Trait Theory. The validity and reliability of the GHDT has been established by the authors – Goodenough (1926a), Harris (1963) – and others (see Ferrien, 1935; Brill, 1935, 1938; McCarthy, 1944; McCurdy, 1947; McHugh, 1945a, 1945b; Smith, 1937; Williams, 1930; Yepsen, 1929) using techniques derived from Classical Test Theory (CTT).

The traditional approach to psychological quantification used by Goodenough and her colleagues did not, however, have "the theoretical foundations nor the psychometric models that could specify explicitly those substantive qualities of test stimuli or of persons that underlie responses" (Embretson, 1985, p. xi). Indeed, Goodenough (1949) stated that in the late nineteenth and early twentieth centuries "the application of statistical methods to problems of tests and testing was neither well understood nor thoroughly appreciated" (p. 85). And whilst Goodenough and her colleagues applied the most sophisticated methods available at the time, those fall short of the theoretical strength of measurement techniques available today.

Recent developments in psychometrics and computer technology offer new approaches to psychological measurement that surpasses many of the limitations of CTT

(Embretson, 1983, 1983). So whilst CTT has established that the GHDT is relatively valid and reliable, an investigation of the underlying psychometric properties of the test from a modern test theory perspective (e.g. Item Response Theory / Latent Trait Theory) has not yet been published. Therefore, to this researcher's knowledge, the GHDT remains unexamined from a modern test theory perspective.

Most psychometricians would assert that any assessment tool or test should be scientifically sound in terms of its psychometric properties (Bond & Fox, 2007; Chien, 2007; Rasch, 1960, 1980; Streiner, 1995; Wade, 1992). Principally, the GHDT has at least four issues that require further investigation: first, thus far it has merely been assumed that each of the 73 DAM items and 71 DAW items all contribute meaningfully towards the investigation of a single underlying construct of intellectual development. Therefore, the issue of test unidimensionality needs to be examined. Second, the current GHDT scoring procedure allows children's drawings to be scored and compared, but only arbitrarily – the scale is ordinal rather than equal-interval. So whilst deductions can be made about drawings that received more or less credit (i.e. one drawing might have a lower raw score than another suggesting that the latter might be somewhat better than the former), this could be misleading until the relevance of all items to the underlying drawing trait and the equal-interval nature of the current scoring scale have been established. Also, without an investigation of the unidimensionality of the test items, the current scoring system could award credit to aspects included in drawings which might not actually contribute meaningfully to the construct of intellectual development. Third, the difficulty estimates of the DAM and DAW items have also not been genuinely calibrated. The current item hierarchy was established on the out-dated idea that children's development is absolutely orderly and sequential. It has been assumed that each item is sequential, equal in difficulty and therefore equally creditable. Fourth, the need for three drawings – one each of a man, a woman, and a self-portrait – has effectively tripled the data collection load on young children, yet the benefits or disadvantages of this innovation have not been critically examined. Furthermore, the assumed inefficacy of Goodenough's idea that a single drawing of a man would be sufficient for inferring young children's levels of intellectual development has not been established empirically.

In addition to verifying the unidimensionality of the GHDT and determining satisfactory fit statistics of the GHDT items, modern test theory has the potential to revolutionize the GHDT from a series of four sub-tests to something more practical and meaningful. A key model in modern test theory – the Rasch model for measurement (Rasch, 1960) – could assist in constructing a single prototype drawing scale or continuum to evaluate children's human figure drawing development in a fashion similar to the centimetre continuum we use to evaluate the development of physical height. After all, there seems little benefit to testing children, regardless of the method, if their development, or progress, cannot be interpreted meaningfully, gauged, and then tracked over time. Furthermore, the convenience of being able to apply a single drawing continuum scoring scale to young children's human figure drawings cannot be over-estimated. The uniqueness of early childhood almost necessitates the freedom of choice for young children to choose the subject of their human figure drawing (i.e. male or female, adult or child). To be able to then evaluate children's human figure drawings using a single scoring scale which has been statistically verified as effective as an evaluative tool would be significant and practical to early childhood educators in Australia.

Aims of this Research

This research aims to investigate the psychometric properties of the GHDT by administering the test to a sample of young Australian children in years prep to five and applying the Rasch model for measurement to the data collected to reveal empirical insights from a modern test theory perspective. The nature of young children's human figure drawings, and the possible links to intellectual development, will be examined using the unidimensionality principle of the Rasch model. The functionality of each of the 144 items in the GHDT will be verified using Rasch output fit statistics, to validate whether each item contributes to the latent construct under investigation. Furthermore, each item's difficulty estimate will be calibrated and plotted along an equal interval scale suitable for comparison (Bond & Fox, 2007; Wright & Masters, 1982; Wright & Stone, 1979).

Insights into the development of young children's drawings over time will be revealed by utilising the children's measures derived from the Rasch model. This will be achieved through a longitudinal study design whereby three phases of data will be collected, each approximately 6 months apart. The effectiveness of each of the four DAM, DAW, SPM and SPW sub-tests will also be examined to determine whether each one contributes significantly towards the understanding of the construct, or whether one or more is essentially redundant.

The intended result of this research, assuming the successful Rasch analysis of the GHDT, is the development of an equal-interval prototype drawing scale or Human Figure Drawing Continuum (HFDC) useful for evaluating young children's human figure drawings of any sort (i.e. drawings of female or male adults or children) meaningfully, potentially offering early childhood educators insights into young children's drawing development, and their levels of intellectual maturity. These insights, as part of a suite of information, could be useful in making judgements about young children's needs and abilities, and for identifying points for early interventions should these be deemed necessary.

Chapter Two

Literature Review

The idea that young children's drawings provide rich insights into their thinking and levels of intellectual development is longstanding and considered by many to be both self-evident and unequivocal (Anning & Ring, 2004; Di Leo, 1970; Goodenough, 1926a; Harris, 1963; Kellogg, 1967; Lewis & Greene, 1983; Selfe, 1983). The notion has been recognised repeatedly by various researchers over so long that few have attempted to challenge it (Selfe, 1983; 1985). Research into young children's drawings has a long and extensive history.

Historical Overview

Since the beginning of recorded time, humans have expressed themselves through drawings. Many anthropologists have made deductions about how life was thousands of years ago based on their examinations of prehistoric cave drawings. Although some conjecture exists about whether such drawings were religious or ceremonial in nature, one thing is almost certain: cave drawings were intended to convey information at a time when gestural or verbal means of communication were not sufficiently established (Houston, 2004). Indeed, these early drawn symbols were the basis of the first crude forms of written language known as proto-writing, and thus, are considered part of an early organized system of communication (Houston, 2004). Historically, drawing has been considered an innate form of expression for humans (Dyson, 1982, 1993; Fowlkes, 1980).

The motivations behind the early cave drawings mentioned above and young children's drawings today can be seen as comparable. Both involve the use of drawings at a time when other forms of communication are not yet fully mastered (Fowlkes, 1980). When young children are still beginning to grasp the concepts of verbal and written communication, they delight in drawing representations of what they understand about the world. Di Leo (1970) also draws comparisons between children's drawings and prehistoric symbols:

> The response of the eighteen-month infant of today to a crayon and a blank sheet of paper repeats the archaic, universal, abiding, and uniquely human response to a blank wall in a cave or place: the urge to cover it with marks, symbols, graffiti, slogans, designs, [and] pictures. The response is so specifically human, that the mere discovery of such graphic activity is prima facie evidence that Man was there – thirty thousand years ago and yesterday. And recapitulating, as it were, the history of the species, the infant of today will take up scribbling with vigor and gusto, eventually leaving behind him – in the manner of his Paleolithic ancestors – a record of his interests, pleasures, and fears. (p. 9)

Interestingly, very few prehistoric drawings made by children have been discovered. Whilst children have been drawing for centuries, very few of their drawings have been preserved. Perhaps the earliest drawing ever discovered is that apparently created by a child in the Minoan period – about 3000 years ago (see Figure 2.1) (Cox, 1992). Cox (1992) suggests that the reason why so few children's drawings were preserved is related to the fact that early childhood has only relatively recently been considered a significant stage of life in its own right. Interest in early childhood began to surge around the eighteenth century when writers such as Jean Jacques Rousseau (1712-1778) regarded childhood as an important stage of life distinct from adulthood (Cox, 1992; Rousseau, 1976). While early childhood has been valued as an important developmental phase for some time now, the enigmatic and archaic nature of young children's drawings did not attract the attention of researchers until about 100 years ago. Despite this, swift progress in the field was made by Cooke (1885), Ricci (1887), and Luquet (1913) and those that were inspired by their work.

Figure 2.1

One of the Earliest Human Figure Drawings by a Child

Note. Drawn on a slate, this human figure drawing is presumed to be the work of a child from the Minoan period, about 3000 years ago. From "Children's Drawings", by Maureen Cox, 1992, p. 3.

Early Investigators of Young Children's Drawings

Records indicate that research into young children's drawings began in the nineteenth century. The earliest work recorded was by Töpffer (1848), who wrote two chapters on children's drawings that were included in a posthumously published book (as cited in Cox, 2005). These ideas were later adopted by Gautier (1856; as cited in Cox, 2005). However, it was towards the end of the nineteenth century before more detailed descriptions of children's drawings were published.

Ebenezer Cooke (1885) published an article on young children's drawings which was based on his observations of the successive stages of children's drawing development. As a result, he recommended the art education in schools be made to align more directly with the mentality, development, and interests of the child (as cited in Goodenough, 1926a; Koppitz, 1968). Considered to be psychologically perceptive for the time, Cooke's article drew much attention and became an influence on educational practice in schools. Thus, it could be said that from as early as 1885 investigators were suggesting that young children's drawings comprised an intellectual rather than aesthetic foundation (Anning & Ring, 2004; Goodenough, 1926a; Harris, 1963; Koppitz, 1968; Luquet, 1913, 1923; Ricci, 1887).

In 1887, Corrado Ricci, a prominent art critic of that time, published a report titled *The Art of Little Children*. The report described some drawings made by a group of young Italian children which Ricci examined during a summer's holiday. The most significant aspect observed by Ricci seems to be how the young children's drawings included elements that existed, but were not actually seen. These drawings are said to show men visible through the hulls of ships, profiles of men astride horses but with both legs showing, and a bell-ringer visible in a bell-tower (Di Leo, 1973). This collection of children’s drawings is claimed to be one of the earliest recorded (Di Leo, 1970; Goodenough, 1926a; Harris, 1963).

During the late nineteenth century, the growth of interest in child studies and children’s drawings produced a number of further investigations. These reports were chiefly descriptive in nature, only a very few articles contained percentages and/or related statistical calculations. In fact, most reports were annotations of young children’s drawing development as observed by their parents. Cox (1992; 2005) states that among the best of these earlier reports were those by Baldwin (1894), Barnes (1891, 1893, 1894), Brown (1897), Clark (1902), Götze (1898), Herrick (1893), Lukens (1896), Maitland (1895), O’Shea (1897), Perez (1888), Shinn (1897), and Sully (1907, 1908). Generally, these reports concurred with those produced before them in that young children’s drawings seem to express some sort of developmental sequence. After 1900, the focus of study on children drawings adopted a more methodical orientation as interest in the field grew stronger.

Between 1900 and 1915, the scientific interest in children’s drawings experienced a peak as two significant international research projects were undertaken. The first of these was implemented by Lamprecht (1906) who devised a plan based on the method used by Barnes (1891, 1893, 1894) in America. Lamprecht’s plan was for children from all over the world to make drawings under standardized conditions. All drawings were to be sent to a centralized bureau at Leipzig for examination and comparison. Of course, Lamprecht’s ambitious proposal generated great interest, and thousands of drawings from all over the world were sent to him. The collection had representations from almost every nation, including drawings made by traditional African groups. Although this investigation held significant possibilities for the fields of psychology, anthropology, ethnology, and education it, unfortunately, has not been completed. Certain parts of the data have been published in summary by Levinstein (1905), who collaborated with Lamprecht on the project; however, a sufficient comparative study of the entire collection has not appeared.

The second significant international research project during this time was undertaken by Claparède (1907), who proposed a plan for the study of children's drawings that was similar to Lamprecht's, yet with a different end in view. Whereas Lamprecht was interested primarily in possible racial similarities and differences, with particular reference to the theory of recapitulation, Claparède was concerned more so with the developmental aspects of drawing. He proposed a study of the developmental stages in children's drawings, with the aim of determining what relationship, if any, exists between aptitude in drawing and general intellectual ability as indicated by school achievement. Claparède's plan was adopted by Ivanoff (1909) who devised a method of scoring the drawings according to a six-point scale "which took into consideration : (a) sense of proportion, (b) imaginative conception, and (c) technical and artistic value – equal weight being given to each of the three criteria" (Goodenough, 1926a, p. 3). Ivanoff compared the score he deduced from each drawing with teachers' evaluations of the child's general ability, achievement in each school subject, and particular moral and social traits. A positive correlation between and amongst the above elements was found in almost all cases (Goodenough, 1926a; Harris, 1963; Ivanoff, 1909), however, a formal scale of any sort was not devised.

Ivanoff's data were later used by Katzaroff (1910) in his study of the subjects drawn most frequently by children. The results corresponded, somewhat, with those obtained by Maitland (1895). Although, only a rough comparison could be made as

Katzaroff did not consider different aged children separately, and his sample was comprised of children who were somewhat older than those studied by Maitland. Whilst Maitland's study established the human figure as the most frequently drawn subject at all ages up to about ten years, Katzaroff's study ranked it third. Children in the latter study chose to draw "miscellaneous objects" and "houses" first and second-most frequently respectively (Katzaroff, 1910). Other studies have concurred with Maitland's findings in that the subject of most children's drawings is the human figure (Goodenough, 1926a; Harris, 1963; Koppitz, 1968; Pikunas & Carberry, 1961; Schuyten, 1901, 1904, 1907).

Schuyten's (1901, 1904, 1907) studies focused primarily on children's human figure drawings, yet he used a different method from those of other investigators of his time. He aimed to establish, over several years, "a standard of excellence for each age … a series of age norms" (Goodenough 1926a, p. 3). The study required children to draw a man "in whatever way they were accustomed to draw it" (Goodenough, 1926a, p. 3), and do so from memory. Schuyten intended to devise an objective method for rating the drawings by means of minute measurements of each of the separate parts of the human body depicted in the drawing, and by comparing these parts with the actual appearance of a human being. Whilst Schuyten was not successful in completing this project, his intention is the earliest known attempt at devising a purely objective drawing scale (Goodenough, 1926a; Harris, 1963).

Lobsien (1905) adopted Schuyten's method in his study of drawings by school children at Kiel. Again, no precise age norms were established as the project was not seen through to completion. Although, Lobsien's preliminary findings concurred with those of Schuyten in that children's levels of development seemed to correspond directly with the level of detail and use of proportions in their drawings of the human figure.

During 1903-05, Kerschensteiner (1905) commenced a comprehensive and carefully controlled study which involved restructuring the drawing course for schools in Munich. To establish a scientific basis for his research, Kerschensteiner spent approximately two years gathering and examining almost 100,000 drawings produced under standardized conditions by children in Munich and surrounding areas. His analysis of these resulted in three different classifications. As reported by Goodenough, the three classifications were:

> 1. Purely schematic drawings. These correspond to what Verworn (1907) has called the 'ideoplastic stage' in drawing; 2. Drawings in terms of visual appearance – Verworn's 'physioplastic stage'; and 3. Drawings in which the child attempts to give an impression of three-dimensional space. (Goodenough, 1926a, p. 4).

Kerschensteiner's study was one of the earliest attempts to formally identify the different stages of young children's drawing development. The study is detailed in his book entitled, *Die Entwickelung der Zeichnerischen Begabung* (*The Development of Drawing Talent*) (1905). It includes numerous illustrations and tables displaying, by grade and sex separately, the percentages of children whose drawings were categorised within each of the three classes described above.

Despite the accumulation of research on young children's drawings – even at this early stage – the most extensive study on children's drawings undertaken at the time was by Rouma (1913). This comprehensive study was based on six distinct sources of drawings which included: drawings made by children aged from seven to about eleven years over a period of eight months; drawings produced during a ten month period by all students of several primary schools (including one for children with special needs) under the direct supervision of Rouma; weekly half-hour observations of the free drawings made by a class of forty children between the ages of about six and eight years; daily half-hour observations of the drawings produced by a class of thirty children with special needs aged between nine and eleven years of age; daily observations of the drawings of children in a special observation class in a school in Brussels; and lastly, drawings collected by teachers on behalf of Rouma from several kindergartens and primary schools, from which he selected particular children for individual observation (Goodenough, 1926a; Harris ,1963).

Similar to Schuyten's (1901, 1904, 1907), Rouma's conclusions were not based on data analysis using statistical methods, but rather on unusually careful observations of each drawing. His study of the development of human figure drawings produced the following stages as outlined by Goodenough (1926a):

> The preliminary stage.
>
> Adaptation of the hand to the instrument
>
> The child gives a definite name to the incoherent lines which he traces
>
> The child announces in advance that which he intends to represent
>
> The child sees a resemblance between the line obtained by chance and certain objects.
>
> Evolution of the representation of the human figure.
>
> First tentative attempts at representation, similar to the preliminary stages
>
> The 'tadpole' stage

Transitional stage

Complete representation of the human figure as seen in full face

Transitional stage between full face and profile

The profile. (pp. 6-7)

Cox (1992, 1997) states that there were several other early accounts of the drawings made by young children: Brown (1897) published a monograph on four studies, Hogan (1898) reported on the drawings created by her young son, and the Sterns (Stern & Stern, 1910) recorded not only the drawings made by their son but also the circumstances under which these were made. Other accounts by Wilhelm Preyer (1899), and Charles Darwin (1877) were also published. Whilst these descriptions were considered useful at the time, they did not genuinely extend upon what was already known about the development of young children's drawings. Conversely, Luquet's (1913, 1923, 1930) research into young children's drawing development has been the keystone for most other theorists' work.

G. H. Luquet

Georges-Henri Luquet's work is often considered to be the most influential of all the early studies on young children's drawings (1913, 1923). Luquet meticulously preserved every drawing produced by his daughter Simone. Each drawing was numbered and dated, and the circumstance under which it was produced – together with any comments made by Simone, or interesting contextual details. Approximately 1500 drawings were collected and analysed over approximately six years.

Through careful observations of these drawings, Luquet classified four distinct stages of drawing development. The first stage, he noted, commenced at ages 2 to 4 years old approximately, and was described as the '*fortuitous realism'* or scribbling stage. Luquet believed that these first marks were somewhat involuntary as the child was, as yet, unaware that the lines the child was producing could be representative of something (Di Leo, 1970). Luquet suggested that the child, having first produced random and, seemingly, purposeless marks on a page, would come to discover meanings in those scribbles (Luquet, 1913, 1923, as cited in Anning & Ring, 2004). Luquet considered this second stage, '*failed realism*', to be a breakthrough in that the child has realized intentional creation. In Luquet's words, "*A ce moment, it pourrait dire: 'Anch'io son pittore*'" (as cited in Di Leo, 1970, p.16) ("At this moment, he could say: 'I, too, am a painter'"). Aware of the nature of early childhood development, however, Luquet took into consideration the numerous 'regressions' in young children's drawing development. Di Leo (1970) affirms this by stating that Luquet's theory was sensitive to the child experiencing "difficulty in synthesizing, in systematizing the details in to a coherent unity" (p. 16).

Between the ages of 4 and 7 years approximately, children would produce drawings characteristic of the stage named '*realisme intellectuel*' ('intellectual realism'). Luquet considered this third stage to be the apogee of the young child's drawing development (Di Leo, 1970). Di Leo (1970) states that this stage "is indeed the expression of one of the most important phenomena in the metal life of the child" (p. 16) and likened it to Piaget's transition from the preoperational to the concrete operational stage of thought. Di Leo (1970) further postulated that "Luquet has given us the key to an understanding of the strange drawings made between the ages of three and seven" (p. 16).

Children aged between about 9 and 10 years were said to be in Luquet's final '*réalisme visuel*' (visual realism) stage, and thus, more able to achieve realistic representations of the objects they chose to draw. Indeed, "the child has now crossed the bridge into the adult world … from now on development [in drawing] will be a matter of technique … many adults will be incapable of drawing any better than they did at ten or twelve" (Di Leo, 1970, p.16).

It has been suggested that Luquet's stage-based theory on young children's drawing development stemmed from the literature on the education of young children by the likes of Rousseau and Froebel (Anning & Ring, 2004). Similar to other stage-based theories, Luquet's has faced criticism and is prone to over-simplification and misinterpretation. Many who criticize Luquet's stage-based theory are unaware that it refers to 'spontaneous' drawings, not drawings produced from models or under guidance from others. Furthermore, Luquet made attempts to explain that whilst his theory was somewhat stage-based it was not at all 'abrupt'; rather, it was sensitive to the gradual, and sometimes regressive, transitions between these stages. The stages themselves, particularly intellectual realism and visual realism, were considered to be different 'styles' or 'systems' of drawing rather than hierarchical sequences (Jolley, Fenn, & Jones, in press). Indeed, Luquet's theory as a whole has been somewhat neglected as there is no full English translation (Crook, 1985; Read, 1943). Regardless, Luquet's work has influenced innumerable researchers in the field of young children's drawings, and informed other developmental theorists such as Piaget (1956).

Jean Piaget

Jean Piaget (1896-1980) is credited with having influenced the field of contemporary child development more than any other (Berk, 2005). In contrast with the behaviourist views of the time, Piaget did not consider young children's learning and development to be a result of reinforcers, or the absorption of teachings from adults. Rather, his cognitive developmental theory claims that young children "actively construct knowledge as they manipulate and explore their world" (Berk, 2005, p. 20). Influenced by his interest in biology, Piaget's theory is based on the biological concepts of adaptation, assimilation and accommodation (Piaget, 1953, 1971; Miller, 2004). Berk (2005) explains, "just as structures of the body are adapted to fit with the environment, so structures of the mind develop to better fit with, or represent, the external world" (p. 20).

Piaget's four stages of cognitive development and his conservation tasks are well known in educational psychology. Similar to the ways his theory provides some explanation for the illogical utterances and actions of young children when faced with his conservation tasks, Piaget's cognitive developmental theory provides a useful framework for interpreting seemingly illogical aspects of young children's human figure drawings.

Piaget's theory in this area built upon the ideas put forth by Luquet (1923) – that young children progress through three principle stages of development in drawing after producing scribbles. As classified by Luquet (1923), these stages are: 1. Synthetic incapacity (also known as fortuitous realism); 2. Intellectual realism and; 3. Visual realism (Piaget & Inhelder, 1956). The following excerpts from *The Child's Conception of Space* (1956) elucidate the nature of Luquet's stages in the words of Piaget and

Inhelder: Stage I Synthetic Incapacity (from the age of about 4 years) – “the drawing fails to correspond with the perception … the clumsiness of the movements … are unable to carry out the child’s intentions … it seems very likely that [these] drawing exhibit more than mere clumsiness in technique” (p. 47); Stage II Intellectual Realism (approximately 6 or 7 years of age) – “consists in drawing not what the child actually sees of the object (this would be visual realism based on perspective) but ‘everything that is there’ … ‘intellectual realism’ constitutes a type of spatial representation in which Euclidean and projective relationships are just beginning to emerge, though as yet in an inchoate form as regards their interconnections” (p. 50); and Stage III Visual Realism (approximately 9 or 10 years of age) – in which “there finally appears a type of drawing which endeavours to take perspective, proportions and distance into account all at once … comprehensive systems replace empirical constructions” (p. 52).

Together with Inhelder, Piaget (1956, 1971) investigated the development of drawing in young children in two separate studies: the conception of space, and the development of imaginal representation in children. Both of these investigations inform the body of knowledge on the development of drawing in young children, and offer complementary insights on why young children’s representations (or drawings) might seem illogical to adults.

In *The Child’s Conception of Space* (1956), Piaget and Inhelder present their theory of the construction of young children’s human figure drawings:

> A drawing is a representation, which means that it implies the constructions of an image, which is something altogether different from perception itself, and there is no evidence that the spatial relationships of which this image is composed are on the same plane as those

> revealed by the corresponding perception. A child may see the nose above the mouth, but when he tries to conjure up these elements and is no longer really perceiving them, he is liable to reverse their order, not simply from want of skill in drawing of lack of attention, but also and more precisely, from the inadequacy of the instruments of spatial representation which are required to reconstruct the order along the vertical axis … we do indeed have a gap between image and drawing. (p. 47)

In *Mental Imagery in the Child* (1971), Piaget and Inhelder offer a further explanation of their view of the development of drawing in young children:

> According to the hypothesis which we have adopted – that knowledge is an assimilative process – the object [that a child may wish to draw] can be known only by being conceptualized to varying degrees. The image is indeed still the product of an attempt to produce a concrete and even *simili-sensible* copy of the object. But this copy is fundamentally symbolic, since the effective signification is to be found in the concept. (p. xix)

In other words, Piaget and Inhelder consider a representation (or drawing) made by a young child to be the product of an attempt to copy that object – whether from memory or observational drawing. This mental copy is distinct from the child's actual perception of the object. As the young child's cognitive structures are still developing into sophisticated logical systems, their mental representation contains only the symbolic elements of the object that have been sufficiently conceptualized and comprehended by the child. The representation, or drawing, made by a young child is, therefore, a symbol of the important concepts understood about that object (Piaget & Inhelder, 1956, 1971).

With regard to human figure drawings made by young children, Piaget's hypothesis provides a framework for interpreting the classic tadpole or cephalopod (as displayed in Figure 1.1, for example): the young child has conceptualized the head, eyes, nose, arms and legs as being important physical attributes of a human being. That is, the child has assimilated these elements into their understanding of the parts of a human. Thus, these human attributes (head, eyes, nose, arms and legs) become the symbolic elements that are included in a representation of a human. For the drawing displayed in Figure 1.1, it could be said that there is evidence of the child experiencing difficulty with "the instruments of spatial representation" (Piaget & Inhelder, 1956, p. 45), as he has placed the nose *between* the eyes rather then *below* the eyes. In any case, as these and other human attributes – such as a neck, hands, ears, torso, and so on – are conceptualized (i.e., assimilated in the child's mental structures) and organized more sophisticatedly by the young child, they will be progressively included in human figure representations as meaningful symbols of the human figure. Further development of the child's spatial operations will allow these other attributes to be placed appropriately in the human figure drawing.

Building upon Luquet's work, the Piaget and Inhelder research resulted in the elaboration of four key stages (including the scribbling phase) and various sub-stages. These are: Stage 0 (up to the age of about 2 years) – "no purpose or aim can be discerned in the drawings. They are simply scribbles … simple rhythmic movements" (Piaget & Inhelder, 1956, p. 55-57); Stage I (approximately 3 to 4 years of age) –"can be divided into two distinct sub-stages. In the first, sub-stage IA, the scribbles appear to vary … the child produces different types of scribble … at the level of sub-stage IB, however, one can begin to speak of real drawings … there is a correct rendering of topological properties" (pp. 55-56); Stage II – "(starting about the age of 4) is marked by progressive differentiation … during sub-stage IIA shapes are gradually distinguished according to their angles and even their dimensions … during sub-stage IIB the rhombus is drawn correct and the circumscribed figures are gradually mastered" (pp. 56-57); Stage III (approximately 7 to 8 years of age) – is considered the stage when "all the problems are overcome" (p. 57). It is "the point where the movements through which the shape is abstracted could be defined as operational. That is, as being flexible and reversible enough to return constantly to the point of reference on which the subsequent construction is based …. [the] construction is anticipated by a mental image drawn up in advance" (p. 76-77)

Piaget and Inhelder (1956) found that, up until the age of about seven or eight years, young children's abilities to attend to detail and proportion in their representations are still emerging. Indeed, around the age of about seven years children's thinking is said to become decentred and more logical, which is evidence of the concrete operational stage of cognitive development. On the other hand, children whose thinking is inferred to be in the preoperational stage are said to experience difficulty in attending to the multiple elements of an object at the one time (McInerney & McInerney, 2006; Miller, 2004; Piaget, 1971). Similarly, Cox (1992) stated that young drawers *do* appear to experience difficulty in organizing their spatial knowledge (i.e. *where* to put marks on the paper) at the same time that they try to organize their conceptual knowledge (i.e. *what* to put on the paper). This seems to align with Piaget's idea that children in the preoperational stage of thought are not yet able to focus on more than one element of a problem at a time; thus, they construct somewhat illogical representations as a result of the mental disequilibrium (Piaget, 1971; Piaget & Inhelder, 1971).

Piaget was not the only researcher to develop Luquet's ideas further. Many others also used Luquet's drawing development theory as a basis for their own work on young children's drawings.

Other Drawing Theorists

Viktor Lowenfeld (1903-1960), an Austrian-American professor of art at the Pennsylvania State University, developed the stage-based concept further by suggesting that there were several more distinct stages of development in young children's drawings. He outlined six stages of drawing development: scribbling (ages 2 to 4 years); pre-schematic (ages 4 to 7 years); schematic (ages 7 to 9 years); dawning realism (ages 9 to 11 years); pseudo realism (ages 11 to 13 years); and period of decision/crisis of adolescence (adolescence and beyond); (Lowenfeld, 1939, 1957 as cited in Anning & Ring, 2004).

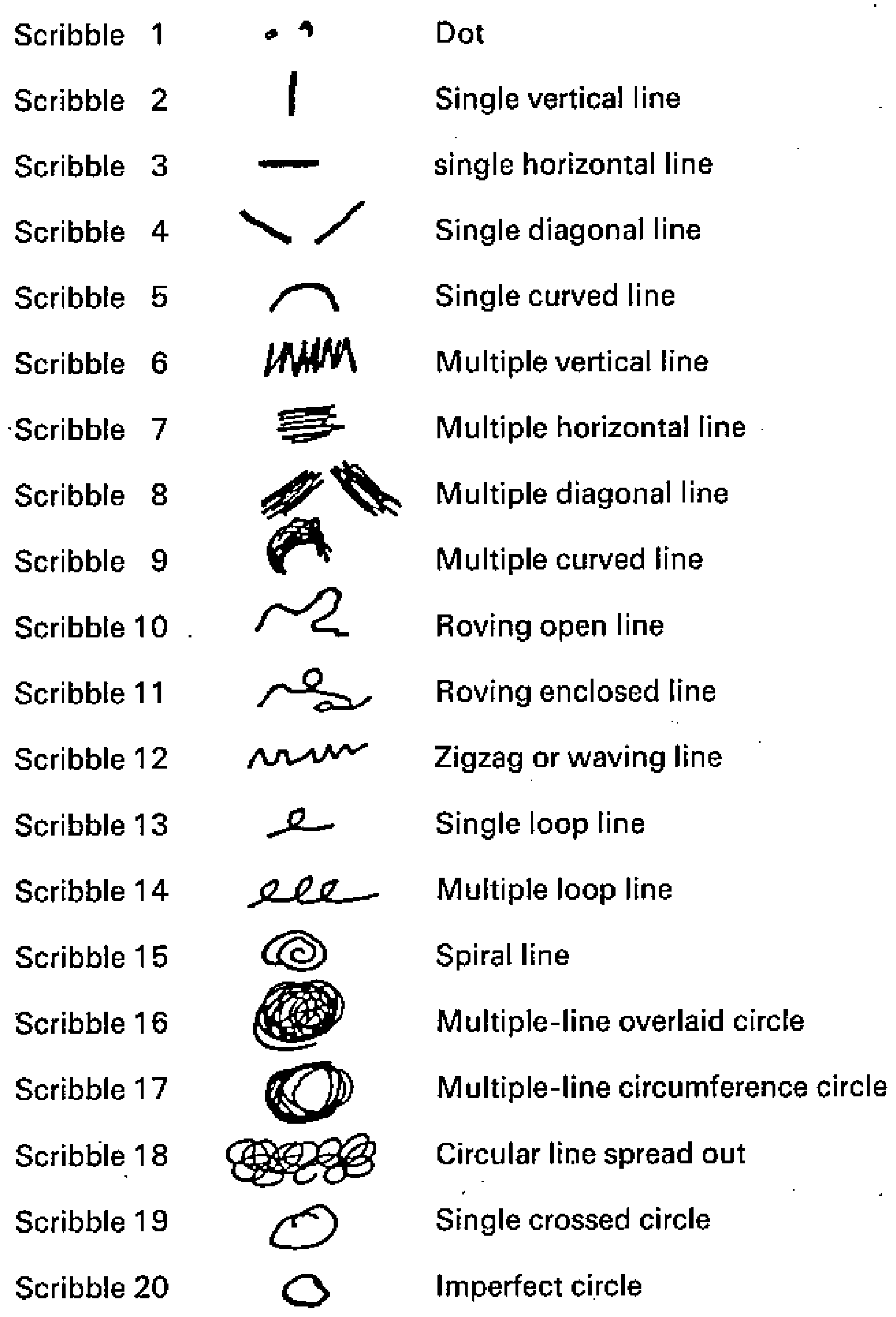
Although he was interested in psychoanalysis, Lowenfeld was an art educator rather than a developmental psychologist. Consequently, he saw drawing as a means of release of tension and feelings, and a quest for realism and creative design. As a result, Lowenfeld's theory had limited relevance to the field of early childhood development. In contrast, Piaget viewed drawing as a principle developmental aspect of early childhood.

Similar to Piaget, Kellogg (1970) was a proponent to the notion of universal, stage-based development in young children's drawings. As a result of her extensive study of children's drawings, Kellogg proposed that there are 20 different types of basic scribbles (see Figure 2.2). Whilst not all children produce all types of scribbles, and many children go on to combine the different types, Kellogg suggested that there was an orderly progression from basic scribbling through to a series of more complex types of drawings (Cox, 1992; Kellogg, 1970; Kellogg & O'Dell, 1967). At around the age of about 4 years, Kellogg states that children begin to use the 20 basic scribble types as a basis to represent the human figure in their drawings (see Figure 2.3). Similar to findings of many other researchers, Kellogg found that the human figure is the most frequently drawn subject in young children's drawings (Kellogg & O'Dell, 1967).

Despite contributing much to the field of young children's drawing development, there are limitations to Kellogg's theory. Golomb (1981, 1992) claims that only around 4% of children actually constructed the sun figures which are said to precede the first human figure drawings made by children (see Figure 2.3). What's more, the children who produced sun figures did so at the same age that they produced their first human figure drawings (Cox, 1992; Golomb, 1981). Other studies have also indicated that children do not always demonstrate an elaborate scribbling phase (Alland, 1983; Cox, 1992; Gardner, 1980).

Figure 2.2

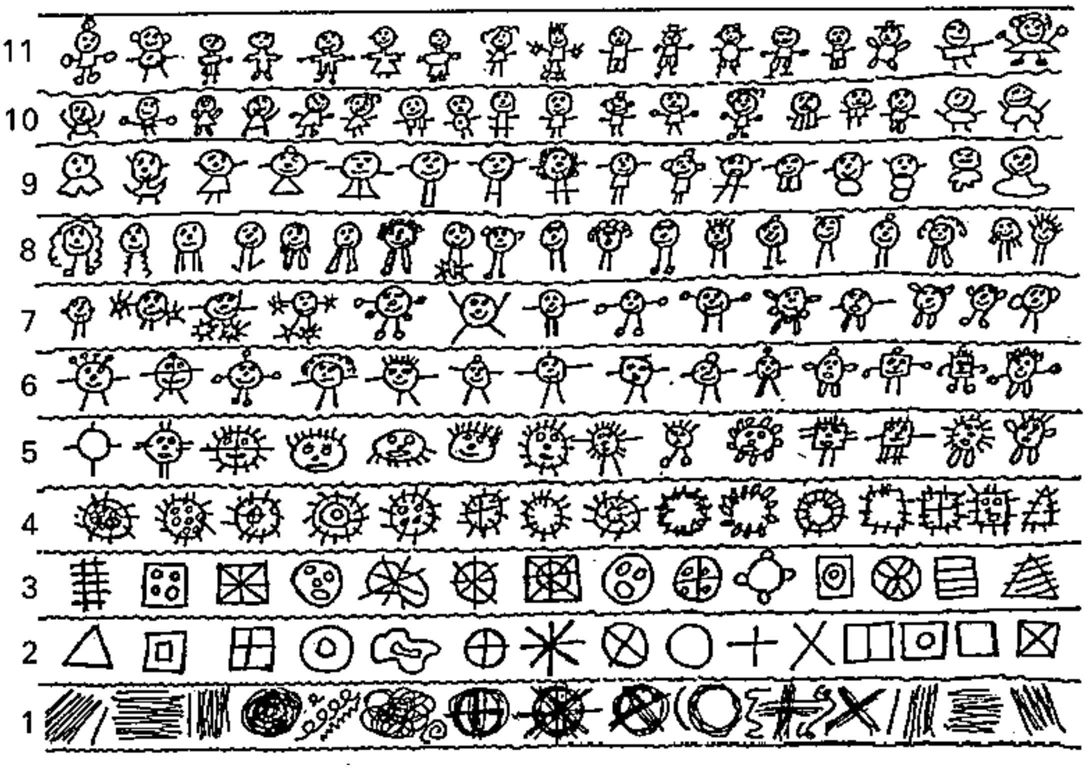
Kellogg's Identification of Twenty Basic Scribbles



Note. From "Children's Drawings", by Maureen Cox, 1992, p. 14.

Figure 2.3

Kellogg's Theory of the Development of Human Figure Drawings



Note. From "Children's Drawings", by Maureen Cox, 1992, p. 15.

The somewhat universal nature of the development of young children's human figure drawings has been espoused by many (Di Leo, 1970, 1973; Goodenough, 1926; Harris, 1963; Kellogg, 1970; Kellogg & O'Dell, 1967; Luquet, 1913, 1923; Piaget & Inhelder, 1956, 1971). In her book, *Measurement of Intelligence by Drawings* (1926), Goodenough explains that children who have had no experience in drawing require only a very small amount of practice before their drawings come to resemble those of other children in their age range and ability level. It would seem that if the development of drawing were based on something else, such as the amount of practice, or level of special skill or creative talent, a child with no experience in drawing would not be able to produce a representation comparable to those of their age range peers. Recognising the potential of the development of drawing in young children as a possible tool for inferring the overall development of the child, Goodenough set about researching children's representations.

Florence Louise Goodenough

In the field of young children's human figure drawings, the work undertaken by Florence Goodenough (1886-1959) is considered by many to be the most prominent. The youngest of nine children, Goodenough's first degree was a Bachelor of Pedagogy completed in 1908. Whilst completing a Bachelor of Science and, afterwards, a Masters degree Goodenough worked as the Director of Research in the Rutherford and Perth Amboy schools, New Jersey. This role is said to be comparable to that of a school psychologist nowadays (Thompson, 1990). In any case, it was during this posting that she collected most of the data for her research on young children's drawings.

In 1921, Goodenough began working with Lewis Terman at Stanford University. Terman was developing further the Stanford-Binet Intelligence Quotient test for children at the time, and Goodenough assisted him in his research (Thompson, 1990). Goodenough's work as Terman's chief field psychologist, and then later his chief research psychologist, contributed to her earning her PhD in 1924.

After completing her doctorate from Stanford University, Goodenough moved to the University of Minnesota. Thereafter she developed the Goodenough Draw-a-Man Test (GDAMT) and published her research in the book entitled *Measurement of Intelligence by Drawings* (1926a). The GDAMT is considered to be Goodenough's single most significant contribution to the field of developmental psychology. The GDAMT, unlike other nonverbal intelligence tests of the time, had high validity and reliability, was easy to administer, and correlated well with other intelligence tests of the period (Goodenough, 1926a; Harris, 1963; Thompson, 1990). It was so highly regarded that even 20 years after its inception, it was ranked the third most frequently used assessment in clinical psychology (Thompson, 1990). Dunn (1972) stated that given that the GDAMT was developed for use with young children, this achievement was particularly notable.

After developing the GDAMT, Goodenough investigated more traditional, verbal tests of intelligence for children. Goodenough's second book (1928a), *The Kuhlman-Binet Tests for Children of Preschool Age: A Critical Study and Evaluation*, investigated the 1922 Kuhlman revision of the Binet Intelligence Quotient test for preschool children. Afterward, she developed her own version of the Stanford-Binet Intelligence Quotient test which she extended downwards to include young children. This became known as the Minnesota Preschool Scale which was useful for inferring the abilities of babies from 18 months to children 6 years of age (Goodenough, Foster & Van Wagenen, 1932).

Shortly after this, Goodenough showcased her extensive knowledge and published *Mental Testing* (1949) a comprehensive text on the history, development, methods, applications and limitations of testing. Acknowledging both the positives and the negatives of testing Goodenough (1949) stated:

> Few if any other fields of psychology have aroused so widespread an interest, not only among psychologists but within such related fields as psychiatry, anthropology, sociology, and education and in certain branches of the law, as well as among the general public. It was perhaps

> inevitable that so rapid and widespread a growth of interest has led to the promulgation of erroneous as well as sound theories, and that the use of the new methods by many enthusiastic but poorly trained persons has not always worked out to the advantage of the tested individuals or of society. Young sciences, like young learners, must try out many pathways in order to discover the true course, and in both cases the early stages of learning are likely to be marked by overhasty conclusions and wishful thinking. (p. vii)

Amongst Goodenough's other contributions to the field of educational psychology were the data collection methods of time sampling and event sampling (Goodenough, 1928b). Time sampling involves setting aside a series of an agreed amount of time for observation (i.e., 5 to 15 minutes), and targeting a behaviour to detect during this period. Completing a series of time samples produces a record of the frequency of a particular behaviour over time. Event sampling, on the other hand, involves waiting for a particular behaviour to occur. If and when the targeted event takes place, the observer then commences formal observation of the behaviour and takes notes as necessary. Both of these observation techniques are considered of utmost practical importance to the fields of education and psychology (Wright, 1960), and in particular, are used extensively in early childhood settings (Bentzen, 2000, 2009).

Goodenough later turned her attention to the study of social and emotional development in young children. She was interested in the idea proposed by Watson (1926), that a newborn displays the three basic emotions of rage, fear and love, and that other emotions stem from these. Goodenough's 1931 report, *Anger in Young Children*, came about from her use of time and event sampling targeting outbursts of anger displayed by young children. She was interested in extending Watson's earlier work by describing the development of emotions beyond infancy (Thompson, 1990).

Despite Goodenough's ventures into other areas of developmental psychology, her work in the field of young children's drawing development is regarded as her most noteworthy. The Goodenough Draw-a-Man Test was considered to be the closest researchers had come to the ideal of a non-verbal, culture-free intelligence test (Di Leo, 1973).

The Goodenough Draw-a-Man Test

Goodenough's accomplishments are considered to have superseded the previous research concerning children's drawings; her book *Measurement of Intelligence by Drawings* (1926a) stands as a classic (Koppitz, 1968). Despite the voluminous literature produced before Goodenough's study, her investigation into children's drawings resulted in original insight and the field taking a new direction. Her research aligned with what others had already suggested - that children's drawings comprise a large intellectual component. In addition, however, Goodenough ascertained that this component links theoretically to the psychometric study of intelligence (Goodenough, 1926a; Harris, 1963).

Goodenough's research aims stemmed from other researchers' previous attempts to classify children's drawings. She was dissatisfied with the small number of classification categories, the informal analysis of drawings, and subjective nature of the procedures as a whole. Whilst distinctions could be made amongst and between groups of drawings using these methods, Goodenough recognized that the individual ratings obtained held little significance. In developing a more suitable scale, Goodenough observed the inadequacies of previous investigators and modified her approach. In her review of the body of knowledge on young children's drawings – most of which is discussed above – Goodenough contended that there were several conclusions that could reasonably be drawn. Some of these were:

> …in young children a close relationship is apparent between concept development as shown in drawing, and general intelligence … drawing, to the child, is primarily a language, a form of expression, rather than a means of creating beauty … in the beginning the child draws what [he or she] knows, rather than what he sees … later on he reaches a stage in which he attempts to draw objects as he sees them. The transition from the first stage to the second one is a gradual and continuous process … The child exaggerates the size of items which seem interesting or important; other parts are minimized or omitted … The order of development in drawing is remarkably constant, even among children of very different social antecedents … this is especially true as regards the human figure, probably because of its universal familiarity … in drawing objects placed before them young children pay little or no attention to the model. Their drawings from the object are not likely to differ in any important respect from their memory drawings … up to about the age of ten years children draw the human figure in preference to any other subject. (Goodenough, 1926a, pp.12-13)

Based on these conclusions, Goodenough set about devising a test that could be used to infer young children's levels of intellectual development. Whilst it would have been unquestionably beneficial to allow each child to select his/her own subject to draw, Goodenough knew from the outset that this was not possible. The numerous and various subjects that children might have chosen to draw would each have required systematic study to determine their relative difficulty levels (Goodenough, 1926a). Furthermore, it was considered impossible to establish with certainty whether a good drawing of an easy subject or a poor drawing of a more difficult subject was an indication of higher ability (Goodenough, 1926a; Harris, 1963). Consequently, the following points derived directly from Goodenough's *Measurement of Intelligence by Drawings* (1926) were of paramount importance in her selection of a universal subject:

> It must be something with which all children are equally familiar. That is, either the situation presented must be an entirely new one, or else all the [participants] must have had as nearly as possible equal opportunity to become familiar with it. For very little children, at least, the latter circumstance is probably the more favourable, since it is less likely to produce mental confusion and has the additional advantage of measuring the learning factor as shown by present accomplishment. It must present as little variability in its essential characteristics as possible. It must be simple in its general outline, so that even very little children will be able to attempt it, yet sufficiently complicated in its detail to tax the abilities of an adult. In order that a proper spirit may be maintained among the children taking the test, the subject chosen must be one of universal interest and appeal. (Goodenough, 1926a, p. 15-16)

Only one subject seemed to satisfy all of the above requirements – the human figure. The general uniformity of men's clothing indicated that a man was especially apt for the drawing test, more so than a woman or a child (Goodenough, 1926a). Thus, the original Draw-a-Man Test came into being.

Experimental Basis of the Goodenough Draw-a-Man Test

Goodenough began the foundations of her study in 1920 at a primary school in Perth Amboy, New Jersey. With the cooperation of the primary grades supervisor, Miss Alice Mehleis and the teachers under her supervision, almost 4000 human figure drawings were obtained from children in the kindergarten and years one to four. All drawings were completed under the standardized Draw-a-Man Test conditions as devised by Goodenough (1926a). A preliminary study was conducted with a sub-group of 100 drawings taken from the total of 4000. Goodenough selected the drawings based on an age-grade classification – 10 drawings each from 10 different age ranges (from 4 years to 9 years 11 months, kindergarten through to year 4). Beyond this, no other limiting factors were applied.

The 100 selected drawings were spread out in order by age (youngest to oldest). Goodenough thoroughly examined each of the drawings looking, systematically, for aspects that distinguished the drawings of older children to that of younger (hence, less developed) children. Specifically, this was “to determine what characteristic changes take place in children’s drawings with increasing age and intellectual development” (Goodenough, 1926a, p.17). Goodenough made no assumptions as to the probable nature of the changes, and artistic aspects of the differences observed were disregarded entirely; “the only point considered was that of comparative differences” (Goodenough, 1926a, p. 17). Also, she engaged several other people to examine the drawings and had them note differences as objectively as possible. As a result of this preliminary analysis, a crude scoring guide of approximately forty points was constructed.

In her book, *Measurement of Intelligence by Drawings* (1926a), Goodenough defines a point as a single unit of the Draw-a-Man scale:

> …it may be based upon the presence or absence of a specified element, upon the method of representation of a given quantitative or spatial relationship, upon eye-hand coordination, or several of these characteristics may have been combined to form a single "point". (p. 17)

Whilst Goodenough did not describe this scale in detail, it became the basis for the Draw-a-Man scale used today. Each of the selected 100 drawings was scored according to it, with each point being recorded and curves being plotted, displaying the number of successes at each age level.

Point Validation

Goodenough used a three-fold criterion for establishing the validity of each point. The requirements were: "(1) a regular and (2) fairly rapid increase in the percentage of children succeeding with the point at successive ages, and (3) a clear differentiation between the performances of children who were of the same age but in different school grades" (Goodenough, 1926a, p. 17-18). Measures were taken to ensure that the criteria for scoring each point was defined in a broad, objective and clear manner so that all situations likely to arise with the scoring of the point were covered.

As to be expected, many of the points on the preliminary scale had to be changed (Goodenough, 1926a). Thus, the scale was revised with other points that had suggested themselves added, and the drawings re-scored according to the new scale. The resultant developmental curves were relatively satisfactory in terms of the small number of drawings under consideration. To test further the limits of this scale, all the drawings from one school were scored in line with it (approximately 800 in total), and new curves were plotted. The findings of this analysis contributed to the scale being revised again and extended.

In total, Goodenough revised and extended the Draw-a-Man scale five times, following the same procedure in each case. She used a different set of drawings for each revision so that the error of validating a point by the means of the same drawings from which it was derived, was avoided. The resultant scale consisted of 51 points, with each conforming satisfactorily to the requirements set (and outlined above) (Goodenough, 1926a).

Goodenough stated that it was unnecessary to present the four preliminary scales. She claimed that the points were essentially the same in all; however, methods and scoring criterion had undergone some modification at each stage. This is particularly the case with regard to several of the points relating to clothing and proportion. To elucidate this further, an excerpt is provided below from Goodenough's (1926a) *Measurement of Intelligence by Drawings.* Regarding point 12a:

> Scale 1. Head smaller than trunk.
>
> Objection: too crude a measure. The curve showed only a slight increase in the percentage of successes at different ages.
>
> Scale 2. Head length not less than one tenth or greater than one fifth of the total body length.

Objection: In many instances there appears to be a negative correlation between the size of the head and the length of the legs. Sometimes this fact is determined by the size of the sheet of paper – the child who makes a very large head and trunk being obliged to make the legs very short in order to get them on the page. Conversely, the child who makes a small head and trunk may thriftily fill up the remaining space with a pair of abnormally drawn-out legs. To compare the several parts with the total often has the effect of penalizing the child twice for a single disproportionate element.

Scale 3. Size of head "not grossly disproportionate to the remainder of the drawing".
Objection: Too indefinite a ruling, leading to subjective errors in scoring. The method was tried largely to see whether the point was worth using at all, as both previous methods had resulted in very unsatisfactory curves. The results obtained in this way showed clearly that the point was one which should be retained in the scale; accordingly another method of scoring was tried.

Scale 4. Both vertical and horizontal measurements of head less than the corresponding measurements of the trunk.
Objection: This method applies very well to full-face drawings, but is not satisfactory with profiles.

> Scale 5. Area of head not more than one half or less than one tenth that of the trunk. This is the method finally adopted. (pp. 19-20)

Evidently, the method employed to validate the scale points was rather "cut and try" (Goodenough, 1926a, p. 20). Any point displaying clear separation of different age and grade groups, was retained. Any point displaying differentiating value was noted, and a new method of scoring devised. When a satisfactory scoring method could not be found, the point was rejected. Some typical examples of rejected points include:

> Teeth shown
>
> Up to about the age of seven, the curve shows a regular increase in the percentage of children who draw the teeth. After this age there is an equally marked decrease, a fact which renders the point useless.
>
> Attempt to show colour by shading
>
> This varies according to the hardness of the lead and the condition of the point of the pencil used.
>
> Attempt to represent movement, as walking or running.
>
> This point was rejected only after several attempts to score it had been made. There is little doubt as to its being, in some degree, a valid indication of intellectual maturity. The difficulty lies in differentiating between real attempts to show movement and mere bad coordination. As a result of poor coordination the drawing may seem to show one leg being raised, as in walking, when nothing of the sort was intended by the child. With more mature drawings it is usually possible to make the

> distinction, but with those of the younger children it is often difficult, if not impossible, to do so. (Goodenough, 1926, p. 20)

Other points, in addition to those listed above, were rejected for similar reasons. Goodenough recognized, however, that further trials might reveal methods of scoring which would override the objections made; although, at that time such methods had not yet been found.

Goodenough Draw-a-Man Test Procedure and Scoring Scale

Goodenough's *Measurement of Intelligence by Drawings* (1926a) details the test administration procedure and scoring guide for the GDAMT. The test requires only the use of a pencil and a test booklet, and can be administered by a classroom teacher. Goodenough explains that children's drawings fall into either one of two classes: Class A – is where the "subject of the drawing cannot be recognised … the drawing consists of merely aimless, uncontrolled scribbling … the total possible score is either 0 or 1" (Goodenough, 1926a, p. 90); Class B – "includes all drawings which can recognised as attempts to represent the human figure, no matter how crude they may be … a credit of 1 is allowed for each point scored, and no half credit are given" (p. 91). Therefore, the GDAMT scoring scale is dichotomous in nature (i.e. no/yes, absent/present, 0/1). An item is deemed either: present and adhering to the scoring criteria (1 credit awarded); or present and not adhering to the scoring criteria, or absent (0 credit awarded) (see Goodenough, 1926a).

Reliability and Validity of the Goodenough Draw-a-Man Test

It is well-known that the utility of any test is dependent on its reliability and validity (Harris, 1963). Goodenough (1926a, 1926b) and others (see Berrien, 1935; Brill, 1935, 1938; Harris, 1963; McCarthy, 1944; McCurdy, 1947; McHugh, 1945a, 1945b; Smith, 1937; Williams, 1930; Yepsen, 1929) established – using the best available methods of the time – the reliability and validity of the Goodenough Draw-a-Man Test. The reliability of the GDAMT has been determined to be between r = .68 and r = .94 using the test-retest method (see Brill, 1935; Goodenough, 1926b; Selfe, 1985).

Test validity is evidenced by the GDAMT's strong correlation with other, more conventional intelligence tests such as the Standford-Binet IQ test (r = .74) (Goodenough, 1926a). Goodenough also found strong correlations between GDAMT scores and children's grade placement, indicating the test had distinct value in predicting future school success. Other researchers found correlations between the GDAMT and the Stanford-Binet of r = .65 (Williams, 1935), r = .60 (Yepsen, 1929), and r = .72 (McElwee, 1932). It should be noted, however, that other researchers did not always adhere closely to Goodenough's test administration instructions or recommended age ranges; resulting in, perhaps, less than otherwise expected reliability estimates (Harris, 1963). Harris (1963) also stated that the validity coefficients, whilst uniformly positive, vary depending on the age of the participants, the age range used in the sample, and the other intelligence test used as a criterion. In any case, Harris (1963) affirmed that the correlation values "show that the test measures intellectual more successfully than aesthetic or personality factors" (p. 36).

There has been much contention concerning the issue of the validity of the GDAMT with regards to the effect of art training on test performance. Harris (1963) also investigated this issue in the study mentioned above by including children who were part of "outstanding programs of art education" (p. 92). Harris found that the results reiterated Goodenough's original conclusion that: "repeated comparisons of the work of children' who have had this type of [elementary school art] training, with that of children from schools in which no drawing at all is taught in the primary grades, have failed to show any consistent differences between the performance of the two groups in drawing the human figure" (Goodenough, 1926a, pp53-54). An extensive study by Phatak (1959) on the relationship between the artistic merit of children's human figures drawings and the score obtained by the GDAMT method also yielded similar results. Phatak's study was further evidence that the GDAMT scoring method and the drawings of human figures by young children is independent of artistic ability.

Harris (1963) investigated inter-rater reliability by having a classroom teacher, and himself, independently administer and score the drawings collected from a group of school children. He found that the test administrator and scorer had very little influence on the scores achieved by the children, or the statistics yielded (Harris, 1963).

With many others having established the reliability and validity of the GDAMT, interest in the field grew immensely. Inspired by Goodenough's use of young children's drawings for inferring intellectual maturity, others began to use drawings for other projective tests.

Projective Tests Based on the GDAMT

Some researchers altered Goodenough's original method and applied it to their own endeavors; for example, to study personality, adjustment problems, social / emotional problems, delinquency and character defects (Anning & Ring, 2004; Di Leo, 1970, 1973; Harris, 1963; Koppitz, 1968). The detail included, or excluded, and the level of overall coherence in the human figure drawings made by children were said to have links to particular personality traits and/or disorders (e.g., Buck, 1948a, 1948b; Koppitz, 1968; Machover, 1949).

In terms of theory, for Harris (1963) and Goodenough (1926a), this research was less scientifically analytical and more intuitive impressionism. These numerous projective tests could be said to have exploited the drawing assessment technique as interest in the field of developmental psychology grew and the number of novice test administrators increased. Indeed, many expressed concern over the reliability and validity of these spin-off projective personality tests and, furthermore, the unauthorized use of Goodenough's DAMT for purposes for which it was not intended (L. Ashton, personal communication, 2008; Ashton, 1999; Crook, 1985). Some felt that the situation was similar to the exploitation of the Rorschach Ink Blot Test (Rorschach, 1942) whereby misunderstanding and misuse of the test resulted in many adults being inaccurately labeled as mentally inept. Similarly, misunderstanding and misuse of the GDAMT was at risk of classifying young children as less than competent if one of their drawings was not looked upon favorably by a teacher or psychologist (L. Ashton, personal communication, 2008). Indeed, Harris (1963) commented on the situation by stating that "all too many investigators appear to have neglected Goodenough's warning" (p. 41) and followed this statement by reiterating Goodenough's original caution from *Measurement of Intelligence by Drawings* (1926a):

> The facts herein reported are by no means intended to convey the impression that the writer is able to diagnose psychopathic tendencies in children by means of a drawing. Certainly no such claim is justified.

> It is believed, however, that by an investigation carried out along the lines which have been indicated a method of scoring might be derived which would throw new light upon eccentricities of mental functioning during childhood. (Goodenough, 1926a, p. 66)

In any case, other investigators continued to apply the GDAMT to their own endeavours, or develop a 'spin off' test of some description, in an effort to capitalize on the surging interest in young children's drawings. These derivatives did little to uphold the reputation of the GDAMT as one of many useful tools for inferring the intellectual developmental levels of young children.

It should be noted that the GDAMT and Rorschach Ink Blot Test were not the only instruments to endure misuse and misinterpretation. Generally, all tests and assessments were prone to inaccurate administration, misunderstanding or overgeneralization of the results (Goodenough, 1949; Gould, 1996). Goodenough was well-aware of the limitations and hazards of testing. In *Mental Testing* (1949) she warned that the surge in interest in intelligence and mental testing "has led to the promulgation of erroneous as well as sound theories, and … the use of the new methods by many enthusiastic but poorly trained persons has not always worked out to the advantage of the tested individuals or of society" (p. vii). Goodenough was acutely aware of the limitations of testing, as well as the implications of scoring and ranking test participants. Correspondingly, she was adamant that the GDAMT was but one of many tools that should be used to comprise a *suite* of information on any individual (Goodenough, 1926a). As one might expect, Goodenough was not supportive of the spin-off projective tests that the growth of interest in psychology and young children's drawings brought about (Harris, 1963).

John Buck

One of the first to alter Goodenough's DAMT concept was John Buck (1906-1983). After being trained in the administration of the Stanford-Binet IQ test for his employment as secretary to John H. Bell M.D. at the State Colony for Epiletics and Feebleminded (later renamed Central Virginia Training Centre), Buck became interested in numerous examination, interview and testing techniques (Rowe, Crews & Finger, 1993). After noticing the benefits of drawing or doodling with his patients at the hospital Buck decided to investigate the literature on drawings in the field of psychology, and realized that there were others espousing that drawing might be useful as a diagnostic aid (Annual Report, 1939).

Buck developed the House-Tree-Person Test (H-T-P), a technique designed to yield information about the sensitivity, maturity, and personality of a person as inferred through their drawings of a house, tree and person (Buck, 1948a, 1948b). The H-T-P was designed as a two-phase test with a non-verbal and then a verbal segment. The non-verbal aspect involved the drawing of the actual house, tree and person; the verbal aspect involved the test administrator asking a series of pre-determined questions about each of the drawings. The test has also been used to detect abuse (Blain, Bergner, Lewis & Goldstein, 1981) and assess levels of self-esteem (Groth-Marnat & Roberts, 1998) in children and adults.

Elizabeth Munsterberg Koppitz

Another test said to be inspired by Goodenough's DAMT was Elizabeth Koppitz's (1919-1983) Human Figure Drawing Test (HFDT). Koppitz details her development of the test in her book *Psychological Evaluation of Children's Human Figure Drawings*

(1968). The test was developed by way of her need for "an integrated, systematic way to interpret HFDs [as] none of the existing methods seemed to tap the full richness of HFDs" (Koppitz, 1968, p. 3). Similar to the GHDT, it can be administered individually or as a group test and requires the drawing of a human figure. In contrast to the GHDT, the HFDT requires only *one* drawing, the sex of which is selected by the child as the administration procedure asks merely for the drawing of "a whole person" (Koppitz, 1968, p. 5). Koppitz (1968) asserts that such non-specific instruction leads the child to "look into himself and into his own feelings when trying to capture the essence of 'a person'" (p. 6) and thus, construct a portrait whereby the inner self and attitude are portrayed.

Completed drawings are analyzed through both objective developmental and projective interpretative methods. Scores are given for items related to the child's age and level of maturation – known as *Developmental Items*; and other items thought to be related to the child's attitudes and concerns – known as *Emotional Items* (Koppitz, 1968, p. 7). Despite Koppitz's well-meaning intentions, the validity and reliability of the HFDT remains somewhat questionable as the clinical interpretation of the children's human figure drawings is based on Koppitz's (1968) "own experience and intuition" (p. 8). Nonetheless, several of her findings correspond with the results of other research projects on young children's human figure drawings. In concurrence with Goodenough (1926a), Harris (1963) and Machover (1949), Koppitz found that the drawings made by primary school-aged girls tended to be more detailed than those of primary school-aged boys.

Like Koppitz, Di Leo (1973) also recognised the potential of young children's drawings to disciplines other than the study of intellectual development. Di Leo (1973) stated that while the value of drawings as expressions of intellectual development should be emphasized, "neither Goodenough nor Harris has sufficiently recognised their affective and perceptuo-motor aspects, since these do not lend themselves to the same degree of objectivity and measurement as does their cognitive aspect" (p. 30).

Joseph H. Di Leo

In his Book, *Children's Drawings as Diagnostic Aids* (1973), Di Leo moves away from his earlier work (1970) – that espoused the developmental progressions and universality of children's drawing development – and focused on using drawings as diagnostic aids for social-emotional traits. He believed that an important element of young children's drawings was overlooked: the emotional aspect.

Di Leo (1973) stated that on the rare occurrence when the Goodenough DAMT result failed to match the child's intellectual capacity, other forms of assessment should be sought to ascertain the reason for the discrepancy. In particular, he felt that the omission of parts of the body in children's human figure drawings was:

> …indicative of feelings and personality traits … insecure, anxious children tend to draw small figures that timidly occupy only a small area of the available space. In contrast, the secure, well adjusted child will draw freely, with joyful abandon, creating a figure that expresses, by its size, sweep, and conspicuous placement on the page, freedom from inhibiting anxiety. (Di Leo, 1970, pp.34-36)

Di Leo's hypothesis that children's drawings could reflect their emotional well-being and project their personality profile has been popular. Numerous psychologists, social workers, and police have adopted his method in an attempt to reveal whether children have learning difficulties, special needs, or have been subject to abuse, assault, and other crimes. However, in the Epilogue of his book (1973) Di Leo states that his endeavors did not have the empirical reliability and validity strength of Goodenough's DAMT (1926a):

> The validity of human figure drawings as expressions of intellectual maturity has been demonstrated by numerous investigators … When drawings are viewed as expressions of feelings, as projections of personality, the same degree of agreement has not been achieved. Here the subjective element in interpreting the drawings introduces a practically uncontrollable variable. Experience, insight, and intelligence play major roles and no two examiners possess these attributes to the same degree. (p. 217)

Thus, Di Leo (1973) acknowledged the limitations of his hypothesis of using children's human figure drawings as diagnostic aids, and he espoused the viability of the GDAMT. Echoing Goodenough's views on assessment, at the close of his book Di Leo imparted that "drawings are but a part of a comprehensive evaluative process" (1973, p. 213).

Suzi Gablik

Suzi Gablik is an artist, art critic, and art historian who has lectured at university level. Her book, *Progress in Art* (1976), presents a theoretical synthesis of many paradigms in an effort to construct what she asserts is a "general theory of art history" (p. 6). Focusing on the overarching concept of art, rather than drawing, Gablik combined the disciplines of cognitive and perceptual science psychology, philosophy of science, systems theory, child development, and art history to create her own theory of art history. In particular, Gablik applied Piaget's theories of cognitive development and conception of space in the child to what she calls megaperiods of art history (i.e. Ancient and Medieval, The Renaissance, Impressionism, Cubism, and so on) in an attempt to demonstrate that they align. In doing so, Gablik suggests that art created by adults in the Ancient and Medieval times is equivalent, cognitively, to that created by children considered to be in the pre-operational stage of thought today.

Indeed, Gablik's research has links to recapitulation theory, otherwise known as ontogeny recapitulates phylogeny. Generally, recapitulation theory suggests that embryonic development parallels the adult stages of earlier life forms (Gould, 1977). Whilst modern biology recognises the links between ontogeny and phylogeny, the universal application of recapitulation theory as a whole is largely rejected (Medicus, 1992). Needless to say, Gablik's theory of art history is considered somewhat controversial.

Recent Studies of Young Children's Drawings

More recent studies on young children's drawings are concerned less with the developmental aspects – or even human figure drawings specifically – but with the socio-cultural element of drawing more generally (Anning & Ring, 2004; Brooks, 2002, 2003, 2004 2005; Lambert, 2006; Ring, 2001). In their book, *Making Sense of Children's Drawings*, Angela Anning and Kathy Ring (2004) do agree that "drawings provide rich insights into young children's thinking" (p. x); however, their principle concern is the "reappraisal of the role of drawing in young children's learning and in their attempts to make sense of and represent the worlds in which they are nurtured and educated" (p. x). Anning and Ring (2004) consider children's drawings to be "under-valued, under-researched, and misunderstood within the domains of childhood studies and early childhood education" (p. xi).

Applying a Vygotskian theoretical perspective, Anning and Ring stated that drawing – and the objects necessary for drawing such as pencils, crayons, paper – together with people and contexts are socio-cultural tools for the child. Their philosophy aligns with the ideas proposed by Kress (1997) who argues that "children act multi-modally, both in the things they use, the objects they make, and in the engagement of their bodies; there is no separation of body and mind" (p. 97). Consequently, Anning and Ring are concerned, not with *how* children draw, but *why* children draw; their aim was "to try to find out what influenced a group of young children's drawing behaviours in the contrasting contexts of home and school" (Anning & Ring, 2004, p. 27). They found that the quality of care and education within the various contexts they investigated impacted, somewhat, on what and how the children in their study drew. Anning and Ring (2004) asserted that drawing is a powerful tool for communicating, and developing further cognitive functioning. Reiterating, somewhat, the underlying theme of this research, Anning and Ring closed with the following statement: "We need a society that can listen to children and recognize that perhaps their drawings may tell us much more about childhood than we ever imagined" (Anning & Ring, 2004, p. 124).

Margaret Brooks (2002, 2003, 2004, 2005) also applied a Vygotskian lens to her research on young children's drawings and their potential role as a learning tool. However, she seemed less than appreciative of the research undertaken by those before her. Brookes asserted that Luquet (1913, 1923), Piaget (1956), and Kellogg (1967, 1970) were far too focused on the developmental aspects of drawing; yet, like many other researchers, she seemed to over-generalise and over-simplify their theories. Similar to the findings put forward by Anning and Ring (2004), Brooks (2002, 2003, 2004, 2005) stated that drawing is a powerful tool that can assist young children in understanding further the world around them, and representing these understandings in meaningful ways before having mastered reading and writing skills.

Lambert (2006) has also had success in using young children's drawings for representing understandings and solving real-life problems. Her case study of one preschooler named Carlos showed that the use of, what she refers to as, diagrammatic representations (drawings) was effective in facilitating Carlos in communicating his thoughts and understandings of real-life issues as they arose. Lambert (2006) states, "there is a place for the use of diagrammatic representation as an aid to problem-solving in any learning environment and … there may be social benefits as well as cognitive ones" (p. 47).

Interestingly, MacPhail and Kinchin (2004) used children's drawings as an assessment tool in the aim of making evaluation in their physical education class more child-centred. They were keen to investigate the children's perceptions of the learning experiences presented, and use this feedback to improve future teaching and learning experiences. They state:

> we attempt[ed] to address the concern that voices of young people are rarely heard in curriculum interventions … In doing so we consider the use of student drawings as a child-centred procedure and evaluation tool … student drawings have the potential to encourage discussion amongst teachers and students about the teaching and learning environment. (MacPhail & Kinchin, 2004, p. 88)

It would seem that, regardless of the context in which they are examined, young children's drawings are complex and meaningful marks that can communicate a great deal of information. Despite the many efforts to extend the body of knowledge on young children's drawings, few could claim to have come close to having an impact comparable to Goodenough's. The only research on young children's drawings to come close to approaching the work of Goodenough was, perhaps, that conducted by her doctoral student, Dale Harris.

Revision and Extension of the Goodenough Draw-a-Man Test

Of all the many projective tests and other research projects spawned by the effectiveness of the GDAMT, Harris's (1963) endeavors were considered the most useful in revising and extending the method. Indeed, Harris worked closely with Goodenough, completing this doctorate under her supervision. The revision and extension was undertaken as a collaborative project between Harris and Goodenough. Whilst Goodenough's health at the time prevented her active participation, Harris (1963) stated that she was a keen supporter of the project: the GDAMT is now known as the Goodenough-Harris Drawing Test (GHDT). His book, *Children's Drawings as Measures of Intellectual Maturity* (1963) is dedicated to the memory of Florence L. Goodenough.

An impetus for Harris's revision and extension of the GDAMT, amongst other things, included a statement by Stewart (1953) in the *Fourth Mental Measurements Yearbook*. She pointed out that the norms for many tests show changes after approximately 30 years; thereby, suggesting that the GDAMT, originally developed in 1926, was due for re-evaluation and re-standardization. Given that Harris' book was published in 1963, it could be concluded that the GHDT is due, yet again, for re-evaluation and re-standardization. Harris was also interested in expanding upon the GDAMT idea to align the test with societal views of the time. There was pressure for the GDAMT to be more culturally and socially inclusive, and to incorporate drawings of women and children in order to promote equality of outcomes for test participants. Also, Harris was keen to see if the test could be extended upwards to include children older than the recommended age of around 10 or 12 years.

Draw-a-Man, Draw-a-Woman and Self-Portrait Sub-Tests

Harris was successful in extending and defining further the Draw-a-Man (DAM) scale, and re-establishing its reliability and validity (Harris, 1963). He extended the original DAM scoring scale from 51 items to 73 items, and complemented it with a Draw-a-Woman sub-test (DAW) and projective self-portrait sub-test (SP) (see Goodenough, 1926a for the GDAMT administration and scoring guide; see Appendix A for the GHDT administration and scoring guide). Harris devised a 71 item scoring scale for the DAW sub-test; however, a dedicated scoring scale for the SP sub-test has not been developed. Harris's extensive study affirmed, principally, that Goodenough's original research was both meticulous in design and execution, and little could be done to enhance it.

In revising the original DAM test, Harris followed three criteria:

> (1) The items should show a regular and fairly rapid increase with age, in the percentage of children passing the point. (2) The items should show a relationship to some general measure of intelligence. (3) The items should differentiate between children scoring high on the scale as a whole and those scoring low on the scale as a whole. (Harris, 1963, p. 74)

Using the above criteria, Harris had around 100 items that seemed useful for the DAM scoring scale. Careful examination of the separation in percentages of children who were high-scorers and children who were low-scorers on each item helped to determine which final items would appear on the DAM scoring scale. In other words, items that were just as easily included in drawings by children who were less-successful as they were by children who were more-successful on the test were not included. All 51 of Goodenough's original items were included in the final 73 item scoring scale.

The same three previously mentioned criteria were applied in the development of the DAW scoring guide. Out of a potential 90 items, Harris devised a DAW scoring guide which included 71 items; 50 of which were common, or equivalent, to items in the DAM scoring guide. As a dedicated scoring scale for the self-portrait sub-test was not devised, test administrators were advised to use the DAM or DAW scoring scale dependent on the sex of the child.

At first, Harris thought that three human figures drawings might be too much for some young children to complete in one sitting. However, preliminary work indicated that children were willing to complete all three drawings, and easily did so within a short amount of time (generally no longer than 10 to 15 minutes). The test administration manual encourages the administrator to provide young children with a short break between the second and third drawings, if necessary. On the other hand, older children (above Goodenough's original recommended age range upper limit of around 10 years of age) were found to be increasingly reluctant to draw at all – especially drawings of themselves (Harris, 1963).

Philosophical Shifts

Harris stated that the nearly forty years of research in psychology since Goodenough's original work resulted in some important philosophical shifts. Foremost, the concept of a single general intelligence was under intense scrutiny and considered problematic; second, the investigators of this ability, or trait, were now sensitive to socio-cultural factors; and the age scale method of expressing levels of any sort of development was being superseded by more sophisticated and accurate measures (Harris, 1963).

From the beginning, Goodenough (1926a) – and later Harris (1963) – were cautious in stating that the GDAMT and the GHDT were tests of intellectual maturity, or intellectual development, rather than tests of *general intelligence*. This was mostly due to the varying correlation statistics of the GHDT and other more conventional tests of intelligence such as the Stanford-Binet; and also because both Goodenough and Harris had both researched the concept of intelligence and understood that it was a complex trait (Goodenough, 1926; Harris, 1963; Thompson, 1990).

The challenging of the idea of a single general intelligence factor, or *g* as it is known, had resulted in the notion that there are several types of intelligence. It has been suggested that there is "probably more to intelligence than just the general factor" (Sternberg, 1999, p. 437). Indeed, various theorists have diverse views on intelligence: Gardner (1983, 1999) maintains there are eight intelligences; Spearman (1904) adheres to the factorially based cognitive theory of intelligence; Carroll (1993) holds a 'Three-stratum theory'; and Sternberg proposes a triachic theory of intelligence (1985, 1986, 1998, 1999), just to name a few notable examples. Thus, Harris (1963) was keen to clarify that whilst results from the GHDT often aligned with those produced from more traditional intelligence tests, the aim of the drawing assessment was not to investigate intelligence, but intellectual development instead.

Further, Harris wanted to move away from the idea of the GDAMT being seen as a predictor of potentialities; as it had been repeatedly recognised that development of abilities could be estimated effectively only after they had actually developed (Harris, 1963). He wanted to diminish the idea of drawing development being determined by normal or typical growth that is age-determined, which was one of the issues with the age-scale concept. Harris (1963) stated:

> The assumed straight line mental age function simply does not aggress with the facts of development. As the process of development is more completely understood, age, or time, seems to be only a crude index. Cumulative changes that occur though accretion or association seem to depend on repetition or reinforcement, not merely lapse of time. (p. 4)

Harris (1963) even mentioned the potential for applying the more sophisticated statistical models that were being created at the time. He stated that they had immense potential for "describing variation and for scaling in reference to such normal variation … these scales [could] accurately place the individual in comparison with a known or defined group. [But] unfortunately, these methods do not readily supply growth measures; they are purely relative" (p. 4). Since then, however, even more sophisticated statistical models have been created (such as the Rasch family of measurement models) that surpass the psychometric limitations that Harris noted.

Goodenough-Harris Drawing Test Scoring Scale

Harris (1963) continued with Goodenough's scoring method whereby items were scored using a simple two-step absent/present system. If an item is absent from the drawing, or present but not within the specified criteria the item is considered absent and no credit is awarded. Alternatively, the item is present and within the specified criteria and therefore it is awarded one credit. The sum of the credits for each drawing is considered to be the total raw score. Therefore each child receives three raw scores, one each for their drawing of a man, a woman, and themselves. The GHDT scoring guides are presented in Appendix A.

There is a problem with this type of scoring method, however, in that the raw counts that make up the raw score are treated as though they are equal-interval measures. As Wright and Mok (2004) point out:

> ....raw counts are only indications of a possible measure … [they] cannot be the measures sought because in their raw state, they have little inferential value. To develop metric meaning, the counts must be incorporated into a stochastic process which constructs inferential stability … raw counts may give the impression that are interval (or ratio) measures of experience. But this is always an illusion. In particular, raw counts at the beginning and end of a raw score scale are problematic because while the counts necessarily terminate at "none" or "all", the measures they might imply have no boundaries. (p. 2-3)

The above excerpt highlights the problematic nature of the GHDT scoring scale, and ordinal counts in general.

Standard Scores and Percentile Ranks

Whilst Harris adhered to the dichotomous scoring of the GDAMT, he also introduced the idea of GHDT standard scores and percentile ranks. Harris constructed tables which could be used to convert children's GHDT raw scores into standard scores (see Harris, 1963, pp. 294-301). The standard score expressed the child's relative standing on the GHDT relative to the child's age and sex, in terms of a mean of 100 and a standard deviation of 15 – the same distribution as adopted for many IQ tests. A standard score could be ascertained by looking up the child's age in years and months together with the raw score achieved for a drawing in the tables provided. The values in these tables were calculated from the sample Harris used for the revision and extension of the GDAMT. How these values were calculated is not explained.

Children's standard scores could also be interpreted in terms of percentile ranks, to show their "relative standing in a theoretical group of 100" (Harris, 1963, p. 311). Harris provided a table for converting standard scores in percentile ranks (see Harris, 1963, page 311). Harris's justification for constructing percentile ranks was that it was considered more readily understood by lay people and was, at the time, often used in school testing.

The adoption of raw scores, standard scores, and percentile ranks is not without issue. The GDAMT and GHDT scoring guides are merely ordinal level scales, at best, and offer a total raw score for a child's drawing of a man, woman or self-portrait. These total raw scores for children's drawings are then compared. Higher raw scores are considered evidence of a better or more detailed drawing than one which was credited with a lesser raw score. The problem herein is the fact that the ordinal nature of the GHDT scoring guides and the equal value of each item have been merely assumed. For example, two children might have received raw scores of 22; however, one child might have achieved this raw score by drawing 22 of the less difficult items to include, whereas the other child might have received this same raw score for drawing 22 of the more difficult items to include in their representation of a man. So when such raw score totals are then used to determine standard scores and percentile ranks, misinterpretation and misunderstanding must ensue. The raw scores are treated as though they are equal-interval measures, when they certainly are not. At least one can conclude that the extent to which these scores might contribute to interval measures has not been investigated empirically. The scoring method of the GHDT presents as a weakness, and requires investigation from a modern test theory perspective.

Goodenough-Harris Drawing Test Ceiling and Floor

One of the most significant findings from the research of Harris (1963, 1977) and others (Harris & Pinder, 1974; Harris & Roberts, 1972; Scott, 1981) to date is that none have "managed to extend the usefulness of the scale into the adolescent years" (Harris, 1963, p. 99). They consistently found that beyond the age of about 10 or 11 years children cease to show meaningful progression in the development of their human figure drawings. Factors other than those dependent on intellectual development seem to impact on the detail of drawing.

Harris's revision and extension of the GDAMT included the aim to raise the ceiling of the test. The standardization of the GHDT indicated that children beyond the age of approximately 11 or 12 years ceased to show significant growth on either the DAM or DAW test (Harris, 1977; Scott, 1981). The separation between children's age ranges beyond about 12 years of age was approximately one raw score point, instead of the usual three to four raw score point annual increases for children below the age of 12 years. With the aim of resolving this issue, Harris added several items to both the DAM and DAW point scales that were exclusive to the drawings of children aged 12 years and older (Scott, 1981). Despite these sound intentions, national probability samples of 13,000 children aged between six and 17 years old (Harris & Pinder, 1974; Harris & Roberts, 1972) showed that whilst the GHDT "reliably discriminates performances at each age from 6 through 12 … no gain in scores could be expected between 13 and 17 years on either the [DAM] or the [DAW] scales" (Scott, 1981, p. 486). This corresponds with Goodenough's original findings that the test is most effective with children aged between about four and 10 years of age (1926a). Indeed, Harris states:

> ... despite the extensive and intensive effort to develop new items that would extend the scale upward in age, it has been noted that few additional items could be found. Some items added are actually elaborations or subdivisions of existing points. None has managed to extend the usefulness of the scale into the adolescent years. It is a tribute to Goodenough's insight and scholarship in her original work with children's drawings that few additional items have been found. (1963, p. 99)

Whilst the rather abrupt plateau of GHDT performances in early adolescence could be because of the limited number of creditable details in a typical human figure drawing, Scott (1981) suggests otherwise. The ceiling of the test seems to parallel the "developmental shift from the habits of concrete conceptualization to abstract models better expressed verbally" (Scott, 1981, p. 486). In other words, the ceiling effect of the test seems to occur at around the time that adolescents are consolidating their concrete operational mental structures and, perhaps, transitioning towards the formal stage. And as Scott (1981) suggested, adherents to Piagetian theory would assert that the abstractions of formal thought would be likely expressed verbally, rather than pictorially.

There have been very few attempts to extend the drawing scales downward. Ford, Stern and Dillon (1974) collected drawings from 58 children aged three, four and five years; although, no comparisons with drawing test norms could be completed as the sex of the human figure drawings was not recorded. Many would consider the idea of extending the GHDT scoring downwards as rather ambitious, as most authorities in the field recognise that typically children begin to produce human figure drawings at the age of around 3 or 4 years – which is the minimum recommended age for the GHDT anyway (Cox, 1992, 1993, 1997; Goodenough, 1926a; Harris, 1963; Kellogg, 1967, 1970; Piaget & Inhelder, 1956, 1971).

Goodenough-Harris Drawing Test Quality Scale Cards

Following the trends at the time (see Burt, 1921; Dunn, 1954; Thorndike, 1913; Lantz, 1955; Kerschensteiner, 1905; Wagner & Schubert, 1955), and in an effort to make the GHDT more user-friendly, Harris developed the Quality Scale Cards (QSCs). Harris (1963) states that, at the time, various types of 'Quality Scales' were provided for children's handwriting samples, creative drawings, paintings, and so on, and were considered a useful concept for quickly ascertaining the approximate 'quality' of productive output. Furthermore, Harris considered that QSCs would assist test administrators in quickly assessing children's drawings; "the quality scale appears to provide a convenient and economical, as well as valid and reliable index to children's drawings of the human figure" (Harris, 1963, p. 110).

The GHDT QSCs were a series of 12 drawings made by children – participants in Harris' study – which were said to represent increments in conceptual maturity (Scott, 1981). Selecting the quality scale card which best resembled a child's drawing in terms of overall detail, proportion, and conceptual maturity was said to enable quick conversion of approximate raw scores into standard scores. Whilst Harris and his assistants attempted to create a, somewhat, objective "equal-appearing interval" (Harris, 1963, p. 110) scale, their methods were not sophisticated enough to render a reliable, equal-interval scale (by today's standards) which was suitable for the task.

The inexact, rudimentary QSC procedure highlights the crucial issue of measurement in the human sciences. Rather than merely guessing or approximating children's abilities and drawing development levels using ordinal scales or quality estimates, modern test theory calls for human attributes to be meticulously examined.

Limitations of the Goodenough-Harris Drawing Test

The reliability and the validity of the GHDT have been established in traditional test theory terms by Goodenough (1926a), Harris, (1963), and their colleagues (see Ferrien, 1935; Brill, 1935, 1938; McCarthy, 1944; McCurdy, 1947; McHugh, 1945a, 1945b; Smith, 1937; Williams, 1930; Yepsen, 1929); although, these qualities remain unexamined from a modern test theory perspective. Consequently, four principle issues await further examination: 1. GHDT test unidimensionality; 2. the possibility of producing equal-interval scales; 3. calibration of item hierarchy; and 4. the possible redundancies in the current four sub-test design. Indeed, it would be significant to investigate what additional information is gathered by the DAW and SPM/SPW sub-tests beyond that already yielded by the original DAM sub-test. Also, the investigation of whether it is possible for children to simply draw a human figure of *any sort* – that is, a male or female child or adult – and it be evaluated by a single 'human figure drawing continuum' scoring scale would be extremely worthwhile. There are also possible issues in that the tracking of children's drawing development over time is somewhat difficult in the absence of interval-level scoring scales. Clearly, there are critical elements in the conception, implementation and interpretation of the GHDT which await thorough investigation from a modern test theory perspective.

The potential that the GHDT holds for evaluating effectively the developmental aspects of young children's drawings – and thus, informing early childhood educators, parents, and psychologists – is quite considerable. However, this potential cannot be verified or realized without examination of the GHDTs underlying psychometric properties from a latent trait theory perspective. The Rasch measurement for dichotomous data model is considered the key model for this research, as it adheres most closely to the idea of producing fundamental measurements (Bond & Fox, 2007; Wright & Stone, 1979).

Applying the Rasch Model

The usefulness of a psychological measurement scale is dependent directly on its reliability (i.e. the consistency with which it measures), and its validity (i.e. the demonstration that it does assess the abilities which it claims) (Harris, 1963). Revision of psychological assessments is necessary to ensure relevance to the evolving sample it is intended for, and consequently, to maintain reliability and validity (Harris, 1963; Stewart, 1953).

Unlike researchers in the commerce and engineering fields who can utilise dependably their physical measures of attributes such as weight, volume, length and the like, researchers in the human sciences typically do not demand the same kind of consistent and reliable measures in their own field (Bond & Fox, 2001). Given that high-stakes decisions about children's education, for instance, are based on the conclusions of researchers in the human sciences, it seems unquestionably critical that these conclusions be based on sound, quality measures. Bond and Fox (2001) argue that researchers in the human sciences are "too narrowly focused on statistical analysis, and not concerned nearly enough about the quality of the measures on which they use these statistics" (p. 1).

The reliability and validity of the GHDT has been established classically by Goodenough (1926), Harris (1963), and others. However, just as society, values, and attitudes have changed since the 1960s so too has research test theory. Thus, to maintain the production of quality, reliable and valid conclusions from the administration of the GHDT, it must be examined from a test theory perspective that encompasses up-to-date, modern technologies.

A key model in modern test theory, the Rasch model is a quantitative, probabilistic measurement tool that can "transform raw data from the human sciences into abstract, equal-interval scales" (Bond & Fox, 2001, p. 7). Unlike traditional relational statistical analysis tools, the Rasch model assumes that task performance is dependent on *both* the ability of the subject, *and* the difficulty of the task item (Bond, 1995; Bond & Fox, 2001; Bond & Bunting, 1995). Indeed, it is the only model "that provides the necessary objectivity for the construction of a scale that is separable from the distribution of the attribute in the persons it measures" (Bond & Fox, 2001, p.7).

The Rasch model for measurement was developed by Georg Rasch (1901-1980), a Danish psychometrician and mathematician. Amongst his several achievements, Rasch developed the Rasch dichotomous (or one parameter) model to apply to data gathered from the administration of intelligence and attainment tests. The Rasch model is most extensively used in education and educational psychology, however, it is becoming increasingly popular in other disciplines such as medicine, occupational therapy (Chien, 2007), and sport (Linacre, 1995).

Other psychometricians have expanded upon the original Rasch one parameter model, and there is now a family of Rasch models (Wright & Mok, 2004). This family includes the rating scale model (Andrich, 1978), the partial credit model (Masters, 1982), the binomial trial model (Wright & Masters, 1982), and the Poisson counts model (Wright & Masters, 1982), amongst others (see Wright & Mok, 2004). This research uses the Rasch one parameter model.

A significant advantage of the Rasch model is its requirement that the items from a scale measure a single unidimensional construct at a time (Rasch, 1960; Wright & Stone, 1982). The Rasch model's unidimensionality principal establishes whether the items within a test or instrument measure a single latent trait or ability (in this case, intellectual development) to create a meaningful continuum of items ranging from comparatively easy to difficult to accomplish (Chien, 2006; Bond, 2004; Bond & Fox, 2001). For example, if all items from an intellectual development assessment adhere to the Rasch model's unidimensionality principal, unidimensionality of the assessment is maintained. As a result, psychologists, researchers, and others can be confident when drawing conclusions from the assessment as the items within it have been *confirmed* to measure what they *intended* to measure. Additionally, perusal of the Rasch model's fit statistics will help identify any items that do not contribute to the measurement of the latent trait under investigation (Linacre, 2002). Consequently, such items can then be revised or removed to enhance the validity of the assessment. Similarly, the Rasch model's fit statistics are useful for detecting misfitting persons. This data is valuable in establishing whether persons responded to the test item in accordance with the latent trait, or whether other answering mechanisms or factors might have interfered (Chien, 2007).

Perhaps the most noteworthy feature of the Rasch model, briefly mentioned above, is its ability to transform raw data into an equal-interval measurement scale (Bond & Fox, 2001; Rasch, 1960; Wright & Masters, 1982). Provided the test satisfies the Rasch model's unidimensionality principle, both test items and test participants can be represented on a single continuum – the logit scale – highlighting relationships between both that would otherwise remain invisible. Thus, the logit scale creates opportunities to make comparisons between and amongst the items and persons, and draw reliable and valid conclusions about both, that otherwise would not be possible (Bond & Fox, 2007; Linacre, 2009). An example of such includes the helpful item-person map that indicates how suitable the test items are for the given sample (Chien, 2007).

Other researchers' successful application of the Rasch model for measurement of developmental data (e.g., Bond, 2001; Bond & Bunting, 1995; Bond & Parkinson, 1996; Bunting, 1993; Chien, 2007; Drake, 1998) confirm it as an apposite tool for the data analysis of this project.

Although, some people from outside objective measurement circles might feel wary of the margin of error associated with the Rasch drawing estimates, it is not unreasonable to think that such high-stakes assessments regarding children's levels of cognitive development should be, somewhat, error-free. Indeed, even the most long-standing and trusted forms of measurement possess a margin of error.

Take the measurement of time, for example. Despite the huge number of clocks and watches in society today, it is unlikely that any significant number of them would display precisely the correct time for any meaningful length of time throughout the day. In fact, it could be said that the measurement of time is habitually viewed as unproblematic and, perhaps, error-free. When a group of people discovers their watches out of sync by a minute or two, often little is said or done to rectify this 'margin of error' in the measurement of time. Few people would bother to adjust their watches or clocks, let alone begin to panic about the inaccuracy and lack of precision in the offending timepiece. Indeed, Bond & Fox (2001) discuss briefly in their text a relevant exemplar from the June 2000 issue of *Discover* magazine:

> …[such are] the difficulties that currently beset those who maintain watch over the world's time. Although the passage of time was previously marked by sunrise and sunset, and more recently by pendulum clocks, which also are calibrated against the rotation of the earth, inconsistencies in our planet's shape and movement make it unreliable as a basis for the measurement of time. From 1967, the standard definition of the period of 1 second moved to the atomic scale based on the radiation of the cesium 133 atom. Although the $650 000 cesium clock in Boulder, Colorado, estimates the passing of 1 second

> with almost unbelievable precision, it does not directly keep time. The keeping of the international time standard is the responsibility of the Paris Observatory. Moreover, it is based on an average. The time estimates of more than 200 atomic clocks from around the world are routinely collected, averaged, and fed back to the originating laboratories so that the computer-generated records on the passage of time at each laboratory can be adjusted regularly for the local clock's deviation from the world average, even if it is somewhat after the fact. (Klinkenborg, 2000, in Bond & Fox, 2001, p.7)

So whilst all forms of measurement possess margins of error to some degree, the error does not go completely unnoticed and unaccounted for. The margin of error becomes significant and, sometimes, troublesome when it exceeds expectation. The Rasch model produces summary statistics to help researchers identify unexpectedly high levels of error. Thus, like almost any entity that measures constructs with any precision (e.g. clothing manufacturers, builders, engineering firms, the world's time-keepers, etc.), those in the business of constructing measures for psychological or educational purposes will accept small margins of error suitable to the use to which the results of testing will be put (Bond & Fox, 2007; Chien, 2007).

The Rasch one parameter model presents as the most apt statistical analysis tool for the task of examining the psychometric properties of the GHDT. The WINSTEPS® computer program, which is based on Rasch model principles, will be applied to the data collected for this project. WINSTEPS® was developed by John Michael 'Mike' Linacre, a world-authority on Rasch analysis. WINSTEPS® is considered to be one of the most user-friendly and practical software programs based on the Rasch model.

Research Questions

The discussion presented here identifies important psychometric issues that affect the validity and utility of the results obtained from the GHDT. It is proposed that the Rasch model for measurement could be useful in verifying and resolving most of these issues to create scientifically-sound drawing continuum appropriate for both boys and girls aged between approximately 4 and 10 years of age.

The following two principal research questions will guide this investigation:

1. Is it possible to Rasch analyse each of the four sub-tests (DAM, DAW, SPM and SPW) of the Goodenough-Harris Drawing Test to produce Rasch measurement scales?
2. To what extent do the generated measurement scales fit the Rasch model's unidimensionality requirement, and its other expectations?

Assuming the successful application of the Rasch model to the analysis of the GHDT, the following research questions become relevant:

3. To what extent do the generated scales adequately summarise/quantify the drawings of young children included in the study and, if they do, to what extent do these scales differ from the original hierarchy?
4. What additional information beyond that from the DAM sub-test, if any, is revealed by the DAW and SPM/SPW sub-tests?
5. How do the children's GHDT drawings change over time?
6. What modifications might update the GHDT to better align it with expectations of current educational users?

7. Is it possible to develop a Human Figure Drawing Continuum (HFDC) that includes some of the items of the current GHDT that *could* be useful for evaluating any human figure drawings made by young children today?

Chapter Three

Method

This research, underpinned by Constructivist and Latent Trait Theory, adopted a cross-sectional and longitudinal research design in order to investigate the Goodenough-Harris Drawing Test (GHDT) using the Rasch model for measurement. The cross-sectional aspect facilitated the gathering of a broad range of human figure drawings produced by children of different ages and abilities at the one time (or one phase of data collection in this instance), which was useful for the initial Rasch analysis and principle research questions. The longitudinal aspect involved the gathering of individual children's drawings produced over a 12 month time frame (i.e. three phases of data collection, each approximately six months apart), which were useful for verifying the results obtained from the initial Rasch analysis, and investigating the development of children's human figure drawings over time.

The cross-sectional and longitudinal approaches were deemed apt for enabling a thorough investigation of the GHDT from a modern test theory perspective; thus, offering the researcher the best possible chance of satisfying each of the research questions outlined in Chapter Two.

Context

All data were collected from an urban preparatory to year 12 school in Townsville, Queensland, Australia. The school setting presented as an appropriate, convenient, and willing host for data collection, and the school community was deemed to be a suitable representation of society from which a sample could be recruited for the research.

The *National Statement on Ethical Conduct in Human Research* (NHMRC, ARC, & AVCC, 2007) states that one of the most significant issues with research, of any kind, is the fact that it often "may lead to … inconveniences for participants and/or others" (p. 16). The researcher's employment as a preschool teacher at the school the year prior to data collection was held to have several conveniences for all stakeholders in this research; in particular, with regard to data collection. The already-established reciprocal relationships with administration staff, teachers, and students were considered a benefit to recruiting participants and scheduling data collection for the research. Similarly, the researcher's insider knowledge of the school's routines, policies and procedures meant that data collection would result in minimal disruption to all involved. Furthermore, the administration staffs of the school were aware of the researcher's project and intention to collect data; thus, they expressed an interest in being the 'host school' of the research. The general convenience of having pre-approval from administration staff to conduct research at the school, together with the other advantages discussed above, presented as an opportune circumstance for data collection.

In addition to the school setting being considered appropriate for and willing to participate in the study, the school community needed to be suitable for sampling. The researcher's previous employment at the school not only resulted in positive relationships and knowledge of the school routines, but also an insider understanding of the school community. The school presented a unique opportunity to access children from a variety of backgrounds within the one setting. The school's location, at the nexus of where long-established housing meets newer master-planned housing estates, resulted in a surprisingly diverse school community. Children from upper-middle class and middle-class type families are enrolled at the school and typically pay the standard rate of fees to attend. Also, lower-income, working-class type families are also part of the school community; and some of these families – for various reasons – are exempt from paying the full rate of fees. This diverse school community presented a unique and convenient opportunity for the researcher to collect data from a representative sample of students, at the one school. Given the aims and scope of this research, this population was deemed suitable for sampling with an acceptable level of external validity (Trochim, 2006).

Sample

In gathering a suitable sample from the host school for this research project, three key elements needed to be considered. First, the cross-sectional aspect of this research required a sample of children of various ages (Cox, 1992). Second, research undertaken by Goodenough (1926) and Harris (1963) indicated that the GHDT was most appropriate for children aged between approximately 4 and 10 years of age. Third, the longitudinal aspect of this research meant that children would be administered the GHDT three times across three phases of data collection, each approximately six months apart.

Consequently, Research Information Sheets and Informed Consent Forms (see Appendix B) were distributed by the class teachers to over 300 children's families across prep and years one to four of the host school. This process yielded a convenience sample of 107 children. Only those children whose parent or caregiver had returned an endorsed consent form were included in the research.

For phase one of data collection, the sample (n = 107) comprised 62 girls and 45 boys from prep through to year four. At the time of this data collection, which took place in March 2007, the children were aged between 4 years and 7 months and 10 years and 1 month. The breakdown of children from each year level is as follows: prep, 18 children (6 males, 12 females); year one, 17 children (9 males, 8 females); year two, 11 children (3 males, 8 females); year three, 35 children (16 males, 19 females); and year four, 29 children (10 males, 19 females).

Phase two of data collection comprised a total of 83 children, (33 males and 50 females). At the time of this data collection, during October and November 2008, the children were aged between 5 years and 1 month, and 10 years and 7 months. The breakdown of children from each year level is as follows: prep, 16 children (6 males, 10 females); year one, 14 children (7 males, 7 females); year two, 10 children (3 males, 7 females); year three, 18 children (7 males, 11 females); year four, 25 children (10 males, 15 females).

Phase three data collection included a sample of 56 children (21 males and 35 females). This final phase of data collection took place during March and April of 2008. At this time, the children were aged between 5 years and 7 months and 11 years and 1 month. The sample comprised: year one, 13 children (5 males, 8 females); year two, 8 children (4 males, 4 females); year three, 3 children (1 male, 2 females); year four, 9 children (3 males, 6 females); year five, 23 children (8 males, 15 females).

Attrition of the sample over the duration of the study was due to some children leaving the school, others were absent from school on the scheduled data collection day, or had other engagements at school at the time of data collection (i.e. guitar lessons, one-on-one testing, reading recovery). Whilst some decrease in sample numbers across the phases was anticipated, the considerable attrition rate (about 30% for phases two and three) was not expected. The researcher made every effort to include all children from the phase one sample, however, on numerous occasions it was simply not possible to obtain the data.

The sample size, whilst comparatively small, was considered sufficient by the researcher and the supervisory team to reflect trends in the data, and remain manageable for the researcher. The administration of the GHDT required detailed data collection procedures, including scheduling, test administration, test scoring, data coding, and data analysis which was an extensive process. A projective data collection time table and administration procedure document was devised in advance. This document indicated that a suitable phase one sample size of between 80 and 120 children would be conducive to thorough and timely data collection and data analysis. Phase one comprised a total of 107 children which sits within this predetermined sample size. Moreover, previous studies on young children's drawings conducted by other researchers (Barrett & Bridson, 1983; Barrett, Beaumont & Jennett, 1985; Blain, Bergner, Lewis, & Goldstein, 1981; Cox, 1997; Davis, 1985; Fowlkes, 1980; Golomb, 1981, 1992; Harris & Pinder, 1974; Ingram, 1985; Kendrick & McKay, 2002; Lambert, 2006; Light, 1985; Luquet, 1923) indicated that this sample size was appropriate for the scope, aims, and time frame of this study.

Informed Consent

Approval to collect data at the Townsville-based primary school was sought, in writing, from the Director of Education in the district, as well as the Principal and the Deputy-Principal of the school. The researcher then approached two to three teachers from each year level (Prep to Year five; about 12 classes in total) and, after discussing the aims of the study, offered the opportunity to participate in the project.

Each child in the selected teachers' class received an Information Sheet and Informed Consent Form distributed with the school Newsletter. Completed Informed Consent Forms were collected by the class teachers and returned to the researcher via a collection tray in the school's Staff Room. A suitable date and time for test administration was then negotiated with each teacher individually. All participating teachers and students were reminded that they had to right to withdraw from the study, or withhold any response to any question, at any time.

The Instrument and Procedure

The Goodenough-Harris Drawing Test (GHDT) is a non-verbal test for young children (aged between approximately 4 to 10 years of age) used for the inference of intellectual development levels. It can be administered to children either individually or in small groups using the same instructions, although, preschool children and children with special needs should be administered the test individually.

The GHDT requires children to be seated individually at desks and, using a 2B pencil, make three drawings in a test booklet: one each of a man, a woman, and a self-portrait. Provided that children attempted to draw a whole person (i.e., head to feet – not a 'bust'-type portrait) to the best of their understanding, there is no 'correct' or 'incorrect' type of drawing. Children are discouraged from holding up, discussing, or verbalizing aspects of their drawings to avoid influencing their own or others' results.

There is no time limit for the test, although children rarely take longer than about 15 minutes to complete all three drawings. The longest that a child took to complete all three drawings in this research project was 20 minutes.

Test Apparatus

The GHDT administration required only two items: a number two pencil and the GHDT booklet. However, several erasers, sharpeners, spare pencils and booklets were included in the data collection kit.

At the commencement of the test administration, after a general discussion about drawing to elicit terminology and help children to feel comfortable, the researcher advised the participants that they would be using "special" pencils and paper to do their drawings. This helped to foster a positive, understanding, and encouraging climate in the room, as well as motivate any children who were, perhaps, less than enthusiastic about drawing.

It should be noted that the researcher supplied all of the pencils and test booklets for the test administrations in order to control for any confounding variables such as pencil hardness, pencil quality, etcetera. It was deemed significant that the items be exactly the same to avoid any adverse effects on the quality of the drawings (i.e., all pencils were the same type, colour, length, hardness, etc.). To thank the children for their participation the researcher distributed stickers appropriate to the varying age levels of the children at the conclusion of the drawing session.

Test Administration

Phase one data collection began in March 2007. The school had offered the use of their Outside School Hours Care room to the researcher to conduct data collection in. This room was not used during school hours, only before and after school. It was an air-conditioned, quiet room which contained enough appropriately-sized table and chairs to comfortably seat approximately 15 children at any one time.

To help ensure method consistency, as well as valid and reliable data collection, each appointment for test administration followed a set format of proceedings. For each data collection appointment, the researcher arrived approximately 15 to 20 minutes earlier than the negotiated time and used this period to set up the necessary apparatus for data collection (e.g., sharpened pencils, arranged test booklets, pencils and erasers on the desk top, etcetera). Methodological and extensive efforts were made on the researcher's behalf to ensure that all data collection sessions followed a similar format, and that all children felt safe, comfortable, calm and positive throughout the entire procedure.

Test administration involved each participant sitting individually at a desk designed for use by children of the particular age group. Each participant was given one GHDT Booklet, a number two pencil, and access to an eraser. Following the procedure outlined in the GHDT manual (Harris, 1963) children were asked to complete the cover page of the test booklet, which requires each child's name, sex, age, date of birth, and school details. Children in prep and year one and children with special needs were assisted with this as necessary.

Adhering to the script in the test manual, participants were then asked to draw the figures of a man, a woman and themselves individually and systematically on separate pages in the test booklet using the number two pencil (Harris, 1963). To elucidate this process, an excerpt of the script is reproduced below:

> I am going to ask you to make three pictures for me today. We will make them one at a time. On this first page I want you to make a picture of a man. Make the very best picture that you can; take your time and work very carefully…. Be sure to make the whole man, not just his head and shoulders. (Harris, 1963, p. 240)

It should be noted that the researcher was conscious of simply 'reading' the script and sounding artificial. Instead, key parts of the script were memorized and communicated at the appropriate time in a natural tone and relaxed manner. It is widely recognized that tense, contrived, and demanding assessment conditions are not conducive to reliable test results and valid examples of children's abilities (Nadelman, 2004). Thus, careful attention was paid to the delivery of the test procedure.

As recommended by Goodenough (1926) and Harris (1963), words of praise and encouragement were spoken during test administration to assist in maintaining interest and motivation whilst the children were drawing. An excerpt from the GHDT administration manual offers an example: "These drawings are very [good]; you boys and girls are doing very well" (Harris, 1963, p. 241). The word 'good' was used instead of the original term 'fine' as the researcher believed the former term was more socially and culturally appropriate to children today. Any adverse comments, criticisms or suggestions were withheld to avoid influencing the nature of the drawings. Any specific questions from participants, or requests for further instructions were responded to with the applicable response outlined in the manual: "Do it whatever way you think is best" (Harris, 1963, p. 241). Similarly, the participants were reminded to refrain from announcing their own suggestions, holding up their work, or providing ‘a running commentary’ of their drawing, e.g., “I’m giving my man a soldier hat” (Harris, 1963, p. 241). Harris (1963) provides an appropriate response that dispels such incidents without affecting participant interest or enthusiasm: “No one must tell about his [or her] picture now. Wait until everybody has finished” (p. 241).

The GHDT has no time limit, thus test administration ceased when each child had completed all three drawings. Participants under the age of eight were given a short rest between their drawings of a woman and the self-portrait. Harris (1963) suggests young children put their pencils down and stretch their arms and fingers to help them relax from the tension created by their concentration and effort.

As recommended by Harris (1963), once test administration was complete individual informal questioning took place, where necessary, to clarify any ambiguous aspects of participants’ drawings. The researcher took considerable care when framing questions to ensure that any assumed answers were not inadvertently suggested in the question itself. Anecdotal notes were written by the researcher directly on the participant’s test booklet in the space provided for ‘Notes’.

As each child finished their drawings they notified the researcher by raising their hand. The researcher then informally reviewed the test booklet to make sure that all essential information was completed on the cover page, and that any ambiguous aspects in the drawings were questioned and / or clarified, if necessary. After this, the test booklet was collected and the child was free to choose a sticker and, in the cases of small group administrations, a book to read whilst other children completed their drawings.

Class disturbance was minimized as data was collected during dedicated appointment times only, and test duration was kept to the minimum necessary for allowing accurate and thorough test administration. To help avoid any outside influence on future participants' performance, participants were encouraged to refrain from discussing the details of the test, or their drawings, with their peers.

Longitudinal Data

To examine developmental changes in the children's human figure drawings, two subsequent data collection phases were planned to follow phase one. These subsequent data collection phases were scheduled approximately six months apart. Therefore, the GHDT was administered to the sample three times over an approximate 12 month period.

Phase two and three GHDT administrations and data collection commenced in October, 2007 and March 2008 respectively. Each followed the same format of proceedings as those established in Phase one. The researcher continued to arrive approximately 15 to 20 minutes ahead of the time negotiated with class teachers, and set up the room in the method previously described. Of course, disturbance to participating children's learning experiences was minimized as test duration was kept to the minimum necessary for allowing accurate and thorough test administration. Again, participants were encouraged to refrain from discussing the details of the test, or their drawing, with their peers to help avoid any outside influence on future participants' performance.

Scoring Procedure

All participants drawing attempts were collected, examined and scored in accordance with the GHDT scoring system (Harris, 1963). As previously discussed, the GHDT comprises two scoring guides with 73 items for the DAM scoring scale, and 71 items for the DAW scoring guide (see Appendix A). The projective self-portrait has no corresponding scoring guide; thus, self-portraits were scored using the DAM or DAW scoring guide in correspondence with the sex of the participant (i.e., DAW scoring guide was used for self portraits drawn by girls, and correspondingly, the DAM scoring guide was used for self portraits drawn by boys).

Scoring Criteria

Harris (1963) affirms that the test can be scored by any person who is "capable of following instructions faithfully" (p. 242). Scoring the GHDT is not difficult; however it does require study, patience, practice and a willingness to follow the instructions meticulously (Harris, 1963). Written criteria are provided for each item in the scoring guide, and most items also include a drawn example.

In accordance with the GHDT manual, each item was scored dichotomously as either: 'absent' if the item was absent from the drawing, or present but not within the item's scoring criteria with zero (0) credit awarded; or 'present' if the item was present in the drawing and adhered to the item's scoring criteria with one (1) credit awarded. The GHDT test booklet has a scoring template beside each of the spaces in which children draw (see example of test booklet in Appendix A). The researcher recorded the applicable score for each item by placing either a '-' (dash) or a '1' (one) beside the item number to indicate an 'absent' or 'present' item respectively. Half credits are not awarded (Harris, 1969).

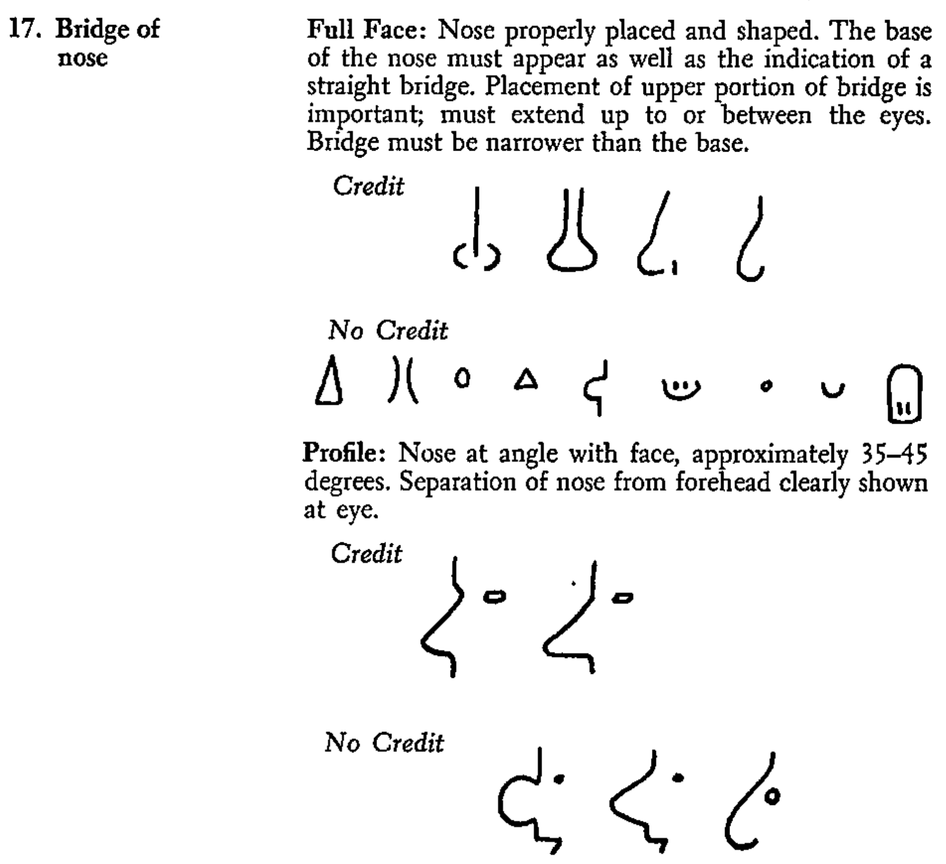
Some items in both the DAM and DAW scoring guides are considered subjective (Goodenough, 1926a; Harris, 1963). Harris (1963) asserted that absolute agreement between two scorers cannot, and should not, be expected. Harris (1963) stated that inter-scorer agreement is high, however, if both are experienced with scoring the GHDT. If uncertainty arose in the scoring of the drawings, the researcher carefully reviewed the scoring criteria and examples for the applicable item.

To elucidate further the scoring process, an excerpt is provided which details the criteria for Item 17 'Bridge of Nose' (see Figure 5) from the Draw-a-Man Scale:

> Full Face: nose properly placed and shaped. The base of the nose must appear as well as the indication of a straight bridge. Placement of upper portion of bridge is important; must extend up to or between the eyes. Bridge must be narrower than the base.
>
> Profile: nose at angle with face, approximately 35-45 degrees. Separation of nose from forehead clearly shown at eye. (Harris, 1963, p. 251)

Figure 3.1

Example from the Draw-a-Man Scoring Scale of Item 17: Bridge of Nose



17. Bridge of nose

Full Face: Nose properly placed and shaped. The base of the nose must appear as well as the indication of a straight bridge. Placement of upper portion of bridge is important; must extend up to or between the eyes. Bridge must be narrower than the base.

Credit

No Credit

Profile: Nose at angle with face, approximately 35–45 degrees. Separation of nose from forehead clearly shown at eye.

Credit

No Credit

Note. From “Children’s Drawings as Measures of Intellectual Maturity” by Dale B. Harris, 1963, p. 251.

Clearly, some subjective judgement is necessary in establishing whether an item is awarded a credit. In regards to Item 17 mentioned above, perceptions of a “properly placed and shaped [nose]” (Harris, 1963, p. 251) would undoubtedly vary amongst scorers. As recommended by Harris (1963), all data in this study was scored by the researcher (i.e., one person); thus, scorer consistency helped to reduce any ‘random error’.

All credits awarded for a drawing were added together to produce a *total* raw *score*. When adding credits to produce a total raw score for each drawing it is not permissible to combine scores from different drawings for the total score, nor is it appropriate to combine the scores from the better-drawn features in different drawings: each feature in each drawing is scored on its own merits.

Calculating Standard Scores, Average Scores, and Percentile Ranks

As discussed in the Literature Review, after all three drawings by each participant had been examined and scored the standard scoring procedure involved calculating the participant's raw, standard, and average scores, and corresponding percentile rank. In this research, only the participants' raw scores were calculated. The reasons for this are discussed below.

The sum of credits awarded for each of the three drawings is deemed to be the participants' *raw score.* Thus, each participant received three raw scores, one for each drawing of a man, a woman, and themselves. These raw scores would be used, together with other values, to calculate the equivalent *standard score, average score*, and *percentile rank.* Standard scores were determined by combining a child's raw score with his/her age (in years and months). A standard score expresses each child's relative standing on the test in regards to his/her own age and sex group, in terms of a mean of 100 and a standard deviation of 15 (Harris, 1963). As well as determining participants' standard score for each drawing, average scores were also established. By summing the standard scores for a child's DAM and DAW drawings and dividing that number by two, an average measure can be determined. Percentile ranks, which show "the relative standing of a child in a theoretical group of 100" (Harris, 1963, p. 311), were determined by looking up a child's standard score in the tables presented in the GHDT manual (see page 311 in Harris, 1963).

As discussed in the Literature Review, modern test theory refutes the use of raw, standard, and average scores as well as percentile ranks as 'measures' of development for several reasons. Most prominently, the scores, or 'measures' as they are often inappropriately referred to as, are based on nothing more than a count of the credits. Moreover, these 'counts' mistakenly assume that every item on the point scale is equal in value, or difficulty. The counts are treated as though they are equal intervals units, like those on a centimetre ruler or thermometer, when there has been no examination or analysis to establish that that is the case. For example, mere 'counts' assume that it is equally difficult to include a head – the most commonly endorsed item in a human figure drawing – as it is to include knees and elbows – amongst the least endorsed items in a human figure drawing – where this is evidently not the case. Every single child in this study included an item creditable as a 'head', whilst comparatively fewer children include items creditable as 'elbows' or 'knees'; clearly, these items are not equal in creditable value, or difficulty. Thus, the standard and average scores, and the corresponding percentile ranks were not used in this thesis.

Data Preparation

As discussed in Chapter Two, one of the two theoretical perspectives adopted in this research was Latent Trait Theory (LTT). Developed on the theoretical basis of LTT, and well-suited to this research, is the Rasch model for measurement. In order to apply the Rasch model to the data in this project, a suitable tool for data exploration needed to be sought. WINSTEPS® (Linacre, 2009a) is a well-known and highly-regarded computer software program based on Rasch model principles. The program was employed as the data analysis tool for this research. Before the data could be submitted to WINSTEPS®, however, it needed to be organized into a format the computer program would recognize.

Dichotomous Data and the Data Lines

The nature of the data meant that scoring the qualitative drawings was straight forward as it was completed using a dichotomous (i.e., two-category 0, 1) system and the Microsoft Excel® software. If the item was absent in the drawing or, perhaps, present but not meeting the item's criteria, zero (0) credit was assigned. Conversely, if the item was present in the drawing and attained to the item's criteria, one (1) credit was assigned. This was in accordance with the principles and procedures espoused by Goodenough (1926a) and Harris (1963). The scoring of each drawing was recorded on the test booklets in the space provided to the right of each drawing box (see Appendix A).

Each data line was constructed by reviewing each of the test booklets. First, the applicable ID code for the child (which started at 100, followed by 101, 102, 103, etc.) was entered into the Excel® spreadsheet. Next, the participant's gender, age and year level were obtained from the front of the test booklet and entered into the data line to complete the construction of the unique ID. The entering of the DAM, DAW and SP raw data followed this. The researcher examined the applicable test booklet and used the codes of '0' and '1' to represent 'no credit' and 'credit' for each item accordingly.

The length of the data lines for male and female children varied, because of the nature of the self-portrait item whereby the scoring scale used to score the drawing is dependent on the gender of the child. Therefore, the data lines of female participants contained 215 items and the data lines of male participants contained 217 items (see both examples below).

The data line for females contained the following types of data: the first three digits indicate the participant's unique identifying code (beginning at 100 through to 207); followed by an 'F' to indicate that the child is female; the next two digits indicate the child's grade and age respectively; the following 73 digits represent the child's score on each of the DAM items; the next 71 digits indicate the child's score on each of the DAW items; and the final digits represent the 71 SPW items (see example below).

107 F 38 11101111010010001100001000101111011001111111110111111110000001 1
1010000001111111010001110101100011101000110111111111101100010001101 1100
1100000001110101000001100110000000100111110001000111000000000000011110000
1000000000

The data line for males contained the following types of data: the first three digits indicate the participant's unique identifying code (beginning at 100 through to 207); followed by an 'M' to indicate that the child is male; the next two digits indicate the child's grade and age respectively; the following 73 digits represent the child's score on each of the DAM items; the next 71 digits indicate the child's score on each of the DAW items; and the final digits represent the 73 SPM items (see example below).

103 M 38 1001010010100000001000001100001000010001000010110000000110000000
0000000000000101001000100010100001000110001110000010100000000000001101100
0000000000100101001010000000100010110000100001000100001011001000110000000
000000000000

Each data line represents the transformation of qualitative data to quantitative data, which can then be subjected to Rasch analysis. The individual data lines were compiled in an Excel® spreadsheet which, collectively, became the data file.

The Data Files

The construction of the main data files was completed by simply working systematically through the pile of examined and scored test booklets from each data collection phase. Within the data files, each data line (or row in the spreadsheet) represents a child, or 'case', and each column represents an item from a GHDT sub-test. The data file for phase one, for example, contained 107 rows and a total of 217 columns that contained all participating children's performances on DAM, DAW, SPM, and SPW sub-tests of the GHDT. The data files for phases two and three were smaller in size as the number of research participants decreased.

There were three main data files; one for each phase of data collection, and each of these data files contained a suite of four worksheets, one for each sub-test (DAM, DAW, SPM and SPW).

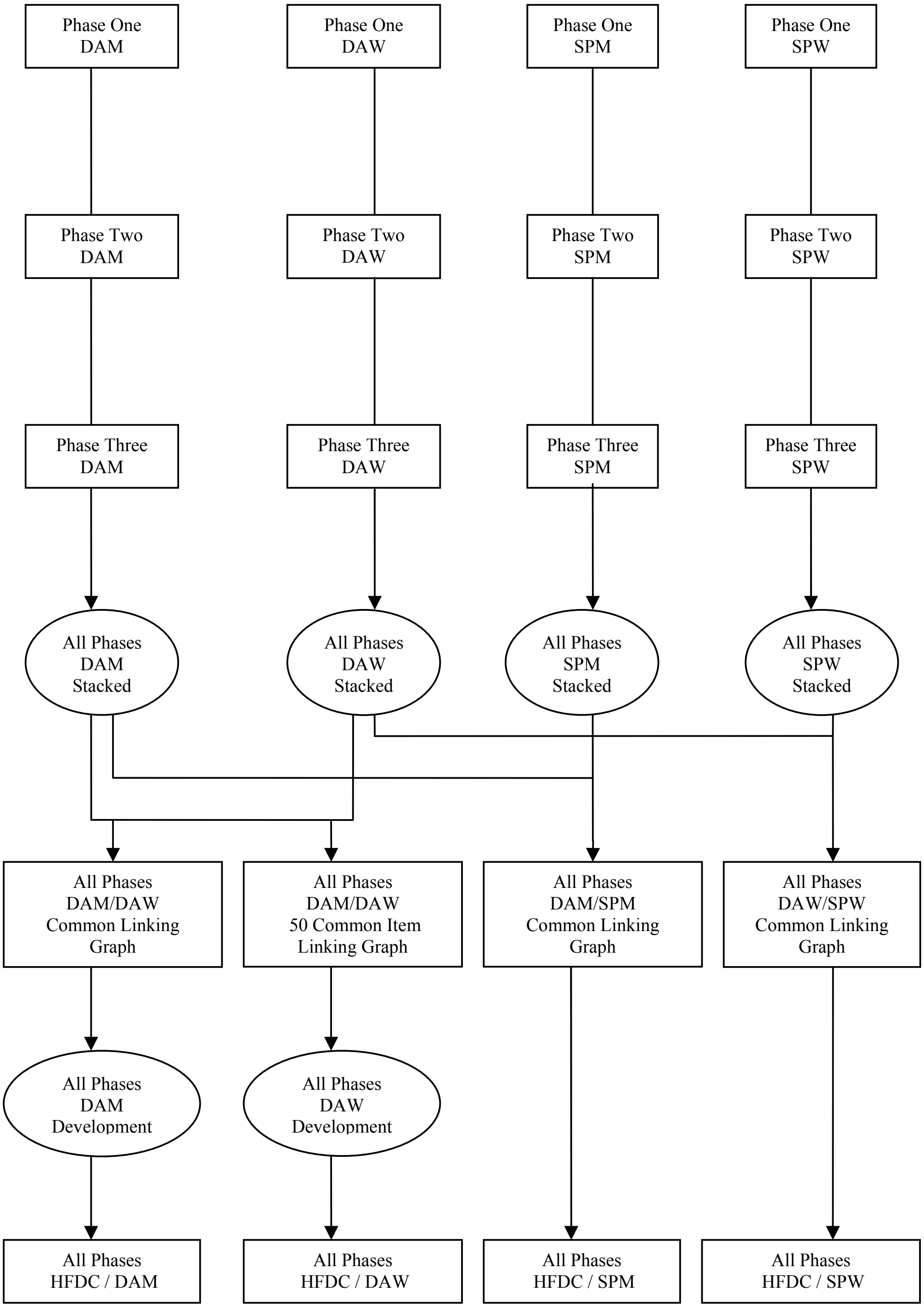
Rasch Analysis

The above described data files were submitted to WINSTEPS® version 3.68.0 (Linacre, 2009a). This version of WINSTEPS® has a data capacity of 10, 000, 000 persons and 30,000 items, with each item being able to have a rating scale with up to 255 categories (Linacre, 2009b). Therefore the data file constructed for this project was clearly within the capabilities of the program.

Using Excel® software, the data files were manipulated so that different parts of the data could be analysed separately or independently of the other data, and also in meaningful configurations. This produced a range of Rasch output which could be used together with Excel® software to construct developmental graphs, and common linking plots. The following diagram illustrates the sequence of data analyses undertaken.

Figure 3.2

Overview of Rasch Analyses Plan



Separate & Complete Data Set Rasch Analyses

The DAM, DAW, SPM and SPW data from phases one, two and three were analysed separately initially. The separate analyses of the data collection from the DAM, DAW, SPM, and SPW sub-tests from each phase of data collection enabled close examination of the underlying psychometric properties of each of the sub-tests which comprise the GHDT. Particular attention was paid to test unidimensionality, item fit statistics, and person fit statistics across the phases within each sub-test. These initial analyses were highly anticipated as they were the very first 'look' at the GHDT from a modern test theory perspective. These separate analyses were followed by the analyses of the data from all phases combined for each of the sub-tests.

To examine each individual sub-test more rigorously, the data from all three phases of data collection were 'stacked' within a modified data file for each sub-test. For example, the DAM sub-test data from phase one, was 'stacked' on top of that received from phases two and three. The ID codes for the children were modified so that, for example, case ID 109M became 109M1, 109M2, and 109M3 with the suffix of 1, 2, and 3 indicating the phase in which the data was gathered. This meant that the sample sizes from each phase were combined to create a 'unique' sample of 246 children (i.e. the total of the samples from phase one n = 107, phase two n = 83, and phase three n = 56). This process was repeated for the DAW, SPM and SPW data. Indeed, Rasch analysis of the three phases of data collected for each sub-test is a more stringent test of unidimensionality than that of analyzing each data set individually.

Developmental Graphs and Common Linking Plots

Another advantage of completing these 'stacked' analyses was the production of key information which could be used in Excel® spreadsheets to construct developmental graphs and common linking plots. The graphs displayed diagrammatically the development of the children's drawing in each sub-test which the Rasch model tracked quantitatively. The plots revealed significant insights on the children's performances across the different sub-tests, and offered evidence of which sub-test might be considered most appropriate for general use.

Analysis of the Human Figure Drawing Continuum

The information revealed by the multiple Rasch analyses culminated in the modification of the GHDT to form, what this researcher terms, the prototype 'Human Figure Drawing Continuum' (HFDC). The HFDC was Rasch analysed and the output was used to construct common linking graphs in Excel® to establish whether it, or the conventional GHDT sub-tests, was most effective in revealing insights about young children's human figure drawings.

This research, and the resultant HFDC, would be the first formal attempt at updating the GHDT since Harris's revision and extension in 1963. All of the Rasch analyses were undertaken in an effort to modify the GHDT to better align the instrument with today's social, cultural and educational dimensions.

Ethics

Any research involving human participants must adhere to particular ethical guidelines to ensure participant wellbeing (Nadelman, 2004). Research projects that involve human participants in Australia must adhere to the *National Statement on Ethical Conduct in Human Research* (2007) developed by National Health and Medical Research Council, Australian Research Council, and Australian Vice-Chancellors' Committee.

In addition to adhering to the values and principles discussed in the National Statement, researchers at James Cook University must also act in accordance with the *James Cook University Statement and Guidelines on Research Practice* (draft) (James Cook University Research Committee, n.d.). This James Cook University-specific policy is based upon the more general *Australian Code for the Responsible Conduct of Research* (National Health and Medical Research Council, Australian Research Council & Universities Australia, 2007). This researcher adhered to the ethos, values, principles, and codes outlined in all three of the above documents to ensure participant wellbeing.

The research aims, sampling methods, and data collection procedures of this research were outlined in a Human Ethics Application and submitted to the James Cook University Ethics Committee. This application was approved without amendment on 10th October 2006 (Approval number H2450, see Appendix C).

All signed Informed Consent Forms and completed GHDT test booklets were kept in a locked cabinet in the researcher's office at the School of Education, James Cook University. This cabinet was only accessible by the researcher. This documentation will be retained for a period of at least five years, after which, it will be destroyed.

Limitations

The limitations to this research include the following issues: sample, scope, and inter-rater reliability. Whilst random selection of participants from a range of school sites would have provided a more representative sample, it was not a feasible option for this study. The twelve month data collection timeframe and need for three data collection phases each approximately six months apart meant that including multiple school sites was not practicable or manageable for a single researcher. However, as explained elsewhere in this chapter, the school selected was deemed to have a student community which was reasonably representative of the larger population.

Whilst it would have been beneficial to have included drawings from children from diverse cultures and backgrounds (such as indigenous students, students with special needs, and students from other countries / cultures), this was also not a feasible option. As the GHDT has not yet been examined from a modern test theory perspective, the researcher believed that the foremost task was to investigate whether it was actually possible to apply the Rasch model to young children's human figure drawings before adding other variables to the research project. In other words, to include multiple other variables at this time (i.e. the very first Rasch analysis of the GHDT) was considered to be confounding to the issue of whether the GHDT was actually apt for Rasch analysis.

A study of inter-rater reliability on the GHDT would have enhanced this research. However, a suitable research assistant in the locality of the project was not found. In any case, the issue of inter-rater reliability on the GHDT is something which has already been investigated by Goodenough (1926a) and Harris (1963), and both found that there was minimal discrepancy in the scoring completed by people experienced in scoring the GHDT. Although, a more recent investigation of this issue would be worthwhile and could easily be conducted in the future using the data collected for this project.

The results produced from the application of the Rasch model for measurement to the GHDT are presented and described in Chapter Four.

Chapter Four

Results

The primary aim of this study was the examination of the well-established and widely used Goodenough-Harris Drawing Test (GHDT) using analyses performed using a key model of modern test theory – the Rasch model. To produce the most thorough representation of the GHDT's performance according to the Rasch model each phase of data *and* each sub-test of the GHDT was analysed individually and in meaningful combinations. The analyses were completed using WINSTEPS® (Linacre, 2009a) software.

The results produced by the Rasch analysis of the DAM sub-test are presented first, followed by the DAW, SPM and the SPW sub-test results. The results from the analyses of each sub-test are presented in a systematic order to enable meaningful comparisons across each of the data collection phases. Therefore, each sub-test section includes the variable maps, summary statistics, item and person statistics produced by the Rasch analysis of each phase separately, followed by some examples of the children's drawings from each phase. Each sub-test section concludes with some results produced from the Rasch analysis of the stacked data set.

After the results for each sub-test are described, the human figure drawing development graphs and common linking plots are presented. Lastly, the results from the comparison of the prototype Human Figure Drawing Continuum (HFDC) and the GHDT sub-tests are described.

Draw-a-Man Sub-Test Results

The DAM sub-test was, generally, the first drawing made by the children in the test administration. This drawing was completed by both male and female research participants (unlike the SPM and SPW sub-tests which were completed by only males and only females respectively). The drawings of men were scored using the 73 item DAM scoring guide.

The variable maps, summary statistics, and item and person statistics produced by the separate Rasch analysis of the DAM sub-test in each phase are presented first. These are followed by some examples of the children's drawings from phases one, two and three which are displayed in Figures 4.4, 4.5, and 4.6 respectively. This DAM section concludes with the key results from the stacked data analysis.

Variable Maps

Figure 4.1 presents the variable map produced by the Rasch analysis of the phase one DAM data. The map displays the estimated locations for the 73 DAM items and the 107 children included in phase one of the study. Item and person locations on this logit scale represent the item difficulty estimates and person ability estimates which are displayed in Tables 4.2 and 4.3. Routinely, a logit value of 0 (zero) is used set as the mean item difficulty, so items of above-average difficulty and persons of ability above that are plotted as positive, and items of below-average difficulty and persons of lower ability are plotted as negative. The person-item variable maps produced by WINSTEPS® display the logit scale down the middle of the map, with the GHDT item difficulty estimates plotted to the right-hand side and the children's ability estimates plotted to the left-hand side. An item difficulty threshold is estimated at the point on the logit scale where a person of that estimated ability has a 50% probability of succeeding on that item. As all of the items in the GHDT are scored dichotomously, each has one item difficulty estimate plotted at the threshold where the probability of scoring either 0 or 1 is 50% (Bond & Fox, 2007). For example, case ID 111M, located at -1.01 logits (see Figure 4.1), would most likely have received credit for items 28, 51 and all of those below his ability estimate, and he would have a 50% probability of receiving credit for DAM item 64 which is in almost complete alignment with his estimated ability. However, case ID 111M most likely did not receive credit for items 19, 36, 6 and others located *above* his location on the logit scale.

Figure 4.1

Phase One Draw-a-Man Person-Item Variable Map

```
                                 Persons -MAP- Items
                                     <more>|<rare>
 5                                         +  59   60   61   69   71   73
                                           |T
                                           |
                                           |
                                           |  38
 4                                         +
                                           |
                                           |  21   37   58   70   72
                                           |
                                           |  13   66
 3                                         +
                                           |  12
                                           |  34   62
                                           |S 26   57   68
                                           |  16
 2                                         +
                                     177F  |  15   20   52
                                     174F T|  27   32
                                     100F  |  23
                                     150M  |  17   65
 1                    109M 108M 107F 196F  +  5    7
                           102M 113F 122F  |  42
                           166F 201F 204F  |  41   49   67   8
            112F 116F 125F 155F 193F 206F  |  29   56
                                     195F S|  22   25   31   43
 0                         120F 124F 146F  +M
            101F 121F 156F 167F 175F 192M  |  40   48
                      153F 172M 179M 205F  |  10   3    33
                 105F 119M 154F 180M 202M  |
                 126F 145F 151F 173M 181M  |  19   36   6
-1                                   111M  +  64
            127F 147F 141F 157F 178M 199F  |  28   51
                                128M 135F M|
  104F 106M 110M 117M 138F 132M 165F 182M  |  54
                              11 12 17 20  |  2    9
-2 114M 118M 129M 133F 134M 136F 159F 161F 163M 183F 184M 194F 197M  +  24   45   63
                           158F 171F 198M  |  14
            103M 139F 140M 142M 148M 152F  |S
                 137M 164F 176M 185F 186M  |  39   53
                           149M 169F 190F S|  18   50   55
-3                                         +
                           162F 168F 187F  |  47
                                143M 144M  |
                           131M 188M 200M  |
                                           |
-4                                         +
                                           |
                                          T|  11
                                     189M  |
                                           |T 44
-5                              160F 191F  +  30
                                           |  46
                                           |
                                           |
                                           |
-6                                   130M  +
                                           |
                                           |  35
                                           |
                                           |
-7                                         +  1    4
                                     <less>|<frequ>
```

The DAM sub-test item difficulties for phase one cover an expansive range of around 12 logits, from -7 to +5 logits. The DAM items were dispersed quite well along this continuum, with some items targeting the bottom of the logit scale, a greater concentration of items through the middle, and several plotted at the top-end of the scale. Clearly, items 59, 60, 61, 69, 71 and 73 (clothing V, profile I, profile II, directed lines and form: facial features, "modeling" technique, and leg movement, respectively) plotted at 5.50 logits are the most difficult items for these children to include in a drawing of a man. In fact, none of the children in phase one received credit for any of these 6 extreme items (see Table 4.2). The DAM items which were least difficult to include in a drawing of a man were items 1 (head present) and 4 (eyes present) located at -8.46 logits. All 107 research participants received credit for both these items (see Table 4.2).

Evidently, the DAM sub-test is quite difficult for this sample. The person range was less than 8 logits towards the lower end of the 12 logits item range. The most successful child, case ID 177F, plotted at +1.76 logits is situated 3 whole logits from the top of the logit scale (where the most difficult items are located). The least successful person was case ID 130M, who was plotted at -6.04 on the logit scale. He was not completely unsuccessful in receiving credit for his drawing of a man; he was credited for items 1 (head present), 4 (eyes present), and 35 (legs present), and 46 (trunk present) (see Figure 4.4).

The variable maps produced from the Rasch analysis of the data collection from the DAM sub-test in phases two and three are displayed in Figures 4.2 and 4.3 respectively. These maps all show remarkable similarity to one another. The logit scales in Figures 4.2 and 4.3 span some 12 logits from -7 to +5 for phase two, and 11 logits from -7 to +4 for phase three. Also, the variable maps for phases two and three show similarly good dispersion of the DAM items along the logit scale as seen in Figure 4.1. Indeed, most items were plotted at remarkably similar locations across all three DAM variable maps.

The case distributions on the left-hand side of each of the logit scales in Figures 4.1, 4.2, and 4.3 are also comparable. Case IDs 100F, 112F, 150FM 174F, and 177F appear amongst the top performing children across all phases of the DAM sub-test. At the other end of the logit scales case IDs 130M, 131M, 143M, 168F, and 188M were located amongst the least successful on the DAM sub-test for all phases.

Rasch analysis variable maps present the relations between person ability and item difficulty only. Error estimates, fit statistics, and the corresponding reliabilities are displayed in other output documents produced by the Rasch model analysis. The summary statistics, item fit statistics, and person fit statistics, all contain crucial information that informs the researcher about data performance according to the Rasch model's expectations (Bond & Fox, 2007; Wright & Stone, 1979). Tables 4.1 through 4.3 present the phase one DAM summary statistics, Tables 4.4 to 4.6 display the item fit statistics, and Tables 4.7 to 4.9 present the person fit statistics for phases one to three correspondingly.

Figure 4.2

Phase Two Draw-a-Man Person-Item Variable Map

```
                           Persons -MAP- Items
                              <more>|<rare>
 5                                  +T 13 21 59 61 62 69 71 73
                                    |
                                    |
                                    |
                                    |
 4                                  +  12 37 58 60
                                    |
                                    |
                                    |  66 72
                                    |
 3                                  +
                                    |  34 38 68 70
                                    |  20
                                    |S
                                    |  26 42 57
 2                                  +  15 27
                           112 174  |
                                   T|  52 65
                                    |  23
                           125 177  |  32 56 7
 1                         196 204  +  16 49
                               205  |  41
                           101 156  |  5
                           166 167  |
                   121 123 150 172 S|  17 3
 0 113 116 120 122 126 155 192 193  +M 8
                           124 175  |  19 22 25 31 40 43
                   146 151 179 180  |  10 48
                   170 173 178 202  |  36
                           181 199  |  29 33 64 67
-1                         114 147  +  51
                   118 135 148 194 M|
                   110 117 152 154  |  28 6  63
       103 115 136 158 165 171 182  |  45
                           137 153  |  54
-2                 127 132 141 163  +  14
                   133 157 186 187  |  24 53
       129 139 149 160 161 169 176  |S 47
           128 168 185 188 189 190 S|  39 50 9
                       142 159 191  |
-3                             184  +  55
                                    |
                               119  |  2
                               143  |
                                    |  18
-4                                 T+
                                    |
                                    |
                                    |
                               131  |
-5                                  +T
                                    |
                               130  |  11
                                    |
                                    |
-6                                  +
                                    |
                                    |
                                    |  4  46
                                    |
-7                                  +  1  30 35 44
                              <less>|<frequ>
```

Figure 4.3

Phase Three Draw-a-Man Person-Item Variable Map

```
                                     Persons -MAP- Items
                                         <more>|<rare>
    4                                          +  21  59  60  61
                                               |  38  58  66  73
                                               |
                                               |
                                               |
    3                                          +  12  13  37  62  69  71
                                               |
                                               |  34
                                          100  |
                                               |S 57
    2                                          +  15  16  72
                                               |
                                              T|  41  68  70
                                     150  177  |  52
                                               |  26
    1                           112  174  179  +  20  23  27  42  65
                                          193  |  49
                                               |  17  3   32
                                167  192  196  |  7
                                          178 S|
    0                                     172  +M 43  48  5
                           121  125  166  175  |  25  8
                           122  123  195  202  |  22  45  56
                                101  116  180  |  31
                                               |
   -1                                139  199  +  29  40
                           110  118  135  164 M|  10  28  33
                                          141  |  51
                                          194  |  2   36  6   64
                           120  140  147  173  |  54
   -2  115  152  159  171  176  182  184  185  +
                                               |S 50  63  67  9
                                129  163  165  |
                                          168 S|  19  53
            131  143  157  160  188  189  191  |  24  55
   -3                                     169  +
                                               |
                                               |  14
                                               |  39
                                               |
   -4                                         T+
                                               |  47
                                               |
                                               |T
                                               |  18
   -5                                     186  +
                                               |
                                               |
                                               |
                                               |
   -6                                          +  11
                                               |
                                               |
                                               |
                                               |
   -7                                          +  1   30  35  4   44  46
                                         <less>|<frequ>
```

Summary Statistics

Table 4.1 presents the summary statistics for the Rasch analysis of the phase one DAM data. The second half of Table 4.1 presents the output data for the 65 items, out of a possible total of 73 items from the DAM scoring guide, that were included in the analysis. Two items (items 1 and 4) were excluded from the analysis as all 107 persons received credit for them, and another 6 items (59, 60, 61, 69, 71 and 73) were excluded as no one in this sample received credit for them. The mean of the item estimates was located at 0 (by default), and the standard deviation was 2.43. The reliability of these item estimates was high at .97. This latter result refers to the ability of the DAM sub-test to define a hierarchy of indicators along the equal-interval scale. Bond and Fox (2001) assert that "[t]he higher the number, the more confidence we can place in the replicability of item placement across other samples" (p.46). Thus, a DAM item reliability index of .97 indicates that researcher could expect the order of the DAM items along the logit scale to be replicated when the test is administered to other appropriate samples.

The top of Table 4.1 contains output data pertinent to the 107 measured persons in phase one data collection. This summary reveals that the mean raw score for the 107 participants was 22.1 out of a maximum of 73 possible credits. Out of a total of 73 items on the DAM scale, only 65 were actually included in this analysis (due to the 8 extreme items omitted). This information, together with the person mean of -1.30, indicates that this test was a little difficult for this sample. A mean person ability estimate closer to 0 (zero) would to indicate that the difficulty level of the test was well-matched to the ability level of the sample (Bond & Fox, 2007). The standard deviation for the persons measured was 1.50, indicating somewhat less spread than that reported for the DAM items (SD = 2.43). The reliability of the person estimates was quite high at .92. A high person reliability index indicates that the person ability estimates would likely be replicated if the sample were given a similar test (Wright & Masters, 1982).

Tables 4.2 and 4.3 display the summary statistics produced by the Rasch analysis of the DAM sub-test data collected in phases two and three respectively. Again, remarkable similarity can be seen amongst the results produced for each phase. Review of the lower halves of Tables 4.1, 4.2 and 4.3 reveal comparable summary statistics for the DAM items. The number of items measurable items was 65, 61 and 63 across phases one, two and three respectively. Based on the children's responses the mean raw score for the items in each phase progressively lessened over phases one (MRS = 36.4), two (MRS= 28.9) and three (MRS = 20). This could be linked to the decreasing sample size suggesting, perhaps, that disproportionately more of the successful drawers were absent from phases two and three. Conversely, the standard deviations across the three phases were remarkably similar: phase one SD = 2.43, phase two SD = 2.46, and phase three SD = 2.27. Also, the item reliability index produced in phase one (r = .92) corresponds well with those produced in phases two (r = .96) and three (r = .94).

Table 4.1

Phase One Draw-a-Man Sub-Test Summary Statistics

SUMMARY OF 107 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	22.1	65.0	-1.30	.39	1.01	.0	1.06	.1
S.D.	10.2	.0	1.50	.08	.25	1.0	1.42	1.1
MAX.	46.0	65.0	1.76	.85	2.33	2.5	9.90	5.6
MIN.	2.0	65.0	-6.04	.34	.55	-2.1	.06	-1.2

REAL RMSE	.43	ADJ.SD	1.44	SEPARATION	3.37	Person RELIABILITY .92
MODEL RMSE	.40	ADJ.SD	1.44	SEPARATION	3.59	Person RELIABILITY .93

S.E. OF Person MEAN = .15

Person RAW SCORE-TO-MEASURE CORRELATION = .98
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .93

SUMMARY OF 65 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	36.4	107.0	.00	.36	.99	.0	1.20	.0
S.D.	31.0	.0	2.43	.17	.19	1.4	1.49	1.5
MAX.	105.0	107.0	4.28	1.01	1.52	3.7	9.90	6.1
MIN.	1.0	107.0	-6.40	.23	.64	-3.8	.21	-2.8

REAL RMSE	.41	ADJ.SD	2.39	SEPARATION	5.81	Item RELIABILITY .97
MODEL RMSE	.40	ADJ.SD	2.40	SEPARATION	5.95	Item RELIABILITY .97

S.E. OF Item MEAN = .30

MAXIMUM EXTREME SCORE: 6 Items
MINIMUM EXTREME SCORE: 2 Items
UMEAN=.000 USCALE=1.000

Table 4.2

Phase Two Draw-a-Man sub-Test Summary Statistics

```
SUMMARY OF 83 MEASURED Persons
+-----------------------------------------------------------------------------+
|           RAW                            MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      21.2      61.0       -1.21     .40      1.00     .0    .91     .0 |
| S.D.       9.5        .0        1.40     .07       .21     .9    .65     .7 |
| MAX.      43.0      61.0        1.72     .82      1.81    2.7   4.81    2.4 |
| MIN.       3.0      61.0       -5.37     .35       .65   -2.0    .33   -1.0 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .42  ADJ.SD    1.34  SEPARATION  3.15  Person RELIABILITY  .91 |
|MODEL RMSE    .40  ADJ.SD    1.34  SEPARATION  3.33  Person RELIABILITY  .92 |
| S.E. OF Person MEAN = .15                                                   |
+-----------------------------------------------------------------------------+
DELETED:     24 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .00

SUMMARY OF 61 MEASURED Items
+-----------------------------------------------------------------------------+
|           RAW                            MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      28.9      83.0         .00     .42       .99     .0    .91    -.1 |
| S.D.      23.8        .0        2.46     .23       .22    1.5    .66    1.3 |
| MAX.      82.0      83.0        4.04    1.04      1.57    3.6   3.61    5.0 |
| MIN.       1.0      83.0       -6.62     .26       .56   -4.5    .05   -3.2 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .50  ADJ.SD    2.41  SEPARATION  4.87  Item   RELIABILITY  .96 |
|MODEL RMSE    .48  ADJ.SD    2.42  SEPARATION  4.98  Item   RELIABILITY  .96 |
| S.E. OF Item MEAN = .32                                                     |
+-----------------------------------------------------------------------------+
MAXIMUM EXTREME SCORE:      8 Items
MINIMUM EXTREME SCORE:      4 Items
UMEAN=.000 USCALE=1.000
```

Table 4.3

Phase Three Draw-a-Man Sub-Test Summary Statistics

SUMMARY OF 56 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	22.5	63.0	-1.14	.38	1.00	.0	.92	.0
S.D.	10.3	.0	1.40	.05	.20	1.0	.44	.7
MAX.	50.0	63.0	2.44	.71	1.43	2.1	2.10	1.6
MIN.	3.0	63.0	-4.92	.35	.59	-2.5	.29	-1.1

REAL RMSE	.40	ADJ.SD	1.34	SEPARATION	3.36	Person RELIABILITY	.92
MODEL RMSE	.38	ADJ.SD	1.35	SEPARATION	3.51	Person RELIABILITY	.92
S.E. OF Person MEAN = .19							

DELETED: 51 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .00

SUMMARY OF 63 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	20.0	56.0	.00	.47	.98	.0	.92	.0
S.D.	15.9	.0	2.27	.21	.26	1.4	.79	1.3
MAX.	55.0	56.0	3.79	1.04	1.87	4.8	4.55	4.0
MIN.	1.0	56.0	-6.00	.31	.59	-3.2	.20	-2.0

REAL RMSE	.53	ADJ.SD	2.21	SEPARATION	4.14	Item RELIABILITY	.94
MODEL RMSE	.52	ADJ.SD	2.21	SEPARATION	4.27	Item RELIABILITY	.95
S.E. OF Item MEAN = .29							

MAXIMUM EXTREME SCORE: 4 Items
MINIMUM EXTREME SCORE: 6 Items
UMEAN=.000 USCALE=1.000

Item Statistics

Table 4.4 displays the item statistics for the phase one DAM sub-test data in measure order. The accepted values for the *t* statistic of the infit and outfit mean square residuals range from -2.0 to +2.0. Similarly, mean squares falling within the range of +.75 to +1.3 are also considered 'a good fit' by the Rasch model. These provide prima facie evidence of test unidimensionality of the data (Bond & Fox, 2007). Conversely, items with *t* statistics with a value larger than +2.0, and mean squares greater than +1.3 are indicative of an underfitting (i.e., erratic) performance. The following 10 items were reported as being too haphazard according to the Rasch model's expectations: item 6 (infit: Mn Sq. = 1.23, *t* = 2.0; outfit: Mn Sq. = 1.68, *t* = 2.60); item 9 (infit: Mn Sq. = 1.28, *t* = 2.5; outfit Mn Sq. = 1.35, *t* = 1.6); item 12 (infit: Mn Sq. = 1.02, *t* = .2; outfit Mn Sq. = 3.87, *t* = 2.2); item 14 (infit Mn Sq. = 1.17, *t* = 1.5; outfit Mn Sq. = 1.78, *t* = 2.7); item 21 (infit: Mn Sq. = .96, *t* = .2; outfit Mn Sq. = 7.38, *t* = 3.0); item 22 (infit: Mn Sq. = 1.42, *t* = 2.9; outfit: Mn Sq. = 1.59, *t* = 1.6); item 25 (infit: Mn Sq. = 1.37, *t* = 2.6; outfit Mn Sq. = 1.77, *t* = 2.0); item 48 (infit Mn Sq. = 1.52, *t* = 3.7; outfit Mn Sq. = 1.61, *t* = 1.8); item 49 (infit Mn Sq. = 1.48, *t* = 2.9; outfit Mn Sq. = 2.24, *t* = 2.4); and item 64 (infit Mn Sq. = 1.29, *t* = 2.5; outfit Mn Sq. = 1.40, *t* = 1.7).

Mean square values less than +.75, and *t* statistics less than -2.0 indicate that an item has been detected as 'overfitting'. Such items are considered by the Rasch model to be 'too predictable' with too little variation in their performance. The following four items from the DAM sub-test were reported as overfitting: item 33 (infit Mn Sq. = .64, *t* = -3.5;outfit Mn Sq. = .58,t *t* = -1.8); item 36 (infit Mn Sq. = .74, *t* = -2.5; outfit Mn Sq. = .65, *t* = -1.7); item 40 (infit Mn Sq. = .71, *t* = -2.6; outfit Mn Sq. = .56, *t* = -1.7); item 54 (infit Mn Sq. = .66, *t* = -3.8; outfit Mn Sq. = .53, *t* = -2.8).

The issue of fit is central in Rasch measurement. In contrast to many test theories – whereby the adherents discuss fit in regards to how well the model fits the data – Rasch measurement advocates insist that it is a researcher's task is to develop tests that produce data that fit the model's expectations. Bond and Fox state, "the Rasch model is a mathematical description of how fundamental measurement should operate with social/psychological variables. Its task is not to account for the data at hand, but rather to specify what kinds of data conform to the strict prescriptions of fundamental measurement" (2007, p. 235). The Rasch model represents the mathematical *ideal* of human attribute measurement; yet, obviously, human attributes are complex and difficult to measure, therefore it is unrealistic to expect 'ideal' or 'perfect' fitting measurement results. Therefore, 14 misfitting items out of a total of 73 represents about 20% of the items – a relatively small number given the complexity of the latent construct of intellectual development which is said to be under investigation by the test.

Table 4.4

Phase One Draw-a-Man Sub-Test Measure Order Item Statistics

Item	Raw		Difficulty	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
59	0	107	5.50	1.83	Maximum Estimated Measure			
60	0	107	5.50	1.83	Maximum Estimated Measure			
61	0	107	5.50	1.83	Maximum Estimated Measure			
69	0	107	5.50	1.83	Maximum Estimated Measure			
71	0	107	5.50	1.83	Maximum Estimated Measure			
73	0	107	5.50	1.83	Maximum Estimated Measure			
38	1	107	4.28	1.01	1.01	0.30	0.44	-0.30
21	2	107	3.56	0.73	0.96	0.20	7.38	3.00
37	2	107	3.56	0.73	0.97	0.20	0.29	-0.60
58	2	107	3.56	0.73	0.99	0.20	0.36	-0.40
70	2	107	3.56	0.73	0.89	0.10	0.21	-0.80
72	2	107	3.56	0.73	1.05	0.30	0.68	0.00
13	3	107	3.12	0.60	0.94	0.00	0.29	-0.70
66	3	107	3.12	0.60	1.08	0.30	3.01	1.70
12	4	107	2.81	0.53	1.02	0.20	3.87	2.20
34	5	107	2.55	0.48	1.15	0.50	9.90	6.10
62	5	107	2.55	0.48	0.93	-0.10	0.39	-0.70
26	6	107	2.34	0.44	0.82	-0.50	0.36	-0.90
57	6	107	2.34	0.44	0.87	-0.30	0.31	-1.00
68	6	107	2.34	0.44	0.85	-0.40	0.27	-1.10
16	7	107	2.16	0.41	1.21	0.70	1.29	0.60
15	10	107	1.72	0.36	1.22	1.00	1.54	1.00
20	10	107	1.72	0.36	0.92	-0.30	0.79	-0.20
52	10	107	1.72	0.36	1.16	0.70	0.99	0.20
27	11	107	1.60	0.35	0.97	-0.10	0.86	-0.10
32	11	107	1.60	0.35	0.81	-0.80	0.38	-1.30
23	13	107	1.37	0.32	1.06	0.40	1.00	0.20
17	15	107	1.17	0.31	1.16	0.90	1.16	0.50
65	15	107	1.17	0.31	0.87	-0.70	1.12	0.40
5	18	107	0.90	0.29	1.19	1.20	0.96	0.10
7	18	107	0.90	0.29	0.76	-1.60	0.44	-1.40
42	20	107	0.74	0.28	0.84	-1.00	0.59	-1.00
8	21	107	0.66	0.28	0.92	-0.50	0.68	-0.70
41	21	107	0.66	0.28	0.87	-0.80	0.56	-1.10
49	21	107	0.66	0.28	1.48	2.90	2.24	2.40
67	23	107	0.51	0.27	0.90	-0.70	0.68	-0.80
56	25	107	0.37	0.26	0.92	-0.50	0.82	-0.40
29	26	107	0.30	0.26	0.93	-0.50	0.70	-0.80
43	27	107	0.23	0.26	0.81	-1.50	0.64	-1.10
22	28	107	0.17	0.26	1.42	2.90	1.59	1.60
31	28	107	0.17	0.26	0.77	-1.90	0.61	-1.20
25	29	107	0.10	0.25	1.37	2.60	1.77	2.00
48	34	107	-0.21	0.24	1.52	3.70	1.61	1.80
40	35	107	-0.27	0.24	0.71	-2.60	0.56	-1.70

10	37	107	-0.38	0.24	1.04	0.40	0.98	0.00
33	38	107	-0.44	0.24	0.64	-3.50	0.58	-1.80
3	39	107	-0.50	0.24	0.82	-1.70	0.68	-1.30
19	43	107	-0.72	0.23	0.78	-2.10	0.67	-1.50
36	44	107	-0.78	0.23	0.74	-2.50	0.65	-1.70
6	45	107	-0.83	0.23	1.23	2.00	1.68	2.60
64	47	107	-0.94	0.23	1.29	2.50	1.40	1.70
28	53	107	-1.26	0.23	0.97	-0.20	0.89	-0.50
51	53	107	-1.26	0.23	1.08	0.80	1.01	0.10
54	60	107	-1.63	0.23	0.66	-3.80	0.53	-2.80
2	62	107	-1.74	0.23	1.11	1.10	1.20	1.00
9	63	107	-1.79	0.23	1.28	2.50	1.35	1.60
45	65	107	-1.90	0.23	0.84	-1.70	0.71	-1.50
63	66	107	-1.96	0.23	1.11	1.10	1.16	0.80
24	68	107	-2.07	0.24	1.02	0.20	1.05	0.30
14	72	107	-2.30	0.24	1.17	1.50	1.78	2.70
53	77	107	-2.60	0.25	0.98	-0.10	0.95	-0.10
39	78	107	-2.66	0.25	0.96	-0.30	0.77	-0.80
55	79	107	-2.73	0.26	0.84	-1.30	0.58	-1.50
18	81	107	-2.86	0.26	0.96	-0.20	0.76	-0.70
50	81	107	-2.86	0.26	1.25	1.80	1.73	2.00
47	85	107	-3.15	0.28	0.83	-1.10	0.56	-1.30
11	98	107	-4.49	0.39	0.84	-0.50	1.60	1.00
44	100	107	-4.84	0.44	0.73	-0.80	0.45	-0.60
30	101	107	-5.04	0.47	0.87	-0.30	1.90	1.10
46	102	107	-5.28	0.51	0.86	-0.30	0.23	-1.00
35	105	107	-6.40	0.76	1.08	0.30	0.52	-0.20
1	107	107	-8.46	1.85	Minimum Estimated Measure			
4	107	107	-8.46	1.85	Minimum Estimated Measure			
Mean	35.40	107.00	0.22	0.53	0.99	0.00	1.20	0.00
SD	33.20	0.00	3.11	0.49	0.19	1.40	1.49	1.50
r	0.97							

Table 4.5 presents the measure order item statistics for the phase two DAM sub-test data analysis. The following four items presented as unpredictable according to the Rasch model: item 25 (infit: Mn Sq. = 1.44, *t* = 3.1; outfit: Mn Sq. = 1.48, t = 1.5); item 48 (infit: Mn Sq. = 1.45, *t* = 3.2; outfit: Mn Sq. = 1.52, *t* = 1.8); item 49 (infit: Mn Sq. = 1.57, *t* = 2.7; outfit: Mn Sq. = 2.14, *t* = 2.0); and item 50 (infit: Mn Sq. = 1.55, *t* = 3.6; outfit: Mn Sq. = 3.43, *t* = 5.0). Item 49 (proportion: head II) was also reported as overly erratic in the phase one DAM sub-test data. Five items were detected as too predictable: item 33 (infit: Mn Sq. = .70, *t* = -2.8; outfit: Mn Sq. = .6, *t* = -2.1); item 36 (infit: Mn Sq.

= .72, *t* = -2.6; outfit: Mn Sq. = .61, *t* = -1.9); item 39 (infit: Mn Sq. = .73; *t* = -2.2; outfit: Mn Sq. = .51, -1.8); item 54 (infit: Mn Sq. = .56, *t* = -4.5; outfit: Mn Sq. = .45, *t* = -3.2); and item 67 (infit: Mn Sq. = .70, *t* = -2.7; outfit: Mn Sq. = .60, *t* -2.0). Items 33 (arms at side or engaged in activity), 36 (hip I: crotch) and 54 (Proportion: limbs in two dimensions) also presented as overly Guttman-like in the phase one DAM sub-test analysis.

Table 4.6 reveals that seven DAM items in phase three were detected as misfitting according to the Rasch model's expectations. Five of these items were reported as being overly haphazard: item 16 (infit: Mn Sq. = 1.53, *t* = 1.4; outfit: Mn Sq. = 4.55, *t* = 2.7); item 22 (infit: Mn Sq. = 1.43, ii 2.5; outfit: Mn Sq. = 1.53, *t* = 1.7); item 48 (infit: Mn Sq. = 1.48, *t* = 2.5; outfit: Mn Sq. = 1.79, *t* = 2.0); item 49 (infit: Mn Sq. = 1.79, *t* = 3.1; outfit: Mn Sq. = 2.76, *t* = 2.5); item 50 (infit: Mn Sq. = 1.87, *t* = 4.8; outfit: Mn Sq. = 3.20, *t* = 4.0). The two Guttman-like response patterns were reported for item 56 (infit: Mn Sq. = .63, *t* = -2.6; outfit: Mn Sq. = .49, *t* = -2.0) and item 67 (infit: Mn Sq. = .59, *t* = -3.2; outfit: Mn Sq. = .44, *t* = -1.9). Items 22, 48 and 49 were also detected as misfitting in the phase one DAM sub-test analysis, and items 48, 49, 50 and 67 in the phase two DAM data analysis.

Table 4.5

Phase Two Draw-a-Man Sub-Test Measure Order Item Statistics

Item	Raw		Difficulty	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
68	0	83	5.31	1.84	Maximum Estimated Measure			
70	0	83	5.31	1.84	Maximum Estimated Measure			
27	1	83	4.08	1.02	1.06	0.40	0.83	0.30
47	1	83	4.08	1.02	0.92	0.20	0.14	-0.70
49	1	83	4.08	1.02	0.92	0.20	0.14	-0.70
66	1	83	4.08	1.02	0.97	0.30	0.20	-0.60
69	1	83	4.08	1.02	0.97	0.30	0.20	-0.60
22	2	83	3.34	0.74	0.92	0.10	0.23	-0.60
67	2	83	3.34	0.74	0.99	0.20	0.33	-0.40
71	2	83	3.34	0.74	0.78	-0.20	0.15	-0.80
50	3	83	2.89	0.61	0.79	-0.30	0.20	-0.80
45	4	83	2.56	0.54	1.14	0.50	0.61	-0.10
54	4	83	2.56	0.54	1.13	0.40	1.18	0.50
65	5	83	2.29	0.49	1.02	0.20	1.20	0.50
39	6	83	2.07	0.46	0.82	-0.50	0.38	-0.80
61	7	83	1.87	0.43	1.04	0.20	0.66	-0.30
8	8	83	1.70	0.41	1.03	0.20	1.25	0.60
12	8	83	1.70	0.41	1.01	0.10	1.20	0.50
16	9	83	1.54	0.39	0.87	-0.50	0.41	-1.00
30	9	83	1.54	0.39	0.93	-0.20	0.54	-0.70
53	9	83	1.54	0.39	0.77	-0.90	0.36	-1.20
62	9	83	1.54	0.39	0.94	-0.20	0.45	-0.90
31	10	83	1.40	0.37	1.22	1.00	1.23	0.60
21	11	83	1.26	0.36	0.82	-0.80	0.42	-1.20
11	12	83	1.14	0.35	1.06	0.40	1.28	0.70
15	12	83	1.14	0.35	0.80	-0.90	0.49	-1.00
23	12	83	1.14	0.35	1.04	0.30	1.16	0.50
52	12	83	1.14	0.35	0.80	-0.90	0.64	-0.60
14	13	83	1.02	0.34	0.85	-0.70	0.62	-0.70
48	13	83	1.02	0.34	1.00	0.10	0.60	-0.80
18	14	83	0.91	0.33	1.01	0.10	1.20	0.60
7	16	83	0.70	0.32	0.88	-0.60	0.61	-0.90
5	17	83	0.60	0.31	0.90	-0.60	0.68	-0.70
29	22	83	0.16	0.29	1.33	2.10	1.38	1.10
38	22	83	0.16	0.29	1.07	0.50	0.87	-0.30
51	22	83	0.16	0.29	0.67	-2.60	0.51	-1.60
3	23	83	0.08	0.28	0.65	-2.80	0.58	-1.30
25	23	83	0.08	0.28	0.66	-2.70	0.48	-1.80
34	25	83	-0.08	0.28	1.97	5.60	3.19	4.30
57	25	83	-0.08	0.28	1.46	3.00	1.84	2.10
60	26	83	-0.16	0.28	1.05	0.40	4.37	6.00
37	28	83	-0.31	0.27	0.85	-1.10	0.69	-1.10
42	29	83	-0.38	0.27	0.71	-2.50	0.70	-1.00
10	31	83	-0.53	0.27	1.06	0.50	0.99	0.10

40	31	83	-0.53	0.27	0.83	-1.40	0.68	-1.30
63	34	83	-0.74	0.26	0.92	-0.60	0.84	-0.60
26	35	83	-0.81	0.26	0.68	-2.90	0.60	-1.90
36	35	83	-0.81	0.26	0.93	-0.60	0.81	-0.70
46	37	83	-0.95	0.26	0.74	-2.40	0.64	-1.70
6	43	83	-1.35	0.26	1.14	1.20	1.11	0.60
32	43	83	-1.35	0.26	1.22	1.80	1.49	2.10
44	46	83	-1.55	0.26	0.94	-0.50	1.08	0.40
64	48	83	-1.69	0.26	1.18	1.50	1.40	1.70
17	49	83	-1.76	0.26	1.30	2.40	1.46	1.80
58	50	83	-1.83	0.26	1.52	3.90	2.36	4.30
2	53	83	-2.04	0.27	1.20	1.60	1.08	0.40
20	55	83	-2.18	0.27	0.88	-1.00	0.77	-0.80
9	58	83	-2.40	0.28	1.16	1.20	1.38	1.20
28	58	83	-2.40	0.28	1.31	2.30	2.07	2.90
35	65	83	-2.99	0.30	0.85	-1.00	0.58	-1.10
43	65	83	-2.99	0.30	0.79	-1.40	0.56	-1.20
56	69	83	-3.38	0.33	0.97	-0.10	1.89	1.60
59	75	83	-4.19	0.41	0.90	-0.20	0.59	-0.40
41	78	83	-4.80	0.51	0.85	-0.30	0.54	-0.30
19	80	83	-5.45	0.64	0.86	-0.10	0.40	-0.40
13	81	83	-5.93	0.77	0.68	-0.40	0.14	-0.90
33	81	83	-5.93	0.77	0.58	-0.60	0.08	-1.10
4	82	83	-6.73	1.05	1.12	0.40	0.25	-0.50
1	83	83	-8.02	1.85	Minimum Estimated Measure			
24	83	83	-8.02	1.85	Minimum Estimated Measure			
55	83	83	-8.02	1.85	Minimum Estimated Measure			
Mean	29.70	83.00	-0.19	0.53	0.98	0.00	0.88	0.00
SD	26.60	0.00	3.10	0.43	0.23	1.50	0.72	1.50
r	0.96							

Table 4.6

Phase Three Draw-a-Man Sub-Test Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
21	0	56	5.05	1.84	Maximum Estimated Measure			
59	0	56	5.05	1.84	Maximum Estimated Measure			
60	0	56	5.05	1.84	Maximum Estimated Measure			
61	0	56	5.05	1.84	Maximum Estimated Measure			
38	1	56	3.79	1.03	1.02	0.30	0.29	-0.10
58	1	56	3.79	1.03	0.96	0.30	0.20	-0.30
66	1	56	3.79	1.03	0.96	0.30	0.20	-0.30
73	1	56	3.79	1.03	1.04	0.40	0.36	0.00
12	2	56	3.02	0.75	0.75	-0.20	0.20	-0.50
13	2	56	3.02	0.75	0.75	-0.20	0.20	-0.50
37	2	56	3.02	0.75	1.09	0.30	0.44	-0.10
62	2	56	3.02	0.75	0.75	-0.20	0.20	-0.50
69	2	56	3.02	0.75	0.75	-0.20	0.20	-0.50
71	2	56	3.02	0.75	0.81	-0.10	0.32	-0.20
34	3	56	2.54	0.63	1.05	0.30	0.43	-0.30
57	4	56	2.19	0.56	0.63	-0.90	0.20	-1.00
15	5	56	1.90	0.51	1.10	0.40	2.44	1.50
16	5	56	1.90	0.51	1.53	1.40	4.55	2.70
72	5	56	1.90	0.51	1.08	0.30	0.55	-0.40
41	6	56	1.65	0.48	0.66	-1.20	0.29	-1.10
68	6	56	1.65	0.48	0.75	-0.80	0.34	-1.00
70	6	56	1.65	0.48	0.82	-0.50	0.37	-0.90
52	7	56	1.44	0.45	0.92	-0.20	0.76	-0.20
26	8	56	1.24	0.43	0.79	-0.80	0.53	-0.70
23	9	56	1.07	0.41	1.02	0.20	1.41	0.80
65	9	56	1.07	0.41	0.92	-0.20	1.31	0.70
20	10	56	0.90	0.40	1.19	0.90	0.77	-0.30
27	10	56	0.90	0.40	1.00	0.10	0.67	-0.50
42	10	56	0.90	0.40	0.69	-1.40	0.46	-1.00
49	11	56	0.75	0.39	1.79	3.10	2.76	2.50
3	12	56	0.60	0.38	1.04	0.20	0.74	-0.40
17	12	56	0.60	0.38	0.90	-0.40	0.85	-0.20
32	12	56	0.60	0.38	0.71	-1.50	0.52	-1.10
7	14	56	0.34	0.36	0.85	-0.80	0.60	-1.00
5	16	56	0.09	0.35	1.07	0.40	1.22	0.70
43	17	56	-0.03	0.34	0.85	-0.90	0.68	-0.90
48	17	56	-0.03	0.34	1.48	2.50	1.79	2.00
8	18	56	-0.14	0.34	0.82	-1.10	0.78	-0.60
25	18	56	-0.14	0.34	1.25	1.50	1.54	1.50
56	20	56	-0.36	0.33	0.63	-2.60	0.49	-2.00
22	21	56	-0.47	0.33	1.43	2.50	1.53	1.70
45	21	56	-0.47	0.33	0.81	-1.30	0.68	-1.20
31	23	56	-0.68	0.32	0.80	-1.40	0.68	-1.30
29	26	56	-0.98	0.32	1.07	0.50	0.98	0.00

40	26	56	-0.98	0.32	0.80	-1.40	0.69	-1.30
10	28	56	-1.18	0.31	0.71	-2.20	0.61	-1.80
28	28	56	-1.18	0.31	0.99	0.00	0.96	-0.10
33	28	56	-1.18	0.31	0.88	-0.90	0.85	-0.60
51	31	56	-1.47	0.31	1.38	2.50	1.40	1.50
2	32	56	-1.57	0.31	1.10	0.70	2.23	3.60
6	33	56	-1.67	0.32	0.93	-0.50	1.15	0.60
36	33	56	-1.67	0.32	0.83	-1.30	0.75	-0.90
64	33	56	-1.67	0.32	1.10	0.80	0.99	0.00
54	34	56	-1.77	0.32	0.77	-1.80	0.63	-1.40
63	38	56	-2.18	0.33	1.39	2.50	1.64	1.70
9	39	56	-2.29	0.33	0.86	-0.90	0.71	-0.80
50	39	56	-2.29	0.33	1.87	4.80	3.20	4.00
67	39	56	-2.29	0.33	0.59	-3.20	0.44	-1.90
19	41	56	-2.51	0.34	1.11	0.70	1.02	0.20
53	41	56	-2.51	0.34	1.01	0.10	1.01	0.20
24	43	56	-2.76	0.35	1.31	1.70	1.97	1.80
55	43	56	-2.76	0.35	0.68	-2.00	0.45	-1.30
14	48	56	-3.49	0.42	1.03	0.20	1.31	0.60
39	49	56	-3.67	0.44	0.92	-0.20	0.57	-0.50
47	51	56	-4.11	0.51	0.75	-0.60	0.33	-0.80
18	53	56	-4.75	0.63	0.81	-0.20	0.48	-0.20
11	55	56	-6.00	1.04	1.15	0.50	1.24	0.60
1	56	56	-7.27	1.85	Minimum Estimated Measure			
4	56	56	-7.27	1.85	Minimum Estimated Measure			
30	56	56	-7.27	1.85	Minimum Estimated Measure			
35	56	56	-7.27	1.85	Minimum Estimated Measure			
44	56	56	-7.27	1.85	Minimum Estimated Measure			
46	56	56	-7.27	1.85	Minimum Estimated Measure			
Mean	21.90	56.00	-0.32	0.66	0.98	0.00	0.92	0.00
SD	18.50	0.00	3.18	0.51	0.26	1.40	0.79	1.30
r	0.94							

Case Statistics

Table 4.7 presents the person statistics in measure order for the phase one DAM sub-test data collection. Out of a total of 107 participants, five performances were found to be misfitting. Only one case was reported to be sitting on the borderline of a response string closer to the Guttman-style (i.e., where the case was successful on easy items and then unsuccessful on all difficult items e.g., the child's response string might resemble the following: 1111010000). This was person ID number 113F (infit Mn Sq. = .67, *t* = -2.1; outfit Mn Sq. = .42, *t* = -1.0). Conversely, four children presented with response strings detected as overly erratic by the Rasch model (i.e., where the child unexpectedly received credit for more difficult items, yet was unsuccessful on less difficult items e.g., the child's response string might resemble the following: 0000101111). These children were: case 101F (infit Mn Sq. = .95, *t* = -.2; outfit fit Mn Sq. = 2.62, *t* = 2.3); case 108M (infit Mn Sq. = 1.35, *t* = 1.9; outfit Mn Sq. = 2.07, *t* = 1.6); case 150M2 (infit Mn Sq. = 1.42, *t* = 2.2; outfit Mn Sq. = 1.32, *t* = 0.7); case 180M (infit Mn Sq. = 1.46, *t* = 2.2; outfit Mn Sq. = 1.46, *t* = .9); case 191F (infit Mn Sq. = 2.33, *t* = 2.5; outfit Mn Sq. = 9.90, *t* = 5.6); and case 193F (infit Mn Sq. = 1.45, *t* = 2.4; outfit Mn Sq. = 1.57, *t* = 1.1).

Table 4.7

Phase One Draw-a-Man Sub-Test Measure Order Person Statistics

Case	Raw		Ability	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
177 F	46	65	1.76	0.36	1.00	0.00	0.71	-0.30
174 F	44	65	1.51	0.35	0.96	-0.10	1.23	0.60
100 F	43	65	1.38	0.35	0.94	-0.30	0.65	-0.50
150 M	41	65	1.14	0.35	1.42	2.20	1.32	0.70
196 F	40	65	1.02	0.34	1.14	0.80	1.12	0.40
107 F	39	65	0.91	0.34	0.80	-1.20	0.51	-0.80
108 M	39	65	0.91	0.34	1.35	1.90	2.07	1.60
109 M	39	65	0.91	0.34	0.96	-0.10	0.79	-0.20
102 F	38	65	0.79	0.34	1.20	1.10	1.16	0.50
113 F	38	65	0.79	0.34	0.67	-2.10	0.42	-1.00
122 F	38	65	0.79	0.34	1.08	0.50	0.78	-0.20
166 F	37	65	0.67	0.34	0.85	-0.90	0.63	-0.50
201 F	37	65	0.67	0.34	0.81	-1.10	0.55	-0.70
204 F	37	65	0.67	0.34	0.92	-0.40	0.84	-0.10
112 F	35	65	0.44	0.34	0.71	-1.80	0.53	-0.80
116 F	35	65	0.44	0.34	0.83	-1.00	0.93	0.10
155 F	35	65	0.44	0.34	1.26	1.40	3.80	3.10
125 F	34	65	0.32	0.34	0.98	-0.10	1.06	0.30
193 F	34	65	0.32	0.34	1.45	2.40	1.57	1.10
206 M	34	65	0.32	0.34	1.05	0.30	0.73	-0.30
195 F	33	65	0.21	0.34	0.92	-0.40	0.71	-0.40
120 F	31	65	-0.03	0.34	0.90	-0.50	0.68	-0.50
124 F	31	65	-0.03	0.34	1.04	0.30	0.97	0.10
146 F	31	65	-0.03	0.34	1.05	0.40	1.08	0.30
101 F	30	65	-0.14	0.34	0.95	-0.20	2.62	2.30
156 F	30	65	-0.14	0.34	1.04	0.30	0.79	-0.20
175 F	30	65	-0.14	0.34	0.84	-0.90	0.60	-0.70
192 M	30	65	-0.14	0.34	0.79	-1.20	0.53	-0.90
121 F	29	65	-0.26	0.35	1.10	0.60	1.01	0.20
167 F	29	65	-0.26	0.35	1.32	1.70	2.19	1.80
153 F	28	65	-0.38	0.35	0.91	-0.40	0.70	-0.40
172 M	28	65	-0.38	0.35	0.82	-1.00	0.51	-0.90
179 F	28	65	-0.38	0.35	0.84	-0.90	0.58	-0.70
205 F	28	65	-0.38	0.35	0.97	-0.10	0.65	-0.60
119 M	27	65	-0.51	0.35	0.85	-0.80	0.54	-0.80
154 F	27	65	-0.51	0.35	0.92	-0.40	0.75	-0.30
180 M	27	65	-0.51	0.35	1.46	2.20	1.46	0.90
202 M	27	65	-0.51	0.35	0.77	-1.30	0.54	-0.80
105 F	26	65	-0.63	0.35	0.93	-0.30	1.02	0.20
126 F	25	65	-0.75	0.35	0.87	-0.60	0.65	-0.50
145 F	25	65	-0.75	0.35	1.21	1.10	0.92	0.10
151 F	25	65	-0.75	0.35	0.77	-1.20	0.68	-0.40
181 M	25	65	-0.75	0.35	0.86	-0.70	0.53	-0.80
173 M	24	65	-0.88	0.36	1.03	0.20	0.75	-0.30

111 M	23	65	-1.01	0.36	1.09	0.50	0.84	-0.10
147 F	22	65	-1.14	0.37	0.90	-0.50	0.67	-0.40
178 M	22	65	-1.14	0.37	1.21	1.00	0.83	-0.10
199 F	22	65	-1.14	0.37	1.08	0.50	0.77	-0.20
127 F	21	65	-1.28	0.37	0.99	0.00	0.86	0.00
141 F	21	65	-1.28	0.37	1.01	0.10	0.73	-0.20
157 F	21	65	-1.28	0.37	0.91	-0.40	0.80	-0.10
128 M	20	65	-1.41	0.37	1.20	1.00	0.98	0.20
135 F	20	65	-1.41	0.37	1.34	1.60	2.54	1.80
104 F	19	65	-1.55	0.38	1.13	0.70	0.66	-0.30
106 M	19	65	-1.55	0.38	0.81	-0.90	0.46	-0.70
110 M	19	65	-1.55	0.38	0.74	-1.30	0.43	-0.80
117 M	19	65	-1.55	0.38	0.84	-0.80	0.65	-0.30
138 F	19	65	-1.55	0.38	1.15	0.70	1.52	0.90
182 M	19	65	-1.55	0.38	0.98	0.00	0.77	-0.10
132 M	18	65	-1.70	0.38	0.77	-1.10	0.55	-0.50
165 F	18	65	-1.70	0.38	0.79	-1.00	0.45	-0.70
115 M	17	65	-1.85	0.39	0.93	-0.30	0.94	0.20
123 F	17	65	-1.85	0.39	1.06	0.40	1.09	0.40
170 M	17	65	-1.85	0.39	1.04	0.20	0.83	0.00
203 M	17	65	-1.85	0.39	0.75	-1.20	0.43	-0.70
114 M	16	65	-2.00	0.40	0.96	-0.10	0.69	-0.10
118 M	16	65	-2.00	0.40	1.10	0.50	0.61	-0.30
129 M	16	65	-2.00	0.40	1.11	0.50	1.47	0.80
133 F	16	65	-2.00	0.40	1.11	0.50	1.12	0.40
134 M	16	65	-2.00	0.40	1.26	1.20	1.26	0.60
136 F	16	65	-2.00	0.40	0.96	-0.10	1.18	0.50
159 F	16	65	-2.00	0.40	1.25	1.10	1.35	0.70
161 F	16	65	-2.00	0.40	0.89	-0.50	0.61	-0.30
163 M	16	65	-2.00	0.40	0.95	-0.20	0.51	-0.40
183 F	16	65	-2.00	0.40	1.12	0.60	0.84	0.10
184 M	16	65	-2.00	0.40	0.75	-1.20	0.43	-0.60
194 F	16	65	-2.00	0.40	0.65	-1.70	0.31	-0.90
197 M	16	65	-2.00	0.40	1.10	0.50	1.07	0.40
158 F	15	65	-2.17	0.41	0.94	-0.20	0.66	-0.10
171 F	15	65	-2.17	0.41	0.95	-0.10	0.70	-0.10
198 M	15	65	-2.17	0.41	0.87	-0.50	0.48	-0.40
103 M	14	65	-2.33	0.41	0.86	-0.50	0.50	-0.40
139 F	14	65	-2.33	0.41	1.06	0.30	0.61	-0.20
140 M	14	65	-2.33	0.41	0.95	-0.10	0.87	0.20
142 M	14	65	-2.33	0.41	0.83	-0.70	0.51	-0.40
148 M	14	65	-2.33	0.41	0.87	-0.50	0.40	-0.60
152 F	14	65	-2.33	0.41	0.62	-1.90	0.39	-0.60
137 M	13	65	-2.51	0.43	1.21	0.90	1.03	0.40
164 F	13	65	-2.51	0.43	1.38	1.50	6.19	3.00
185 F	13	65	-2.51	0.43	1.01	0.10	0.88	0.20
176 M	12	65	-2.69	0.44	1.01	0.10	0.80	0.10
186 M	12	65	-2.69	0.44	1.34	1.30	0.94	0.30
149 M	11	65	-2.89	0.45	0.73	-1.00	0.39	-0.50
169 F	11	65	-2.89	0.45	1.10	0.40	1.09	0.40

190 F	11	65	-2.89	0.45	0.85	-0.50	0.37	-0.50
162 F	10	65	-3.10	0.47	1.43	1.50	9.90	5.20
168 F	10	65	-3.10	0.47	0.77	-0.80	0.32	-0.60
187 F	10	65	-3.10	0.47	1.32	1.10	0.65	-0.10
143 M	9	65	-3.33	0.49	0.55	-1.70	0.18	-0.90
144 M	9	65	-3.33	0.49	0.80	-0.60	0.31	-0.60
131 M	8	65	-3.58	0.51	0.96	0.00	0.36	-0.50
188 M	8	65	-3.58	0.51	0.60	-1.30	0.21	-0.80
200 M	8	65	-3.58	0.51	1.51	1.50	1.23	0.60
189 M	5	65	-4.53	0.61	1.17	0.50	0.49	-0.20
160 F	4	65	-4.93	0.66	1.63	1.40	1.29	0.60
191 F	4	65	-4.93	0.66	2.33	2.50	9.90	5.60
130 M	2	65	-6.04	0.85	0.57	-0.80	0.06	-1.20
Mean	22.1	65	-1.30	0.39	1.01	0.00	1.06	0.10
SD	10.2	0	1.50	0.08	0.25	1.00	1.42	1.10
r	0.97							

Table 4.8 presents the item statistics in measure order for the phase two DAM data analysis. The phase two data sample comprised 83 children, and out of this total, only two cases presented as marginally misfitting. Both cases were detected as showing more variation than expected by the Rasch model. Case ID 112F (infit: Mn Sq. = 1.63. t = 2.7; outfit: Mn Sq. = 1.59, t = 1.0) and case ID 193F (infit: Mn Sq. = 1.46, t = 2.3; outfit: Mn Sq. = 1.82, t = 1.4) both reported results which indicated that they were only slightly unpredictable. Out of the two misfitting cases described above, case ID 193F's drawing performance was reported as unpredictable in the phase one DAM (infit: Mn Sq. = 1.45, t = 2.4; outfit: Mn Sq. = 1.57, t = 1.1) sub-test analysis as well.

Table 4.8

Phase Two Draw-a-Man Sub-Test Measure Order Person Statistics

Case #	Raw Score	Count	Ability Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
112 F	43	61	1.72	0.38	1.63	2.70	1.59	1.00
174 F	43	61	1.72	0.38	0.95	-0.20	1.73	1.20
125 F	39	61	1.17	0.36	1.19	1.00	0.88	-0.10
177 F	39	61	1.17	0.36	0.93	-0.30	0.62	-0.70
204 F	38	61	1.04	0.36	1.37	1.70	1.76	1.40
196 F	37	61	0.91	0.36	0.83	-0.90	0.57	-0.90
205 F	36	61	0.78	0.36	1.34	1.70	1.97	1.70
101 F	35	61	0.65	0.36	1.18	1.00	1.38	0.90
156 F	34	61	0.52	0.36	1.12	0.70	0.90	0.00
166 F	33	61	0.40	0.35	0.95	-0.20	0.68	-0.50
167 F	33	61	0.40	0.35	1.13	0.80	1.19	0.50
123 F	32	61	0.27	0.35	0.97	-0.10	1.02	0.20
121 F	31	61	0.15	0.35	0.89	-0.50	0.75	-0.30
150 M	31	61	0.15	0.35	1.24	1.30	1.51	1.00
172 M	31	61	0.15	0.35	0.82	-1.00	0.74	-0.40
113 F	30	61	0.03	0.35	1.06	0.40	0.95	0.10
116 F	30	61	0.03	0.35	1.25	1.40	1.30	0.70
155 F	30	61	0.03	0.35	1.03	0.20	0.79	-0.20
192 M	30	61	0.03	0.35	0.83	-0.90	0.64	-0.60
193 F	30	61	0.03	0.35	1.46	2.30	1.82	1.40
120 F	29	61	-0.10	0.35	1.07	0.40	1.00	0.20
122 F	29	61	-0.10	0.35	0.74	-1.60	0.48	-1.00
126 F	29	61	-0.10	0.35	1.15	0.90	0.88	-0.10
124 F	28	61	-0.22	0.35	1.10	0.60	0.87	-0.10
175 F	28	61	-0.22	0.35	0.97	-0.10	0.76	-0.30
146 F	27	61	-0.35	0.35	0.81	-1.10	0.63	-0.60
180 M	27	61	-0.35	0.35	0.76	-1.40	0.48	-1.00
151 F	26	61	-0.47	0.35	1.06	0.40	1.50	0.90
179 F	26	61	-0.47	0.35	0.92	-0.40	0.64	-0.50
170 M	25	61	-0.60	0.35	1.00	0.10	0.74	-0.20
173 M	25	61	-0.60	0.35	0.96	-0.20	0.70	-0.30
178 M	25	61	-0.60	0.35	0.88	-0.60	0.76	-0.20
202 M	25	61	-0.60	0.35	0.93	-0.30	0.94	0.10
181 M	24	61	-0.72	0.36	0.97	-0.10	0.86	0.00
199 F	24	61	-0.72	0.36	0.98	-0.10	0.80	-0.10
114 M	22	61	-0.98	0.36	0.94	-0.30	0.83	0.00
147 F	22	61	-0.98	0.36	1.12	0.70	1.01	0.30
135 F	21	61	-1.11	0.36	0.96	-0.10	0.68	-0.30
118 M	20	61	-1.24	0.37	0.95	-0.20	0.60	-0.40
148 M	20	61	-1.24	0.37	1.09	0.50	0.79	-0.10
194 F	20	61	-1.24	0.37	0.85	-0.80	0.57	-0.40
110 M	19	61	-1.38	0.37	1.23	1.20	1.11	0.40
117 M	19	61	-1.38	0.37	1.22	1.10	1.28	0.60
152 F	19	61	-1.38	0.37	1.05	0.30	1.54	0.90

154 F	19	61	-1.38	0.37	0.65	-2.00	0.49	-0.60
115 M	18	61	-1.52	0.38	1.24	1.20	1.15	0.50
158 F	18	61	-1.52	0.38	0.93	-0.30	0.66	-0.20
182 M	18	61	-1.52	0.38	0.96	-0.10	0.62	-0.30
103 M	17	61	-1.66	0.38	0.85	-0.70	0.53	-0.40
136 F	17	61	-1.66	0.38	0.81	-1.00	0.55	-0.40
165 F	17	61	-1.66	0.38	0.67	-1.80	0.40	-0.60
171 F	17	61	-1.66	0.38	0.93	-0.30	0.55	-0.40
137 M	16	61	-1.81	0.39	1.12	0.60	0.81	0.10
153 F	16	61	-1.81	0.39	1.12	0.70	2.55	1.60
127 F	15	61	-1.97	0.40	1.18	0.90	3.18	1.90
132 M	15	61	-1.97	0.40	0.77	-1.10	0.44	-0.50
141 F	15	61	-1.97	0.40	0.78	-1.00	0.67	-0.10
163 M	15	61	-1.97	0.40	1.18	0.90	0.79	0.00
133 F	14	61	-2.13	0.41	0.96	-0.10	0.57	-0.30
157 F	14	61	-2.13	0.41	0.66	-1.70	0.33	-0.80
186 M	14	61	-2.13	0.41	1.05	0.30	0.86	0.10
187 F	13	61	-2.30	0.42	0.80	-0.90	0.44	-0.50
129 M	12	61	-2.47	0.43	0.88	-0.40	0.96	0.30
139 F	12	61	-2.47	0.43	1.06	0.30	0.80	0.10
149 M	12	61	-2.47	0.43	0.89	-0.40	0.95	0.30
160 F	12	61	-2.47	0.43	0.80	-0.80	0.44	-0.40
161 F	12	61	-2.47	0.43	0.77	-1.00	0.45	-0.40
169 F	12	61	-2.47	0.43	0.69	-1.40	0.34	-0.60
176 M	12	61	-2.47	0.43	1.07	0.40	0.74	0.00
128 M	11	61	-2.66	0.44	0.73	-1.10	0.44	-0.40
168 F	11	61	-2.66	0.44	0.95	-0.10	1.01	0.30
185 F	11	61	-2.66	0.44	0.85	-0.50	0.37	-0.50
188 M	11	61	-2.66	0.44	1.44	1.60	4.81	2.40
189 M	11	61	-2.66	0.44	1.02	0.20	0.92	0.20
190 F	11	61	-2.66	0.44	1.15	0.60	0.82	0.10
142 M	10	61	-2.87	0.46	0.89	-0.30	0.37	-0.50
159 F	10	61	-2.87	0.46	0.93	-0.20	0.59	-0.10
191 F	10	61	-2.87	0.46	0.81	-0.70	0.37	-0.50
184 M	9	61	-3.09	0.48	1.18	0.70	0.51	-0.20
119 M	8	61	-3.33	0.51	0.81	-0.50	0.60	-0.10
143 M	7	61	-3.61	0.54	0.88	-0.20	0.47	-0.30
131 M	4	61	-4.77	0.73	1.11	0.40	0.33	-0.40
130 M	3	61	-5.37	0.82	1.81	1.30	0.64	0.00
Mean	29	61	-0.10	0.35	1.15	0.90	0.88	-0.10
SD	28	61	-0.22	0.35	1.10	0.60	0.87	-0.10
r	0.96							

Table 4.9 displays the measure order person statistics for the phase three DAM sub-test data analysis. This analysis detected only two marginally misfitting performances. The slightly erratic performance was reported as those by: case ID 178M (infit: Mn Sq. = 1.43, *t* = 2.1; outfit: Mn Sq. = 1.77, *t* = 1.5). The second of these case performances was interesting in that only the outfit mean square was unexpectedly high; the other fit statistics are relatively acceptable; case ID 147F (infit: Mn Sq. = .59, *t* = -2.5; outfit: Mn Sq. = .32, *t* = -1.0) was reported as having a Guttman-style response string. This is the first time that either of these children's drawings yielded fit statistics detected as misfitting by the Rasch model.

Table 4.9

Phase Three Draw-a-Man Sub-Test Measure Order Person Statistics

Case #	Raw Score	Count	Ability Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
100 F	50	63	2.44	0.39	1.29	1.30	1.64	1.00
177 F	43	63	1.46	0.36	1.28	1.50	1.27	0.60
150 M	42	63	1.33	0.36	0.88	-0.60	0.56	-0.70
174 F	40	63	1.08	0.35	1.08	0.50	1.39	0.80
112 F	39	63	0.95	0.35	0.85	-0.80	0.92	0.00
179 F	39	63	0.95	0.35	1.35	1.80	1.26	0.60
193 F	38	63	0.83	0.35	1.32	1.70	1.92	1.60
192 M	35	63	0.47	0.35	1.00	0.10	0.86	-0.10
167 F	34	63	0.35	0.35	1.27	1.40	1.60	1.20
196 F	34	63	0.35	0.35	1.07	0.40	0.98	0.10
178 M	32	63	0.11	0.35	1.43	2.10	1.77	1.50
172 M	31	63	-0.01	0.35	0.78	-1.20	0.62	-0.70
121 F	29	63	-0.25	0.35	0.78	-1.20	0.57	-0.80
125 F	29	63	-0.25	0.35	0.97	-0.10	0.73	-0.40
166 F	29	63	-0.25	0.35	1.00	0.00	0.80	-0.30
175 F	29	63	-0.25	0.35	0.80	-1.10	0.58	-0.80
195 F	28	63	-0.37	0.35	1.13	0.80	0.90	0.00
122 F	27	63	-0.49	0.35	0.74	-1.50	0.49	-1.10
123 F	27	63	-0.49	0.35	0.98	0.00	0.89	-0.10
202 M	27	63	-0.49	0.35	1.24	1.30	1.02	0.20
101 F	26	63	-0.61	0.35	0.84	-0.90	0.64	-0.60
116 F	26	63	-0.61	0.35	1.05	0.30	0.77	-0.30
180 M	26	63	-0.61	0.35	0.78	-1.20	0.54	-0.90
139 F	23	63	-0.98	0.36	0.79	-1.20	0.56	-0.80
199 F	23	63	-0.98	0.36	0.98	-0.10	0.66	-0.50
110 M	22	63	-1.11	0.36	1.10	0.60	0.88	0.00
118 M	22	63	-1.11	0.36	1.03	0.20	0.80	-0.20
135 F	21	63	-1.24	0.36	0.91	-0.40	0.67	-0.40
164 F	21	63	-1.24	0.36	1.16	0.90	1.14	0.40
141 F	20	63	-1.37	0.36	0.76	-1.30	0.50	-0.70
194 F	19	63	-1.50	0.37	1.30	1.50	1.20	0.50
120 F	17	63	-1.77	0.38	0.91	-0.40	0.55	-0.40
140 M	17	63	-1.77	0.38	0.85	-0.70	0.50	-0.50
147 F	17	63	-1.77	0.38	0.59	-2.50	0.32	-1.00
173 M	17	63	-1.77	0.38	0.87	-0.60	1.61	0.90
115 M	16	63	-1.92	0.38	1.18	0.90	0.76	0.00
152 F	16	63	-1.92	0.38	0.80	-1.10	1.06	0.40
159 F	16	63	-1.92	0.38	0.70	-1.70	0.41	-0.70
171 F	16	63	-1.92	0.38	1.04	0.30	0.79	0.00
182 M	16	63	-1.92	0.38	0.96	-0.10	1.20	0.50
184 M	16	63	-1.92	0.38	1.08	0.50	0.84	0.10
176 M	15	63	-2.06	0.39	0.98	0.00	0.62	-0.20
185 F	15	63	-2.06	0.39	0.75	-1.30	0.40	-0.60
129 M	13	63	-2.37	0.40	0.95	-0.20	1.60	0.90

163 M	13	63	-2.37	0.40	0.97	-0.10	1.77	1.00
165 F	13	63	-2.37	0.40	0.74	-1.30	0.36	-0.70
168 F	12	63	-2.54	0.41	0.93	-0.30	0.58	-0.20
143 M	11	63	-2.71	0.42	1.07	0.40	0.91	0.20
157 F	11	63	-2.71	0.42	1.30	1.20	1.41	0.70
188 M	11	63	-2.71	0.42	1.30	1.30	1.35	0.70
189 M	11	63	-2.71	0.42	0.93	-0.20	2.10	1.20
131 M	10	63	-2.90	0.44	1.01	0.10	0.51	-0.30
160 F	10	63	-2.90	0.44	0.71	-1.20	0.29	-0.70
191 F	10	63	-2.90	0.44	1.18	0.80	1.39	0.70
169 F	9	63	-3.10	0.45	0.94	-0.20	0.61	-0.10
186 M	3	63	-4.92	0.71	1.15	0.50	0.68	0.10
Mean	22.5	63	-1.14	0.38	1.00	0.00	0.92	0.00
SD	10.3	0	1.40	0.05	0.20	1.00	0.44	0.70
r	.97							

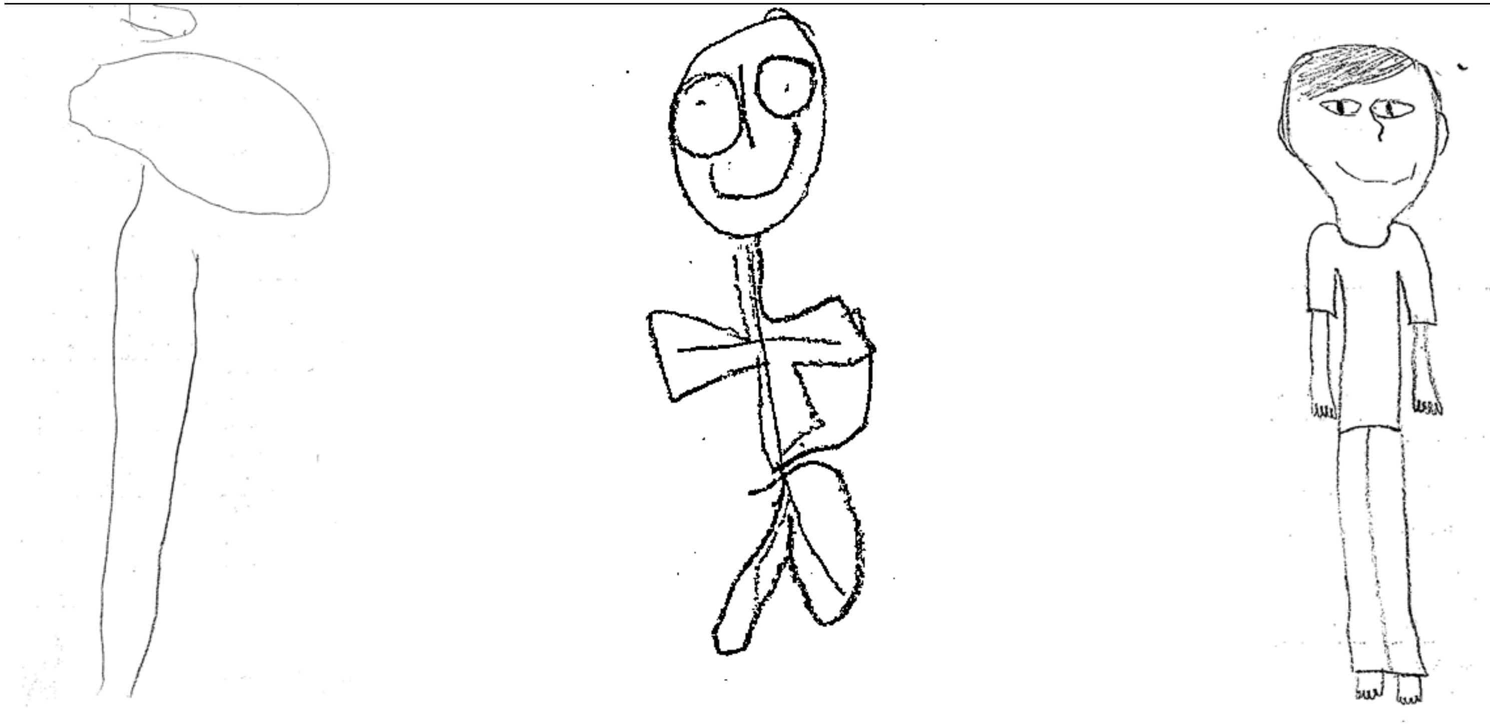
Examples of Children's Draw-a-Man Drawings

Figures 4.4, 4.5 and 4.6 present examples of the children's DAM drawings from phase one, two and three respectively. In each figure, an example of the least successful, mean and most successful drawing performances, according the Rasch analysis results, are displayed.

Review of Figures 4.4 to 4.6 reveals remarkable development of the human figure drawings – regardless of whether it is amongst the least successful, mean or most successful drawing performances. In particular, there is evidence that the human figure drawings develop over the three phases of data collection. That is, a larger number of human attributes have been conceptualized as the children mature intellectually and, thus, more detail comes through to the child's drawn representations of a human figure.

Figure 4.4

Phase One Draw-a-Man Drawings: Examples of Least Successful, Mean, and Most Successful



Case ID: 130M Ability Estimate = -6.04; SE = .85	Case ID: 135F Ability Estimate = -1.41; SE = .37	Case ID: 177F Ability Estimate = 1.76; SE = .36

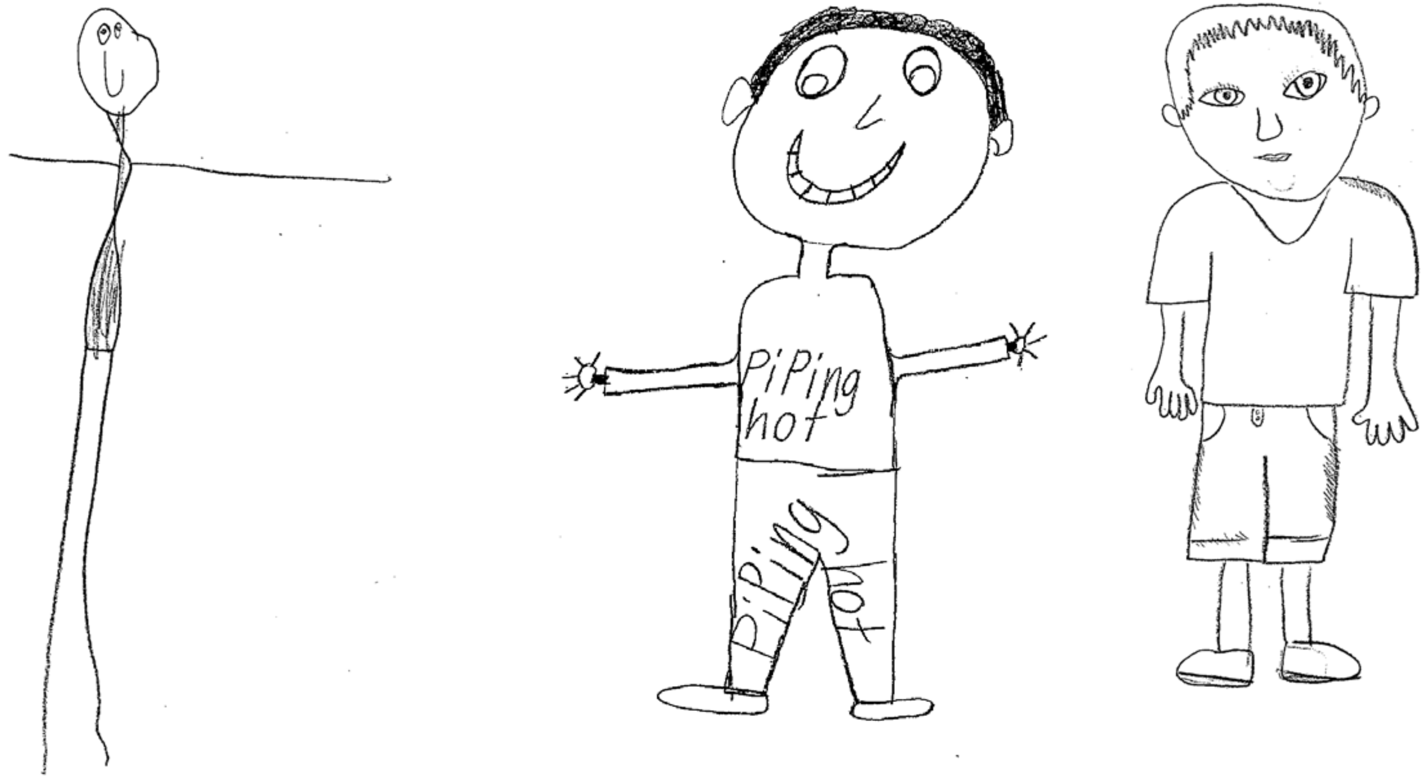
Figure 4.5

Phase Two Draw-a-Man Drawings: Examples of Least Successful, Mean, and Most Successful

Case ID: 130M Ability Estimate = -5.37; SE = .82	Case ID: 148M Ability Estimate = -1.24; SE = .37	Case ID: 174F Ability Estimate = 1.72; SE = .38

Figure 4.6

Phase Three Draw-a-Man Drawings: Examples of Least Successful, Mean, and Most Successful



Case ID: 186M
Ability Estimate = -4.92; SE = .71

Case ID: 110M
Ability Estimate = -1.11; SE = .36

Case ID: 100F
Ability Estimate = 2.44; SE = .39

Draw-a-Man Results from Analysis of Complete Data Set

Whilst unidimensionality of the DAM sub-test items was implied by the fit statistics produced by the separate Rasch analyses discussed in the previous sections, an analysis of all three phases of data together was also undertaken as a more rigorous investigation of this Rasch model principle. Indeed, Rasch analysis of the three phases of 'stacked' data provides a more thorough examination of the data than that provided by analyzing each data set individually (Bond & Fox, 2007; Drake, 1998).

As the use of the complete data set linked all data in one analytical framework, Rasch analysis of the whole DAM sub-test data set produced item difficulty and case ability estimates that were instrumental in investigating the development of the children's drawings over time. These results are described – together with those produced by the DAW sub-test results – in the 'Children's Drawing Development across the Phases' section later in this chapter. The key results produced by the Rasch analysis of the stacked DAM data set, however, are described below.

Variable Map

Figure 4.7 presents the variable map produced by the Rasch analysis of the stacked DAM sub-test data. This analysis included the 73 DAM items and all 246 'unique' cases from the three phases of data collection. Unlike the other variable maps that have been presented in this chapter, this variable map shows the estimated locations of items specifically, and those of the cases more generally. Since the focus here was on the items, rather than the cases, the case IDs were excluded from the variable maps. The children are represented on the left-hand side of the logit scale as ‘#’ which indicates that approximately 2 children were located at that estimated ability level.

Figure 4.7

Stacked Data Draw-a-Man Person-Item Variable Map

```
       Persons MAP OF Items
              <more>|<rare>
   6                +  59     61
                    |T
                    |
                    |  60     73
   5                +
                    |
                    |  21     69     71
                    |
   4                +
                    |  58
                    |  13     37     38
                    |  12     62     66
   3                +S
                 .  |  34     70     72
                    |
                    |  57     68
   2                +  26
                 . T|  15     20
                ##  |  16     27     52
               .##  |  23     32     42     65
   1            ##  +  41
           .######  |  17     49     7
               ###  |  5      56
             ##### S|  8
   0       #######  +M 22     25     3      43
         #########  |  31     48
             #####  |  10     29     40
          .#######  |  33     67
  -1         #####  +  19     36     64
             .#### M|  28     51     6
       ###########  |  45
          ########  |  54     63
  -2        .#####  +
     .############  |  2      24     9
          ########  |  14     53
          .####### S|  39     50     55
  -3            ##  +S 47
                ##  |
                .#  |  18
                .#  |
  -4                +
                 . T|
                 .  |
                    |
  -5             #  +  11
                    |
                 #  |
                    |T 44
  -6                +  30     46
                    |
                 .  |
                    |
  -7                +
                    |  35
                    |
                    |
  -8                +  1      4
              <less>|<frequ>
 EACH '#' IS 2.
```

The logit scale produced in Figure 4.7 is the most expansive of the DAM variable maps yet. However, at 14 logits (from -8 to +6 logits) it was still comparable to the other variable maps shown in Figures 4.1, 4.2 and 4.3 (which were 12, 12 and 11 logits in range respectively). Also resembling the DAM variable maps for phases one, two and three, the items in Figure 4.7 were well dispersed along the logit scale. Figure 4.7 shows items 1 (head present), 4 (eyes present), 11 (mouth present), 30 (arms present), 35 (legs present), and 44 (attachment of arms and legs) were located towards the bottom of the logit scale indicating that these items were amongst the least difficult to include in drawings of men. Also, the top of the logit scale shows that items 21 (hair IV), 59 (clothing V), 60 (profile I), 61 (profile II), 71 ('modeling' technique), and 73 (leg movement) were some of the most difficult items to include in drawings of men in the stacked data analysis. These results strongly resemble those produced by the Rasch analysis of the DAM data in phases one, two and three.

Item Fit Statistics

Table 4.10 displays the measure order item fit statistics generated from the Rasch analysis of the stacked DAM data. Out of a total of 73 DAM sub-test items, 10 were found to be underfitting according to the Rasch model's expectations: items 2 (neck present), 12 (lips two dimensions), 21 (hair IV), 22 (ears present), 25 (correct number of fingers shown), 34 (elbow joint shown), 48 (attachment of arms and legs II), 49 (proportion: head II), 50 (proportion: face) and 63 (motor coordination and lines). Most of these items yielded large outfit statistics suggesting that some children received, or did not receive, credit for these items rather unexpectedly. Indeed, Linacre (2002) states that underfitting outfit statistics can be indicative of "lucky guesses and careless mistakes … these are usually easy to diagnose and remedy" (p. 878). In other words, the estimated difficulty of the items suggested that some lower ability children would not include these features in their drawings of men, yet some of these children did; or alternatively, some higher-ability children unexpectedly did not include these features in their drawings of men (Linacre, 2002).

Whilst many of the underfitting items detected in the DAM stacked data analysis were also reported as misfitting in the individual phase analyses, some of these items were not. This was the first time that items 2 (infit: Mn Sq. = 1.22; $t = 2.90$; outfit: Mn Sq. = 2.03; $t = 5.20$), 34 (infit: Mn Sq. = 1.13, $t = 0.60$; outfit: Mn Sq. = 9.90, $t = 7.80$), and 63 (infit: Mn Sq. = 1.23, $t = 3.20$; outfit: Mn Sq. = 1.30, $t = 2.10$) were detected as underfitting, whereas items 12, 21, 22, 25, 48, 49, and 50 have repeatedly been reported as underfitting in the individual analyses of the DAM sub-test phases (see Tables 4.4, 4.5, and 4.6).

Table 4.10 reveals that nine DAM items were reported as overfitting in the stacked data Rasch analysis. These items were: 4 (eyes present), 31 (shoulders I), 32 (shoulders II), 33 (arms at side or engaged in activity), 36 (hip I (crotch)), 40 (feet II: proportion), 54 (proportion: limbs in two dimensions), 55 (clothing I), and 67 (directed lines and form: trunk outline). The fit statistics of each of these items suggest that children with lower ability predictability did not, and children with higher ability predictability did, receive credit for these items. Of these nine items, 33, 36, 40, 54, 55, and 67 were also reported as overfitting in the output produced by the separate Rasch analysis of each data collection phase as well.

Bond and Fox (2007) suggest that overfitting items should not necessarily be removed from a test; indeed, "omitting the overfitting items ($t < -2.0$ or Mn Sq. < 0.70)

could rob the test of its best items – the other items are not as good as these" (p. 241). Therefore, the results presented here, as well as those generated by the Rasch analyses of the individual phases were carefully considered in order to determine which items would, and would not be removed from the DAM sub-test in order to produce the best assessment of children's DAM drawings.

Table 4.10

Stacked Data Draw-a-Man Sub-Test Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
59	0	246	6.40	1.83	Maximum Estimated Measure			
61	0	246	6.40	1.83	Maximum Estimated Measure			
60	1	246	5.18	1.01	0.99	0.30	0.17	-1.70
73	1	246	5.18	1.01	1.00	0.30	0.24	-1.40
21	2	246	4.47	0.72	1.01	0.20	9.67	5.00
69	2	246	4.47	0.72	0.87	0.00	0.13	-1.80
71	2	246	4.47	0.72	0.89	0.10	0.22	-1.40
58	4	246	3.75	0.52	0.99	0.10	0.33	-1.00
13	5	246	3.51	0.46	0.89	-0.10	0.27	-1.30
37	5	246	3.51	0.46	0.99	0.10	0.29	-1.20
38	5	246	3.51	0.46	1.06	0.30	1.39	0.70
66	6	246	3.31	0.43	1.01	0.20	1.97	1.40
12	7	246	3.14	0.40	0.94	-0.10	2.92	2.20
62	7	246	3.14	0.40	0.92	-0.10	0.37	-1.10
72	9	246	2.86	0.35	1.01	0.10	0.50	-0.80
34	11	246	2.63	0.32	1.13	0.60	9.90	7.80
70	11	246	2.63	0.32	0.87	-0.40	0.31	-1.60
57	15	246	2.26	0.28	0.83	-0.80	0.32	-1.80
68	15	246	2.26	0.28	0.79	-1.00	0.27	-2.00
26	19	246	1.97	0.26	0.86	-0.70	0.45	-1.50
15	21	246	1.85	0.25	1.09	0.60	1.37	1.00
20	24	246	1.67	0.23	0.99	0.00	0.69	-0.80
16	25	246	1.62	0.23	1.28	1.70	1.72	1.70
52	26	246	1.57	0.23	1.08	0.60	0.88	-0.20
27	27	246	1.52	0.22	0.98	-0.10	0.79	-0.50
23	32	246	1.28	0.21	0.98	-0.10	0.95	-0.10
65	33	246	1.24	0.21	0.92	-0.60	1.10	0.40
32	35	246	1.16	0.20	0.76	-2.10	0.43	-2.30
42	35	246	1.16	0.20	0.86	-1.10	0.58	-1.50
41	42	246	0.89	0.19	0.95	-0.40	0.64	-1.40
7	43	246	0.85	0.19	0.81	-1.90	0.50	-2.20
49	46	246	0.75	0.18	1.58	4.90	2.33	3.80
17	48	246	0.68	0.18	1.11	1.10	1.14	0.60
5	52	246	0.55	0.18	1.10	1.00	1.04	0.30
56	57	246	0.39	0.17	0.80	-2.30	0.66	-1.60
8	63	246	0.22	0.17	0.83	-2.10	0.70	-1.50
43	71	246	0.00	0.16	0.85	-1.90	0.68	-1.80
3	73	246	-0.05	0.16	0.87	-1.60	0.78	-1.20
25	74	246	-0.08	0.16	1.36	4.10	1.59	2.80
22	75	246	-0.11	0.16	1.35	4.10	1.48	2.40
31	77	246	-0.16	0.16	0.77	-3.20	0.62	-2.40
48	80	246	-0.24	0.16	1.49	5.50	1.63	3.10
29	88	246	-0.43	0.16	1.00	0.10	0.88	-0.70
40	89	246	-0.46	0.16	0.75	-3.80	0.60	-2.90

10	94	246	-0.58	0.15	0.97	-0.40	0.90	-0.60
67	97	246	-0.65	0.15	0.82	-2.70	0.68	-2.30
33	102	246	-0.77	0.15	0.71	-4.50	0.64	-2.70
36	111	246	-0.98	0.15	0.75	-3.90	0.66	-2.80
19	112	246	-1.00	0.15	0.96	-0.50	0.91	-0.60
64	116	246	-1.09	0.15	1.14	2.00	1.20	1.40
6	123	246	-1.25	0.15	1.07	1.10	1.31	2.20
51	123	246	-1.25	0.15	1.14	2.00	1.13	1.00
28	126	246	-1.32	0.15	1.05	0.80	1.06	0.50
45	133	246	-1.48	0.15	0.91	-1.30	0.86	-1.10
54	143	246	-1.71	0.15	0.66	-5.80	0.53	-4.20
63	147	246	-1.80	0.15	1.23	3.20	1.30	2.10
9	162	246	-2.16	0.16	1.14	1.90	1.22	1.40
2	163	246	-2.18	0.16	1.22	2.90	2.03	5.20
24	168	246	-2.30	0.16	1.11	1.60	1.35	2.00
53	173	246	-2.43	0.16	1.05	0.60	1.05	0.40
14	174	246	-2.46	0.16	1.12	1.60	1.41	2.10
50	181	246	-2.65	0.17	1.49	5.40	2.87	6.60
55	187	246	-2.82	0.17	0.83	-2.00	0.57	-2.30
39	188	246	-2.85	0.17	0.87	-1.60	0.65	-1.80
47	194	246	-3.03	0.18	0.82	-2.00	0.60	-1.90
18	207	246	-3.48	0.20	0.94	-0.40	0.70	-1.10
11	233	246	-5.00	0.31	0.84	-0.60	1.27	0.60
44	239	246	-5.78	0.42	0.77	-0.60	0.40	-1.00
30	240	246	-5.96	0.45	0.84	-0.30	1.73	1.10
46	240	246	-5.96	0.45	0.94	0.00	0.39	-1.00
35	244	246	-7.22	0.74	1.04	0.30	0.42	-0.80
4	245	246	-7.97	1.03	1.03	0.30	0.09	-2.10
1	246	246	-9.22	1.85	Minimum Estimated Measure			
Mean	85.9	246	0.05	0.36	0.99	0.00	1.16	-0.10
SD	77.8	0	3.22	0.37	0.18	2.10	1.59	2.30
r	0.98							

The following section describes the results produced from the Rasch analysis of the DAW sub-test data collected over the three phases of data collection.

Draw-a-Woman Sub-Test Results

The DAW sub-test component of the GHDT comprised, what was generally, the second drawing completed by the children. As for the DAM sub-test, both male and female children completed this component (unlike the SPM and SPW sub-tests which were completed by only males and females correspondingly). The drawings of women were scored using the 71 point DAW scoring guide.

Initially, the three phases of data collected for DAW sub-test were analysed separately. The variable maps, summary statistics, and the item and case fit statistics produced are described in the following sub-sections. Following this are Figures 4.11, 4.12 and 4.13 which display examples of the least successful, mean and most successful drawings of women produced by children in each of the three phases. This DAW section closes with some key results of the Rasch analysis of the stacked DAW sub-test data.

Variable Maps

Figure 4.8 displays the first variable map produced for the DAW sub-test data. The logit scale produced for the DAW data covers a span of 13 logits (from -8 to +5), about the same range as that produced by the DAM data (which range from 11 to 14 logits). Similar to the DAM person-item variable maps, the DAW variable map has the 71 DAW items dispersed quite well along the equal-interval scale. The most difficult items are 22 (hair IV – use of directed lines to indicate part, texture, or combing. Superior style achieved) and 71 (directed lines and form: facial features – facial features must be symmetrical in all respects). Item 71 from the DAW sub-test aligns with one of the most difficult items on the DAM sub-test – item 69 (directed line and form: facial features).

Two of the least difficult items in the DAW sub-test were 1 (head present) and 4 (eyes present), exactly as revealed in the DAM sub-test variable maps

The DAW person-item variable map displays case IDs 100F, 174F and 196F as the three top performers. Case ID 130M located at -5.24 was revealed as the child who received the least amount of credit for his drawings of a woman. His DAM sub-test drawings were also amongst the least successful in the DAM sub-test analyses. However, case ID 130M's drawing of a woman in phase one attracted marginally more credit than his drawing of a man did, as he was credited for items 1 (head present), 4 (eyes present), 13 (mouth present), 19 (hair I), 33 (legs present), and 55 (trunk present). DAM items 11 (mouth present), and 18 (hair present) were not included in his drawing of a man in phase one.

Figure 4.8

Phase One Draw-a-Woman Person-Item Variable Map

```
                                             Persons -MAP- Items
                                                 <more>|<rare>
    5                                                  +  22  71
                                                       |
                                                       |
                                                       |  50  66  67  68  70
    4                                                  +
                                                       |
                                                       |  27  49  54
                                                       |
    3                                                  +
                                                       |  47
                                                       |S 30  45
                                                       |  12  18  61  69  8
    2                                                  +  31  48
                                           100 174 196  |  21  53
                                                   177 T|  16  65
                                                   150  |  23
    1                                      109 112 204  +  14  15
                               102 107 113 124 146 193  |  11  34  39  7
                                           116 122 166  |  51  62
                                   120 121 179 195 201  |  38  5
    0                                              155 S+M 25  29  52
                           101 125 154 156 172 175 206  |  10  3
                               108 127 151 153 202 205  |  37  46
                                               105 181  |  26  36  40  57  6
   -1                                  110 111 167 178  +  44  63
                                           117 136 145 M|  20  32  60
       104 106 119 123 126 173 180 182 192 194 197 203  |  64
                                   114 138 158 187 199  |  2   42
   -2              103 118 132 133 137 148 157 164 165  +  9
               115 135 139 140 142 147 152 163 183 198  |  17  28
                                   134 141 143 169 170  |S 35  58
                                   128 129 162 176 184 S|
   -3                      149 159 161 168 171 186 190  +  43
                                           185 188 200  |
                                                   144  |  56
                                               131 189  |  59
   -4                                                  +
                                                      T|  19
                                                   191  |  41
                                                   160  |
   -5                                                  +  55
                                                   130  |T
                                                       |  13  24
                                                       |
   -6                                                  +
                                                       |
                                                       |
                                                       |
   -7                                                  +  33
                                                       |
                                                       |
                                                       |
   -8                                                  +  1   4
                                                 <less>|<frequ>
```

Figures 4.9 and 4.10 present the variables maps produced by the Rasch analysis of the DAW sub-test collected in phases two and three. In the same way that the three variables maps produced for the DAM sub-test data showed remarkable similarities, so too do the DAW variable maps. The logit scale produced for phase two (12 logits from -7 to +5) and three (10 logits from -6 to +4) span a comparable number of logits as that produced for the phase one DAW data (13 logits from -8 to +5). And although the phase three DAW logit scale is the least expansive of the three phases it shows, perhaps, the most even spread of items across the entire logit scale that has been seen thus far (see Figure 4.10).

The three variable maps show that DAW items 1 (head present), 4 (eyes present), 13 (mouth present), 24 (arms present), 33 (legs present), and 55 (trunk present) were consistently reported as being amongst the least difficult items to include in drawings of women. On the other hand, items 22 (hair IV), 47 (neckline II: collar), 49 (waist II), 50 (skirt ‘modeled’ to indicate pleats or draping), 66 (directed lines and form: head outline), 68 (directed lines and form: hip contour), 69 (directed lines and form: arms taper), 70 (directed lines and form: calf of leg), and 71(directed lines and form: facial features) were consistently plotted towards the top of the logit scale indicating that they were amongst the most difficult items for these children to include in their drawings of women.

Figure 4.9

Phase Two Draw-a-Woman Person-Item Variable Map

```
                                     Persons -MAP- Items
                                         <more>|<rare>
    5                                          +  68  70
                                               |
                                               |
                                               |
                                               |
    4                                          +  27  47  49  66  69
                                               |
                                               |
                                               |  22  67  71
                                               |
    3                                          +
                                               |  50
                                               |S 45  54
                                               |
                                          174  |  65
    2                                          +  39
                                               |  61
                                          204 T|  12  16  30  53  62  8
                                     112  177  |  31
                                               |  11  15  21  23  52
    1                                          +  14  18  48
                                125  146  175  |  7
                                          172  |  5
                      101  116  124  196  205  |
                                     120  167 S|  29  38  51
    0                      113  122  155  166  +M 25  3   34  57
                           123  150  179  193  |  60
                           121  151  154  178  |  37  42
                                114  156  202  |  10  40
                           117  126  180  194  |  26  36  63
   -1                                          +  46
       110  136  147  157  170  176  192  199 M|
                                     173  181  |  32  6
       103  115  127  132  135  141  148  153  |  44  64
                                118  137  187  |  17  58
   -2                           165  171  182  +  2
                           133  139  158  163  |  20
                                152  160  190  |  28  9
                      128  149  169  186  189 S|S
                                129  159  184  |
   -3                                188  191  +  35  43
                      119  142  143  168  185  |
                                               |  56
                                               |
                                          130  |
   -4                                          +
                                              T|  59
                                          161  |
                                               |
                                               |  41
   -5                                          +
                                               |T
                                               |  19
                                               |
                                          131  |
   -6                                          +  13  33
                                               |
                                               |
                                               |
                                               |  4
   -7                                          +  1   24  55
                                         <less>|<frequ>
```

Figure 4.10

Phase Three Draw-a-Woman Person-Item Variable Map

```
                                Persons -MAP- Items
                                    <more>|<rare>
   4                                      +  66  70
                                          |  45  50  69  71
                                          |
                                          |
                                          |
                                          |  22  47
   3                                      +
                                          |
                                          |  16  27  30  49  54  68
                                          |S
                                          |  14  15  53
                                 100 174  |
   2                                      +  67
                                         T|  23  39
                                          |  12  31  61  65  8
                                          |  18
                                          |
                                          |  62
   1                         112 179 195  +  11
                             167 177 193  |  21
                                          |
                         123 150 166 196  |  29  3   48  52
                                     116 S|  38  7
                                 175 178  |  57
   0                             101 172  +M 25  51
                                 121 125  |  5
                                     141  |  10  32  42
                                     157  |  26  34  37  60
                                          |
                         110 120 122 199  |
  -1                                 194 M+  2
                             140 147 202  |  36  40
                                 171 180  |  44  6
                                 159 164  |  46
     118 129 135 139 152 176 182 189 192  |  28  63
                                     165  |  58
  -2                             115 185  +  64  9
                             169 188 191  |
                             131 143 168 S|
                                          |S
                                 160 173  |
                                     184  |
  -3                                      +  20
                                          |  17  43
                                          |
                                     163  |  35
                                         T|  59
                                          |
  -4                                 186  +
                                          |  19
                                          |
                                          |  41  56
                                          |
                                          |
  -5                                      +T
                                          |
                                          |
                                          |
                                          |  13  24  33
                                          |
  -6                                      +  1   4   55
                                    <less>|<frequ>
```

In the same ways that the DAW item locations have been remarkably similar across the three data collection phases, so too have the children's performances. For instance, the performances of case IDs 130M and 131M were reported as being amongst the least successful in phases one and two on the DAW sub-test. These two boys' performances on the DAM sub-test yielded similar ability estimate locations (see Figures 4.1, 4.2 and 4.3). In phase three, however, case ID 186M yielded an ability estimate of -4.05 which reported him to be the least successful on the DAW sub-test for this phase. Case ID 186M's previous ability estimates for phases one and two were a little higher than this at -3.07 and -2.45 respectively. In this instance, case ID 131M produced a drawing which received more credit than his previous drawings, and case ID 130M was absent from the sample.

At the other end of the scale, however, the drawings of women produced by case IDs 100F, 112F, 174F and 177F have yielded ability estimates (see Tables 4.17 to 4.19) which present them as amongst the most successful achievers on the DAW sub-test across all three phases. These girls performed similarly well on the DAM sub-test too (see Figures 4.1, 4.2 and 4.3).

Summary Statistics

Table 4.11 presents the summary statistics for the phase one DAW sub-test analysis. A review of the latter half of this table reveals the information pertinent to the DAW items. The mean of the item estimates was located at 0 (by default), and the standard deviation was 2.61 – about the same as that reported for the phase one DAM sub-test at 2.43. The reliability of these estimates was the same as that produced for the phase one DAM sub-test at .97. A total of 67 items were useful in the analysis, which produced an item mean raw score of 36.0 which is almost exactly the same as the phase one DAM item mean raw score of 36.4. The note underneath the information box provides a reminder that four items were excluded from the analysis as two were endorsed by all children, and the other two were not endorsed by any of the children in the phase one sample. This is slightly fewer than the eight items that were excluded from the DAM sub-test.

The top half of Table 4.11 displays the summary statistics of the 107 children in the sample. With a person mean raw score of 22.6 out of a possible 71 credits, the children appear to have produced drawings of women which received about the same amount of credit as their drawings of men (DAM sub-test person mean raw score = 22.1). This is significant as the DAW drawings were judged against some differing criteria (i.e., only 50 out of a total of 73 items are common to both the DAM and DAW scoring scale). The phase one DAW sub-test data analysis produced a person ability mean of -1.37 which indicates that the sub-test was about as challenging for these children as was the phase one DAM sub-test (M = -1.30). The phase one DAW sub-test produced a standard deviation of person estimates of 1.50 (phase one DAM SD = 1.49). Furthermore, the person reliability index is exactly the same as that produced for the phase one DAM sub-test at .92.

The strong relationships between the Rasch output data for the DAM and DAW items will be discussed in further detail in Chapter Five.

Table 4.11

Phase One Draw-a-Woman Sub-Test Summary Statistics

SUMMARY OF 107 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	22.6	67.0	-1.37	.39	.99	.0	1.05	.1
S.D.	10.5	.0	1.49	.06	.23	1.0	1.23	.9
MAX.	47.0	67.0	1.76	.68	1.81	3.7	9.87	5.2
MIN.	4.0	67.0	-5.24	.34	.59	-1.9	.12	-1.0

REAL RMSE .42 ADJ.SD 1.43 SEPARATION 3.40 Person RELIABILITY .92

MODEL RMSE .40 ADJ.SD 1.43 SEPARATION 3.58 Person RELIABILITY .93

S.E. OF Person MEAN = .14

Person RAW SCORE-TO-MEASURE CORRELATION = .99

CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .93

SUMMARY OF 67 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	36.0	107.0	.00	.39	.99	.0	1.05	.0
S.D.	31.4	.0	2.61	.23	.20	1.5	.85	1.5
MAX.	106.0	107.0	4.28	1.02	1.64	4.5	5.81	3.5
MIN.	1.0	107.0	-7.03	.23	.61	-3.7	.18	-2.9

REAL RMSE .47 ADJ.SD 2.56 SEPARATION 5.51 Item RELIABILITY .97

MODEL RMSE .46 ADJ.SD 2.57 SEPARATION 5.62 Item RELIABILITY .97

S.E. OF Item MEAN = .32

MAXIMUM EXTREME SCORE: 2 Items

MINIMUM EXTREME SCORE: 2 Items

UMEAN=.000 USCALE=1.000

Tables 4.12 and 4.13 present the summary statistics for phases two and three of the DAW sub-test. Review of the lower halves of these tables shows that all three phases have reported a similar number of measurable items in the analyses, being 67, 66 and 66 for phases one, two, and three respectively. The standard deviations for each phase were also similar at 2.61, 2.56 and 2.48 for phases one, two and three. Similarly, the item reliability indexes remain the same across phases one (r = .97), two (r = .96), and three (r = .95). One dissimilarity, however, was the mean raw scores for the DAW sub-test across the three phases: 36.0, 28.2 and 20.9 for phases one, two and three. Over the twelve month data collection timeframe one could assume that the children were becoming older, more experienced and, perhaps, more likely to produce detailed human figure drawings, which would produce an increase in the mean raw score for items over the phases. On the other hand, however, was the simple fact that the sample size decreased quite dramatically over the data collection period which could have brought this result.

The upper halves of Tables 4.12 and 4.13 present the summary statistics for the measured cases in phases two and three of the DAW sub-test data. The number of measurable children decreased across the three phases as the sample size had a rather large attrition rate (around 30% for each phase). The number of measurable cases was 107, 83 and 56 for phases one, two and three respectively. Given the rather large discrepancy in the number of measurable cases, however, the results produced by the three DAW Rasch analyses were reasonably comparable. The standard deviations for each phase were extremely similar with 1.49 for phase one, 1.45 for phase two, and 1.35 for phase three. The person reliability indexes across the phases were practically identical at r = .92, .91 and .91 for phases one, two and three respectively. The item mean raw scores for the three phases of DAW sub-test data were reasonably well-aligned across phases one (22.6), two (22.4), and three (24.6).

Table 4.12

Phase Two Draw-a-Woman Sub-Test Summary Statistics

SUMMARY OF 83 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	22.4	66.0	-1.25	.39	1.01	.0	.88	.0
S.D.	10.0	.0	1.45	.07	.24	1.1	.68	.7
MAX.	50.0	66.0	2.21	.76	1.74	3.7	5.05	3.2
MIN.	3.0	66.0	-5.86	.33	.61	-2.5	.25	-1.1

REAL RMSE .42 ADJ.SD 1.38 SEPARATION 3.26 Person RELIABILITY .91
MODEL RMSE .40 ADJ.SD 1.39 SEPARATION 3.47 Person RELIABILITY .92
S.E. OF Person MEAN = .16

DELETED: 24 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .98
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .00

SUMMARY OF 66 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	28.2	83.0	.00	.43	.98	.0	.88	.0
S.D.	24.5	.0	2.56	.23	.23	1.5	.72	1.5
MAX.	82.0	83.0	4.08	1.05	1.97	5.6	4.37	6.0
MIN.	1.0	83.0	-6.73	.26	.58	-2.9	.08	-1.9

REAL RMSE .50 ADJ.SD 2.51 SEPARATION 5.01 Item RELIABILITY .96
MODEL RMSE .49 ADJ.SD 2.52 SEPARATION 5.11 Item RELIABILITY .96
S.E. OF Item MEAN = .32

MAXIMUM EXTREME SCORE: 2 Items
MINIMUM EXTREME SCORE: 3 Items
UMEAN=.000 USCALE=1.000

Table 4.13

Phase Three Draw-a-Woman Sub-Test Summary Statistics

SUMMARY OF 56 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	24.6	66.0	-.94	.38	.99	.0	.93	.0
S.D.	10.0	.0	1.35	.04	.23	1.1	.72	.8
MAX.	49.0	66.0	2.10	.55	1.44	2.3	4.92	2.7
MIN.	7.0	66.0	-4.05	.34	.56	-2.2	.24	-1.3

REAL RMSE .40 ADJ.SD 1.29 SEPARATION 3.22 Person RELIABILITY .91
MODEL RMSE .38 ADJ.SD 1.29 SEPARATION 3.37 Person RELIABILITY .92
S.E. OF Person MEAN = .18

DELETED: 51 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .00

SUMMARY OF 66 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	20.9	56.0	.00	.49	.98	.0	.93	-.1
S.D.	17.2	.0	2.48	.22	.26	1.4	.79	1.3
MAX.	55.0	56.0	3.91	1.03	1.91	5.3	4.79	5.1
MIN.	1.0	56.0	-5.70	.31	.56	-3.3	.13	-2.7

REAL RMSE .55 ADJ.SD 2.42 SEPARATION 4.40 Item RELIABILITY .95
MODEL RMSE .54 ADJ.SD 2.43 SEPARATION 4.51 Item RELIABILITY .95
S.E. OF Item MEAN = .31

MAXIMUM EXTREME SCORE: 2 Items
MINIMUM EXTREME SCORE: 3 Items
UMEAN=.000 USCALE=1.000

While most of the locations of the items and persons were similar across each of the three variable maps produced by the Rasch analysis of the DAW sub-test data, it was important to examine the item and person fit statistics to find out which should be reported as misfitting the Rasch model's expectations. The item fit statistics are presented and described first, followed by the person fit statistics.

Item Fit Statistics

Table 4.14 displays the entry order item statistics produced by the Rasch analysis of the phase one DAW sub-test data. A total of 16 items yielded fit statistics which indicated that they were misfitting according to the expectations of the Rasch model. Ten of these items were considered overly erratic: item 5 (infit Mn Sq. = .99, *t* = -0.1; outfit Mn Sq. = 2.06, *t* 2.4); item 6 (infit Mn Sq. = 1.11, *t* = 0.9; outfit Mn Sq. = 1.58, *t* = 2.4); item 9 (infit Mn Sq. = 1.33, *t* = 3.0; oufit Mn Sq. = 1.48, *t* = 2.1); item 17 (infit Mn Sq. = 1.24, *t* = 2.2; outfit Mn Sq. = 1.67, *t* = 2.4); item 29 (infit Mn Sq. = 1.42, *t* = 2.9; outfit Mn Sq. = 1.64, *t* = 1.9); item 34 (infit Mn Sq. = 1.28, *t* = 1.7; outfit Mn Sq. = 2.29, *t* = 2.4); item 57 (infit Mn Sq. = 1.64, *t* = 4.5; outfit Mn Sq. = 1.97, *t* = 3.5); item 58 (infit Mn Sq. = 1.50, *t* = 4.0; outfit Mn Sq. = 2.25, *t* = 3.5); and item 64 (infit Mn Sq. = 1.23, *t* = 2.1; Outfit Mn Sq. = 1.51, *t* = 2.5).

Six items yielded statistics typical of more Guttman-like response patterns: item 25 (infit Mn Sq. = .70, *t* = -2.4; outfit Mn Sq. = .49, *t* = -1.9); item 26 (infit Mn Sq. = .69, *t* = -2.9; outfit Mn Sq. = .57, *t* = -2.4); item 36 (infit Mn Sq. = .68, *t* = -3.0; outfit Mn Sq. = .59, *t* = -2.2); item 40 (infit Mn Sq. = .73, *t* = -2.6; outfit Mn Sq. = .65, *t* = -1.9); and item 46 (infit: Mn Sq. = .74, *t* = -2.2; outfit: Mn Sq. = .58, *t* = -1.9).

Out of a total of 28 misfitting items from both the phase one DAM and DAW sub-tests combined, six items were common to both sub-tests. These were: item 6 (eye detail: pupil); item 9 (nose present); item 25/29 (DAM/DAW item numbers respectively) (correct number of fingers shown); item 36/34 (hip), item 40/36 (feet II: proportion); and item 49/57 (proportion: head II/ head-trunk proportion).

Table 4.14

Phase One Draw-a-Woman Sub-Test Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
22	0	107	5.51	Maximum Estimated Measure				
71	0	107	5.51	Maximum Estimated Measure				
50	1	107	4.28	0.96	0.30	0.18	-0.80	0.96
66	1	107	4.28	1.01	0.30	0.36	-0.40	1.01
67	1	107	4.28	1.04	0.40	0.81	0.20	1.04
68	1	107	4.28	1.04	0.40	1.02	0.40	1.04
70	1	107	4.28	1.01	0.30	0.36	-0.40	1.01
27	2	107	3.56	1.00	0.20	0.32	-0.50	1.00
49	2	107	3.56	0.94	0.10	0.25	-0.70	0.94
54	2	107	3.56	0.99	0.20	5.81	2.60	0.99
47	4	107	2.80	1.00	0.10	0.94	0.30	1.00
30	5	107	2.55	0.99	0.10	0.46	-0.50	0.99
45	5	107	2.55	1.04	0.20	0.51	-0.40	1.04
8	6	107	2.33	1.05	0.30	1.50	0.80	1.05
61	6	107	2.33	0.93	-0.10	0.89	0.10	0.93
69	6	107	2.33	1.00	0.10	0.54	-0.40	1.00
12	7	107	2.15	0.98	0.00	1.29	0.60	0.98
18	7	107	2.15	1.18	0.70	2.35	1.60	1.18
31	8	107	1.98	1.02	0.20	1.90	1.20	1.02
48	8	107	1.98	0.99	0.10	0.90	0.10	0.99
53	9	107	1.83	0.86	-0.50	0.38	-1.00	0.86
21	10	107	1.69	0.82	-0.80	0.48	-0.80	0.82
16	11	107	1.57	0.81	-0.90	0.34	-1.30	0.81
65	11	107	1.57	1.14	0.70	1.07	0.30	1.14
23	15	107	1.13	1.05	0.30	1.35	0.80	1.05
14	17	107	0.94	0.80	-1.20	0.42	-1.40	0.80
15	17	107	0.94	0.80	-1.20	0.42	-1.40	0.80
11	18	107	0.85	1.04	0.30	1.13	0.40	1.04
39	18	107	0.85	0.94	-0.30	0.51	-1.10	0.94
7	19	107	0.77	0.79	-1.40	0.64	-0.80	0.79
34	20	107	0.69	1.28	1.70	2.29	2.40	1.28
62	21	107	0.61	1.13	0.80	0.91	-0.10	1.13
51	24	107	0.38	0.74	-1.90	0.47	-1.70	0.74
5	25	107	0.30	0.99	-0.10	2.06	2.40	0.99
38	26	107	0.23	0.75	-1.90	0.49	-1.70	0.75
25	28	107	0.10	0.70	-2.40	0.49	-1.90	0.70
52	29	107	0.03	1.35	2.40	1.20	0.70	1.35
29	30	107	-0.04	1.42	2.90	1.64	1.90	1.42
10	34	107	-0.29	1.02	0.20	0.92	-0.20	1.02
3	35	107	-0.35	0.80	-1.60	0.69	-1.30	0.80
46	36	107	-0.41	0.74	-2.20	0.58	-1.90	0.74
37	39	107	-0.59	0.61	-3.70	0.47	-2.90	0.61
57	40	107	-0.65	1.64	4.50	1.97	3.50	1.64
6	42	107	-0.76	1.11	0.90	1.58	2.40	1.11

26	42	107	-0.76	0.69	-2.90	0.57	-2.40	0.69	
36	42	107	-0.76	0.68	-3.00	0.59	-2.20	0.68	
40	44	107	-0.87	0.73	-2.60	0.65	-1.90	0.73	
44	46	107	-0.98	0.88	-1.10	0.83	-0.90	0.88	
63	48	107	-1.09	0.90	-0.90	0.90	-0.50	0.90	
32	49	107	-1.15	1.02	0.20	0.91	-0.50	1.02	
60	49	107	-1.15	1.12	1.10	1.08	0.50	1.12	
20	53	107	-1.36	0.97	-0.20	0.84	-0.80	0.97	
64	56	107	-1.52	1.23	2.10	1.51	2.50	1.23	
2	58	107	-1.63	1.10	1.00	1.03	0.20	1.10	
42	61	107	-1.79	1.07	0.70	1.30	1.50	1.07	
9	65	107	-2.01	1.33	3.00	1.48	2.10	1.33	
28	69	107	-2.23	1.03	0.40	1.17	0.80	1.03	
17	70	107	-2.29	1.24	2.20	1.67	2.40	1.24	
35	73	107	-2.46	0.88	-1.10	0.90	-0.30	0.88	
58	75	107	-2.58	1.50	4.00	2.25	3.50	1.50	
43	81	107	-2.95	0.78	-1.80	0.53	-1.60	0.78	
56	88	107	-3.47	0.80	-1.20	0.54	-1.10	0.80	
59	91	107	-3.73	1.24	1.30	2.54	2.30	1.24	
19	96	107	-4.25	0.91	-0.30	0.51	-0.70	0.91	
41	98	107	-4.52	0.79	-0.70	0.39	-0.90	0.79	
55	101	107	-5.03	0.92	-0.10	2.99	1.90	0.92	
13	103	107	-5.51	1.14	0.50	0.90	0.20	1.14	
24	103	107	-5.51	0.83	-0.30	0.49	-0.40	0.83	
33	106	107	-7.03	1.07	0.40	0.75	0.10	1.07	
1	107	107	-8.27	Minimum Estimated Measure					
4	107	107	-8.27	Minimum Estimated Measure					
Mean	37	107	-0.08	0.99	0.00	1.05	0.00	0.99	
SD	33.3	0	3.03	0.20	1.50	0.85	1.50	0.20	
r	0.97								

Table 4.15

Phase Two Draw-a-Woman Sub-Test Measure Order Item Statistics

Item	Raw		Difficulty	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
68	0	83	5.31	1.84	Maximum Estimated Measure			
70	0	83	5.31	1.84	Maximum Estimated Measure			
27	1	83	4.08	1.02	1.06	0.40	0.83	0.30
47	1	83	4.08	1.02	0.92	0.20	0.14	-0.70
49	1	83	4.08	1.02	0.92	0.20	0.14	-0.70
66	1	83	4.08	1.02	0.97	0.30	0.20	-0.60
69	1	83	4.08	1.02	0.97	0.30	0.20	-0.60
22	2	83	3.34	0.74	0.92	0.10	0.23	-0.60
67	2	83	3.34	0.74	0.99	0.20	0.33	-0.40
71	2	83	3.34	0.74	0.78	-0.20	0.15	-0.80
50	3	83	2.89	0.61	0.79	-0.30	0.20	-0.80
45	4	83	2.56	0.54	1.14	0.50	0.61	-0.10
54	4	83	2.56	0.54	1.13	0.40	1.18	0.50
65	5	83	2.29	0.49	1.02	0.20	1.20	0.50
39	6	83	2.07	0.46	0.82	-0.50	0.38	-0.80
61	7	83	1.87	0.43	1.04	0.20	0.66	-0.30
8	8	83	1.70	0.41	1.03	0.20	1.25	0.60
12	8	83	1.70	0.41	1.01	0.10	1.20	0.50
16	9	83	1.54	0.39	0.87	-0.50	0.41	-1.00
30	9	83	1.54	0.39	0.93	-0.20	0.54	-0.70
53	9	83	1.54	0.39	0.77	-0.90	0.36	-1.20
62	9	83	1.54	0.39	0.94	-0.20	0.45	-0.90
31	10	83	1.40	0.37	1.22	1.00	1.23	0.60
21	11	83	1.26	0.36	0.82	-0.80	0.42	-1.20
11	12	83	1.14	0.35	1.06	0.40	1.28	0.70
15	12	83	1.14	0.35	0.80	-0.90	0.49	-1.00
23	12	83	1.14	0.35	1.04	0.30	1.16	0.50
52	12	83	1.14	0.35	0.80	-0.90	0.64	-0.60
14	13	83	1.02	0.34	0.85	-0.70	0.62	-0.70
48	13	83	1.02	0.34	1.00	0.10	0.60	-0.80
18	14	83	0.91	0.33	1.01	0.10	1.20	0.60
7	16	83	0.70	0.32	0.88	-0.60	0.61	-0.90
5	17	83	0.60	0.31	0.90	-0.60	0.68	-0.70
29	22	83	0.16	0.29	1.33	2.10	1.38	1.10
38	22	83	0.16	0.29	1.07	0.50	0.87	-0.30
51	22	83	0.16	0.29	-0.67	2.60	0.51	-1.60
3	23	83	0.08	0.28	-0.65	2.80	0.58	-1.30
25	23	83	0.08	0.28	-0.66	2.70	0.48	-1.80
34	25	83	-0.08	0.28	1.97	5.60	3.19	4.30
57	25	83	-0.08	0.28	1.46	3.00	1.84	2.10
60	26	83	-0.16	0.28	1.05	0.40	4.37	6.00
37	28	83	-0.31	0.27	-0.85	1.10	0.69	-1.10
42	29	83	-0.38	0.27	-0.71	2.50	0.70	-1.00
10	31	83	-0.53	0.27	1.06	0.50	0.99	0.10

40	31	83	-0.53	0.27	-0.83	1.40	0.68	-1.30
63	34	83	-0.74	0.26	0.92	-0.60	0.84	-0.60
26	35	83	-0.81	0.26	-0.68	2.90	0.60	-1.90
36	35	83	-0.81	0.26	0.93	-0.60	0.81	-0.70
46	37	83	-0.95	0.26	-0.74	2.40	0.64	-1.70
6	43	83	-1.35	0.26	1.14	1.20	1.11	0.60
32	43	83	-1.35	0.26	1.22	1.80	1.49	2.10
44	46	83	-1.55	0.26	0.94	-0.50	1.08	0.40
64	48	83	-1.69	0.26	1.18	1.50	1.40	1.70
17	49	83	-1.76	0.26	1.30	2.40	1.46	1.80
58	50	83	-1.83	0.26	1.52	3.90	2.36	4.30
2	53	83	-2.04	0.27	1.20	1.60	1.08	0.40
20	55	83	-2.18	0.27	-0.88	1.00	0.77	-0.80
9	58	83	-2.40	0.28	1.16	1.20	1.38	1.20
28	58	83	-2.40	0.28	1.31	2.30	2.07	2.90
35	65	83	-2.99	0.30	-0.85	1.00	0.58	-1.10
43	65	83	-2.99	0.30	-0.79	1.40	0.56	-1.20
56	69	83	-3.38	0.33	0.97	-0.10	1.89	1.60
59	75	83	-4.19	0.41	0.90	-0.20	0.59	-0.40
41	78	83	-4.80	0.51	0.85	-0.30	0.54	-0.30
19	80	83	-5.45	0.64	0.86	-0.10	0.40	-0.40
13	81	83	-5.93	0.77	0.68	-0.40	0.14	-0.90
33	81	83	-5.93	0.77	0.58	-0.60	0.08	-1.10
4	82	83	-6.73	1.05	1.12	0.40	0.25	-0.50
1	83	83	-8.02	1.85	Minimum Estimated Measure			
24	83	83	-8.02	1.85	Minimum Estimated Measure			
55	83	83	-8.02	1.85	Minimum Estimated Measure			
Mean	29.7	83	-0.19	0.53	0.98	0.00	0.88	0.00
SD	26.6	0	3.10	0.43	0.23	1.50	0.72	1.50
r	0.96							

Table 4.15 displays the item fit statistics for the phase two DAW data. Ten of the 71 items were identified as misfitting. The five items reported as having too much variation were: 17 (infit: Mn Sq. = 1.30, *t* = 4.2; outfit: Mn Sq. = 1.46, *t* = 1.8); 28 (infit: Mn Sq. = 1.31, *t* = 2.3; outfit: Mn Sq. = 2.07, *t* = 2.9); 34 (infit: Mn Sq. = 1.97, *t* = 5.6; outfit: Mn Sq. = 5.6; outfit: Mn Sq. = 3.19, *t* = 4.3); 57 (infit: Mn Sq. = 1.46, *t* = 3.0; outfit: Mn Sq. = 1.84, *t* = 2.1); and 58 (infit: Mn Sq. = 1.52, *t* = 3.9; outfit: Mn Sq. = 2.36, *t* = 4.3).

The remaining five misfitting items yielded statistics typical of more Guttman-like response patterns. These items were: 3 (infit: Mn Sq. = .65, *t* = -2.8; outfit: Mn Sq. =

.58, $t = -1.3$); 25 (infit: Mn Sq. = .66, $t = -2.7$; outfit: Mn Sq. = .48, $t = -1.8$); 26 (infit: Mn Sq. = .68, $t = -2.9$; outfit: Mn Sq. = .60, $t = -1.9$); 46 (infit: Mn Sq. = .74, $t = -2.4$; outfit: Mn Sq. = .64, $t = -1.7$); and 51 (infit: Mn Sq. = .67, $t = -2.6$; outfit: Mn Sq. = .51, $t = -1.6$).

Table 4.16

Phase Three Draw-a-Woman Sub-Test Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
66	0	56	5.16	1.84	Maximum Estimated Measure			
70	0	56	5.16	1.84	Maximum Estimated Measure			
45	1	56	3.91	1.03	0.82	0.10	0.13	-0.50
50	1	56	3.91	1.03	0.82	0.10	0.13	-0.50
69	1	56	3.91	1.03	0.82	0.10	0.13	-0.50
71	1	56	3.91	1.03	1.01	0.30	0.34	-0.10
22	2	56	3.16	0.75	0.95	0.10	4.79	2.00
47	2	56	3.16	0.75	0.85	-0.10	0.26	-0.40
16	3	56	2.69	0.63	0.88	-0.10	0.33	-0.50
27	3	56	2.69	0.63	0.90	-0.10	0.36	-0.40
30	3	56	2.69	0.63	0.91	0.00	0.41	-0.40
49	3	56	2.69	0.63	0.69	-0.60	0.21	-0.80
54	3	56	2.69	0.63	0.95	0.10	0.84	0.20
68	3	56	2.69	0.63	1.15	0.50	1.08	0.40
14	4	56	2.34	0.56	0.91	-0.10	0.39	-0.60
15	4	56	2.34	0.56	0.91	-0.10	0.39	-0.60
53	4	56	2.34	0.56	0.85	-0.30	0.32	-0.70
67	5	56	2.06	0.51	1.20	0.70	1.85	1.10
23	6	56	1.82	0.47	0.85	-0.40	0.73	-0.10
39	6	56	1.82	0.47	0.76	-0.70	0.37	-0.90
8	7	56	1.61	0.45	1.09	0.40	0.60	-0.40
12	7	56	1.61	0.45	0.95	-0.10	0.66	-0.30
31	7	56	1.61	0.45	1.09	0.40	0.76	-0.10
61	7	56	1.61	0.45	0.93	-0.20	0.61	-0.40
65	7	56	1.61	0.45	0.85	-0.50	1.18	0.50
18	8	56	1.42	0.43	1.52	1.90	2.55	2.00
62	9	56	1.25	0.41	0.82	-0.70	0.56	-0.70
11	11	56	0.93	0.38	0.92	-0.30	0.69	-0.50
21	12	56	0.79	0.37	0.77	-1.20	0.84	-0.20
3	14	56	0.53	0.36	0.88	-0.60	0.74	-0.60
29	14	56	0.53	0.36	1.23	1.20	1.52	1.30
48	14	56	0.53	0.36	0.92	-0.40	0.81	-0.40
52	14	56	0.53	0.36	1.06	0.40	0.94	0.00
7	15	56	0.40	0.35	0.89	-0.60	0.76	-0.60
38	15	56	0.40	0.35	0.88	-0.60	0.68	-0.90
57	17	56	0.17	0.34	1.84	4.00	2.21	3.00
25	18	56	0.05	0.34	0.83	-1.00	0.63	-1.30
51	18	56	0.05	0.34	0.79	-1.30	0.71	-1.00
5	20	56	-0.17	0.33	0.96	-0.20	0.89	-0.30
32	21	56	-0.27	0.32	1.17	1.10	1.16	0.70
10	22	56	-0.38	0.32	0.99	0.00	1.00	0.10
42	22	56	-0.38	0.32	0.56	-3.30	0.47	-2.70
34	23	56	-0.48	0.32	1.91	4.80	2.34	4.30
37	23	56	-0.48	0.32	0.77	-1.60	0.67	-1.50

26	24	56	-0.58	0.32	0.73	-1.90	0.70	-1.40
60	24	56	-0.58	0.32	0.89	-0.70	0.87	-0.50
2	29	56	-1.08	0.31	0.91	-0.60	0.88	-0.50
36	30	56	-1.17	0.31	0.81	-1.50	0.86	-0.60
40	30	56	-1.17	0.31	0.65	-2.90	0.54	-2.40
6	31	56	-1.27	0.31	0.74	-2.10	0.61	-1.90
44	31	56	-1.27	0.31	0.91	-0.70	0.84	-0.60
46	34	56	-1.57	0.32	0.79	-1.70	0.71	-1.10
28	35	56	-1.67	0.32	1.32	2.20	1.59	1.90
63	35	56	-1.67	0.32	1.15	1.10	1.09	0.40
58	37	56	-1.87	0.32	1.90	5.30	3.40	5.10
9	38	56	-1.98	0.33	1.00	0.00	0.98	0.10
64	38	56	-1.98	0.33	1.17	1.20	1.23	0.80
20	46	56	-2.95	0.38	0.97	-0.10	0.69	-0.40
17	47	56	-3.10	0.40	1.15	0.70	1.85	1.40
43	47	56	-3.10	0.40	0.83	-0.70	0.55	-0.70
35	49	56	-3.45	0.44	0.73	-0.90	0.41	-0.80
59	50	56	-3.65	0.46	1.18	0.60	0.84	0.10
19	52	56	-4.15	0.55	0.89	-0.10	0.52	-0.30
41	53	56	-4.49	0.62	1.24	0.60	2.30	1.30
56	53	56	-4.49	0.62	1.08	0.30	0.54	-0.20
13	55	56	-5.70	1.02	1.05	0.40	0.70	0.20
24	55	56	-5.70	1.02	1.09	0.40	1.69	0.90
33	55	56	-5.70	1.02	0.90	0.20	0.18	-0.50
1	56	56	-6.94	1.83	Minimum Estimated Measure			
4	56	56	-6.94	1.83	Minimum Estimated Measure			
55	56	56	-6.94	1.83	Minimum Estimated Measure			
Mean	21.8	56	-0.15	0.59	0.98	0.00	0.93	-0.10
SD	18.4	0	2.92	0.40	0.26	1.40	0.79	1.30
r	0.95							

Table 4.16 presents the phase three DAW measure order item statistics. In this instance, only seven DAW items were reported as misfitting the Rasch model's expectations. The overly erratic items were: 18 (infit: Mn Sq. = 1.52, t = 1.9; outfit: Mn Sq. = 2.55, t = 2.0); 22 (Infit: Mn Sq. = .95, t = .1; outfit: Mn Sq. = 4.79, t = 2.0); 34 (infit: Mn Sq. = 1.91, t = 4.8; outfit: Mn Sq. = 2.34, t = 4.3); 57 (infit: Mn Sq. = 1.84, t = 4.0; outfit: Mn Sq. 2.21, t = 3.0); 58 (infit: Mn Sq. = 1.90, t = 5.3; outfit: Mn Sq. = 3.4, t = 5.1). The two overfitting items were: 40 (infit: Mn Sq. = .65, t = -2.9; outfit: Mn Sq. = .54, t = -2.4); and 42 (infit: Mn Sq. = .56, t = -3.3; outfit: .47, t = -2.7).

Items 34, 40, 42, 57 and 58 were also detected as misfitting in the Rasch analysis of the DAW sub-test data in phases one and two. The DAW items 34 (hip), 57 (head-trunk proportion) and 58 (head: proportion) have equivalent DAM items (36, 49, and 50 respectively) which were also reported as misfitting in the DAM Rasch analyses. Hence, there is an underlying qualitative aspect of these items that the Rasch model has detected quantitatively, as these items were continually reported as misfitting according to the expectations of the Rasch model, in *both* the DAM and DAW sub-tests.

Obviously, the item fit statistics produced by the Rasch model are extremely useful to the aims of this thesis. In order to further develop the GHDT, items need to be examined, and then modified, refined, or removed completely if they do not contribute meaningfully to the investigation of the latent construct. While the examination of the item statistics has been helpful, it is also important to examine the case statistics to find out how well the DAW cases performances fit the Rasch model's expectations.

Case Statistics

Table 4.17 presents the measure order person statistics for the Rasch analysis of the phase one DAW sub-test. The fit statistics indicate that only one child presented with a response string that was considered more erratic than expected by the Rasch model. The single 'noisy' case was: 102F (infit: Mn Sq. = 1.81, t = 3.7; outfit: Mn Sq. = 8.22, t = 5.2). Case 102F's outfit t of 5.2 indicates some 620% more variation in her response string than was predicted by the Rasch model. The DAM person statistics in Tables 4.7, 4.8 and 4.9 show that case ID 102F was not reported as misfitting in the DAM sub-test.

None of the children's DAW sub-test performances from phase one were reported to be overly predictable or Guttman-like.

Table 4.17

Phase One Draw-a-Woman Sub-Test Measure Order Person Statistics

Case #	Raw Score	Count	Ability Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
196 F	47	67	1.76	0.36	0.87	-0.60	1.11	0.40
100 F	46	67	1.63	0.36	1.11	0.60	1.00	0.20
174 F	46	67	1.63	0.36	0.95	-0.20	1.02	0.20
177 F	44	67	1.38	0.35	1.16	0.90	0.95	0.10
150 M	43	67	1.25	0.35	1.18	1.00	0.88	0.00
109 M	40	67	0.89	0.35	0.81	-1.10	0.52	-0.80
112 F	40	67	0.89	0.35	0.95	-0.20	0.90	0.00
204 F	40	67	0.89	0.35	0.97	-0.10	0.93	0.10
113 F	39	67	0.77	0.34	0.85	-0.80	0.57	-0.60
124 F	39	67	0.77	0.34	0.90	-0.50	0.96	0.10
102 F	38	67	0.65	0.34	1.81	3.70	8.22	5.20
107 F	38	67	0.65	0.34	0.77	-1.30	0.55	-0.70
146 F	38	67	0.65	0.34	0.94	-0.30	1.98	1.50
193 F	38	67	0.65	0.34	1.28	1.50	1.44	0.90
116 F	37	67	0.54	0.34	0.96	-0.20	1.07	0.30
166 F	37	67	0.54	0.34	1.00	0.10	0.75	-0.30
122 F	36	67	0.42	0.34	0.86	-0.80	0.93	0.10
120 F	35	67	0.30	0.34	1.17	1.00	0.93	0.10
121 F	35	67	0.30	0.34	1.05	0.40	0.92	0.10
179 F	34	67	0.19	0.34	1.08	0.50	0.89	0.00
195 F	34	67	0.19	0.34	1.13	0.80	0.84	-0.10
201 F	34	67	0.19	0.34	1.13	0.80	1.11	0.40
155 F	33	67	0.07	0.34	1.13	0.80	1.00	0.20
125 F	31	67	-0.16	0.34	0.82	-1.00	0.67	-0.40
154 F	31	67	-0.16	0.34	1.24	1.30	1.11	0.40
156 F	31	67	-0.16	0.34	1.24	1.30	2.10	1.60
101 F	30	67	-0.28	0.34	0.95	-0.20	0.68	-0.40
172 M	30	67	-0.28	0.34	1.12	0.70	1.04	0.30
175 F	30	67	-0.28	0.34	1.20	1.20	1.22	0.60
206 M	30	67	-0.28	0.34	0.99	0.00	0.70	-0.40
108 M	29	67	-0.40	0.34	1.32	1.70	1.29	0.60
153 F	29	67	-0.40	0.34	1.21	1.20	2.37	1.90
127 F	28	67	-0.52	0.35	1.24	1.40	0.99	0.20
151 F	28	67	-0.52	0.35	0.91	-0.50	0.68	-0.40
202 M	28	67	-0.52	0.35	1.00	0.10	0.88	0.00
205 F	28	67	-0.52	0.35	1.17	1.00	0.86	0.00
105 F	27	67	-0.64	0.35	1.23	1.30	0.88	0.00
181 M	26	67	-0.76	0.35	0.98	-0.10	1.07	0.30
110 M	24	67	-1.01	0.36	0.78	-1.30	0.59	-0.40
111 M	24	67	-1.01	0.36	0.78	-1.30	0.59	-0.40
167 F	24	67	-1.01	0.36	1.01	0.10	0.84	0.00
178 M	24	67	-1.01	0.36	0.98	-0.10	0.57	-0.50
117 M	22	67	-1.26	0.36	0.70	-1.80	0.42	-0.70
136 F	22	67	-1.26	0.36	0.74	-1.50	0.43	-0.70

145 F	22	67	-1.26	0.36	1.00	0.00	0.63	-0.30
104 F	21	67	-1.40	0.37	0.78	-1.20	0.42	-0.70
106 M	21	67	-1.40	0.37	0.98	0.00	0.60	-0.30
119 M	21	67	-1.40	0.37	0.76	-1.30	0.41	-0.70
123 F	21	67	-1.40	0.37	1.21	1.10	1.05	0.30
126 F	21	67	-1.40	0.37	1.33	1.60	2.11	1.30
182 M	21	67	-1.40	0.37	0.93	-0.30	0.60	-0.30
192 M	21	67	-1.40	0.37	0.70	-1.70	0.38	-0.80
173 M	20	67	-1.53	0.37	0.96	-0.20	0.52	-0.40
180 M	20	67	-1.53	0.37	0.86	-0.70	0.64	-0.20
194 F	20	67	-1.53	0.37	0.74	-1.40	0.38	-0.70
197 M	20	67	-1.53	0.37	0.99	0.00	0.64	-0.20
203 M	20	67	-1.53	0.37	0.94	-0.30	1.58	0.90
114 M	19	67	-1.68	0.38	0.85	-0.70	0.52	-0.40
199 F	19	67	-1.68	0.38	0.93	-0.30	0.72	-0.10
138 F	18	67	-1.82	0.39	1.18	0.90	1.12	0.40
158 F	18	67	-1.82	0.39	0.98	0.00	2.00	1.20
187 F	18	67	-1.82	0.39	0.64	-1.90	0.34	-0.90
103 M	17	67	-1.97	0.39	0.73	-1.30	0.55	-0.40
118 M	17	67	-1.97	0.39	0.91	-0.30	0.46	-0.50
132 M	17	67	-1.97	0.39	0.77	-1.10	0.43	-0.60
133 F	17	67	-1.97	0.39	0.89	-0.40	0.44	-0.60
137 M	17	67	-1.97	0.39	1.25	1.10	0.98	0.30
148 M	17	67	-1.97	0.39	0.94	-0.20	0.65	-0.20
157 F	17	67	-1.97	0.39	0.98	0.00	0.60	-0.30
164 F	17	67	-1.97	0.39	0.93	-0.30	0.98	0.30
165 F	17	67	-1.97	0.39	0.81	-0.90	0.43	-0.60
147 F	16	67	-2.13	0.40	0.67	-1.60	0.32	-0.80
152 F	16	67	-2.13	0.40	0.76	-1.10	1.37	0.70
198 M	16	67	-2.13	0.40	0.81	-0.80	0.45	-0.50
115 M	15	67	-2.30	0.41	0.66	-1.50	0.32	-0.80
135 F	15	67	-2.30	0.41	1.17	0.80	2.08	1.30
139 F	15	67	-2.30	0.41	0.94	-0.20	0.42	-0.60
140 M	15	67	-2.30	0.41	0.76	-1.00	0.34	-0.80
142 M	15	67	-2.30	0.41	0.69	-1.40	0.34	-0.80
163 M	15	67	-2.30	0.41	1.23	1.00	0.95	0.20
183 F	15	67	-2.30	0.41	0.81	-0.70	0.47	-0.50
134 M	14	67	-2.48	0.42	0.74	-1.00	0.37	-0.70
141 F	14	67	-2.48	0.42	1.48	1.70	3.07	1.90
143 M	14	67	-2.48	0.42	0.70	-1.30	0.35	-0.80
169 F	14	67	-2.48	0.42	0.86	-0.50	0.68	-0.10
170 M	14	67	-2.48	0.42	1.08	0.40	0.91	0.20
128 M	13	67	-2.66	0.44	1.37	1.30	1.82	1.10
129 M	13	67	-2.66	0.44	0.99	0.10	0.46	-0.50
176 M	13	67	-2.66	0.44	0.80	-0.70	2.07	1.20
162 F	12	67	-2.86	0.45	1.55	1.80	9.87	4.10
184 M	12	67	-2.86	0.45	1.44	1.50	0.98	0.30
149 M	11	67	-3.07	0.47	0.74	-0.90	1.23	0.60
159 F	11	67	-3.07	0.47	0.78	-0.70	0.47	-0.40
161 F	11	67	-3.07	0.47	1.14	0.60	2.67	1.50

168 F	11	67	-3.07	0.47	0.99	0.10	0.99	0.30
171 F	11	67	-3.07	0.47	0.59	-1.50	0.78	0.10
186 M	11	67	-3.07	0.47	1.54	1.70	1.10	0.40
190 F	11	67	-3.07	0.47	0.92	-0.20	0.92	0.20
185 F	10	67	-3.30	0.49	0.82	-0.50	0.89	0.20
188 M	10	67	-3.30	0.49	0.77	-0.70	0.29	-0.70
200 M	10	67	-3.30	0.49	1.12	0.50	0.75	0.10
144 M	9	67	-3.55	0.51	0.65	-1.10	0.25	-0.70
131 M	8	67	-3.82	0.53	1.21	0.70	0.50	-0.20
189 M	8	67	-3.82	0.53	1.41	1.20	0.53	-0.20
191 F	6	67	-4.45	0.59	1.18	0.60	1.95	1.10
160 F	5	67	-4.82	0.63	1.59	1.40	1.37	0.70
130 M	4	67	-5.24	0.68	0.71	-0.60	0.12	-1.00
Mean	22.6	67.0	-1.37	0.39	0.99	0.00	1.05	0.10
SD	10.5	0	1.49	0.06	0.23	1.00	1.23	0.90
r	0.97							

Table 4.18 displays the measure order person statistics for the phase two DAW sub-test data analysis. A total of five cases were detected as misfitting in this phase, with three reported as too unpredictable and the remaining two reported as too predictable according to the Rasch model's expectations. The three marginally 'noisy' cases were: 112F (infit: Mn Sq. = 1.74, *t* = 3.7; outfit: Mn Sq. = 3.84, *t* = 3.2); 131M (infit: Mn Sq. = 1.65, *t* = 1.4; outfit: Mn Sq. = 5.05, *t* = 2.2); and 167F (infit: Mn Sq. = 1.67, *t* 3.4; outfit: Mn Sq. = 1.63, *t* = 1.1). The two slightly overfitting cases were: 178F (infit: Mn Sq. = .62, *t* = -2.5; outfit: Mn Sq. = .39, *t* = -1.1), and 180M (infit: Mn Sq. = .61, *t* = -2.4; outfit: Mn Sq. = .35, *t* = -1.1).

There were no common misfitting cases between the phase one and two DAW sub-test results. The only common misfitting case for this phase two DAW sub-test analysis and the other sub-tests analyses thus far was case ID 180M, who was detected as misfitting in the phase one DAM sub-test results.

Table 4.18

Phase Two Draw-a-Woman Sub-Test Measure Order Person Statistics

Case #	Raw Score	Count	Ability Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
174 F	50	66	2.21	0.37	0.92	-0.40	1.29	0.60
204 F	46	66	1.69	0.35	1.37	1.90	1.22	0.50
112 F	43	66	1.33	0.34	1.74	3.70	3.87	3.20
177 F	43	66	1.33	0.34	1.03	0.20	0.69	-0.40
125 F	39	66	0.88	0.34	1.22	1.30	1.01	0.20
175 F	39	66	0.88	0.34	1.03	0.30	1.08	0.30
146 F	38	66	0.76	0.34	1.01	0.10	0.68	-0.40
172 M	37	66	0.65	0.33	0.93	-0.40	0.82	-0.10
101 F	35	66	0.43	0.33	0.96	-0.20	1.08	0.30
116 F	35	66	0.43	0.33	1.11	0.70	1.06	0.30
124 F	35	66	0.43	0.33	1.10	0.70	1.08	0.30
205 F	35	66	0.43	0.33	1.24	1.40	1.34	0.70
196 F	34	66	0.32	0.34	1.19	1.10	0.93	0.10
120 F	33	66	0.20	0.34	1.11	0.70	0.85	-0.10
167 F	33	66	0.20	0.34	1.67	3.40	1.63	1.10
113 F	32	66	0.09	0.34	1.09	0.60	0.96	0.10
122 F	32	66	0.09	0.34	1.05	0.30	0.82	-0.10
155 F	32	66	0.09	0.34	0.75	-1.60	0.49	-0.80
166 F	32	66	0.09	0.34	1.03	0.20	0.79	-0.20
150 M	30	66	-0.14	0.34	1.03	0.20	0.87	0.00
179 F	30	66	-0.14	0.34	1.08	0.50	0.78	-0.20
193 F	30	66	-0.14	0.34	1.32	1.80	2.10	1.60
123 F	29	66	-0.25	0.34	0.81	-1.10	0.59	-0.50
121 F	28	66	-0.37	0.34	0.87	-0.70	0.59	-0.50
151 F	28	66	-0.37	0.34	1.07	0.50	0.94	0.10
154 F	28	66	-0.37	0.34	0.91	-0.50	0.74	-0.20
178 M	28	66	-0.37	0.34	0.62	-2.50	0.39	-1.10
114 M	26	66	-0.61	0.35	0.96	-0.20	0.70	-0.30
156 F	26	66	-0.61	0.35	0.87	-0.70	0.52	-0.70
202 M	26	66	-0.61	0.35	0.86	-0.80	0.58	-0.60
117 M	25	66	-0.73	0.35	1.03	0.20	0.88	0.00
194 F	25	66	-0.73	0.35	0.79	-1.20	0.60	-0.50
126 F	24	66	-0.86	0.36	0.87	-0.70	0.53	-0.60
180 M	24	66	-0.86	0.36	0.61	-2.40	0.35	-1.10
170 M	22	66	-1.12	0.37	0.99	0.00	0.69	-0.30
176 M	22	66	-1.12	0.37	1.02	0.20	1.01	0.30
192 M	22	66	-1.12	0.37	0.78	-1.20	0.61	-0.40
110 M	21	66	-1.25	0.37	1.11	0.60	0.84	0.00
136 F	21	66	-1.25	0.37	1.03	0.20	1.29	0.60
147 F	21	66	-1.25	0.37	1.14	0.70	0.99	0.20
157 F	21	66	-1.25	0.37	1.17	0.90	0.84	0.00
199 F	21	66	-1.25	0.37	1.01	0.10	0.61	-0.40
173 M	20	66	-1.39	0.38	0.87	-0.60	0.58	-0.40
181 M	20	66	-1.39	0.38	0.75	-1.30	0.40	-0.80

103 M	19	66	-1.54	0.38	0.90	-0.40	0.73	-0.10
115 M	19	66	-1.54	0.38	0.97	-0.10	0.66	-0.20
127 F	19	66	-1.54	0.38	1.01	0.10	1.24	0.50
135 F	19	66	-1.54	0.38	0.67	-1.70	0.36	-0.80
148 M	19	66	-1.54	0.38	0.77	-1.10	0.44	-0.60
132 M	18	66	-1.68	0.39	1.02	0.20	0.87	0.10
141 F	18	66	-1.68	0.39	1.12	0.60	1.74	1.00
153 F	18	66	-1.68	0.39	0.71	-1.40	0.38	-0.70
118 M	17	66	-1.84	0.40	1.22	1.00	0.90	0.20
137 M	17	66	-1.84	0.40	0.88	-0.50	0.45	-0.60
187 F	17	66	-1.84	0.40	0.73	-1.30	0.66	-0.20
165 F	16	66	-2.00	0.41	0.63	-1.70	0.36	-0.80
171 F	16	66	-2.00	0.41	1.04	0.30	0.68	-0.20
182 M	16	66	-2.00	0.41	1.17	0.80	0.80	0.00
133 F	15	66	-2.17	0.42	1.03	0.20	0.86	0.10
139 F	15	66	-2.17	0.42	1.37	1.40	1.67	1.00
158 F	15	66	-2.17	0.42	0.74	-1.10	0.40	-0.70
163 M	15	66	-2.17	0.42	0.94	-0.20	0.84	0.10
152 F	14	66	-2.35	0.43	1.05	0.30	1.59	0.90
160 F	14	66	-2.35	0.43	0.96	-0.10	0.47	-0.50
190 F	14	66	-2.35	0.43	0.79	-0.80	0.37	-0.70
128 M	13	66	-2.54	0.44	1.15	0.60	0.98	0.30
149 M	13	66	-2.54	0.44	0.79	-0.80	0.42	-0.60
169 F	13	66	-2.54	0.44	0.83	-0.60	0.44	-0.50
186 M	13	66	-2.54	0.44	1.05	0.30	0.92	0.20
189 M	13	66	-2.54	0.44	0.91	-0.20	0.37	-0.70
129 M	12	66	-2.74	0.46	0.67	-1.20	0.25	-1.00
159 F	12	66	-2.74	0.46	0.84	-0.50	0.51	-0.40
184 M	12	66	-2.74	0.46	0.82	-0.60	0.52	-0.40
188 M	11	66	-2.96	0.48	1.69	2.00	1.28	0.60
191 F	11	66	-2.96	0.48	0.74	-0.90	0.25	-0.90
119 M	10	66	-3.20	0.50	1.03	0.20	0.65	-0.10
142 M	10	66	-3.20	0.50	1.13	0.50	0.67	-0.10
143 M	10	66	-3.20	0.50	0.99	0.10	0.61	-0.10
168 F	10	66	-3.20	0.50	1.07	0.30	0.65	-0.10
185 F	10	66	-3.20	0.50	0.66	-1.10	0.25	-0.80
130 M	8	66	-3.75	0.55	1.45	1.20	1.22	0.60
161 F	6	66	-4.44	0.62	1.54	1.30	0.70	0.10
131 M	3	66	-5.86	0.76	1.65	1.40	5.05	2.20
Mean	22.40	66.00	-1.25	0.39	1.01	0.00	0.88	0.00
SD	10.00	0.00	1.45	0.07	0.24	1.10	0.68	0.70
r	0.91							

Table 4.19

Phase Three Draw-a-Woman Sub-Test Entry Order Person Statistics

Case #	Raw Score	Count	Ability Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
100 F	49	66	2.10	0.36	1.44	2.20	1.20	0.50
174 F	49	66	2.10	0.36	1.26	1.40	1.70	1.10
112 F	40	66	1.01	0.34	1.29	1.60	1.24	0.60
179 F	40	66	1.01	0.34	1.16	0.90	1.00	0.20
195 F	40	66	1.01	0.34	0.74	-1.50	0.51	-0.80
177 F	39	66	0.89	0.34	1.27	1.40	1.04	0.30
193 F	39	66	0.89	0.34	1.19	1.10	1.22	0.50
167 F	38	66	0.78	0.34	1.00	0.10	0.70	-0.40
123 F	36	66	0.55	0.34	0.97	-0.10	0.76	-0.20
166 F	36	66	0.55	0.34	1.06	0.40	0.92	0.10
150 M	35	66	0.43	0.34	0.69	-1.90	0.44	-1.00
196 F	35	66	0.43	0.34	1.44	2.30	1.88	1.40
116 F	34	66	0.31	0.34	1.36	1.90	1.25	0.60
175 F	33	66	0.20	0.34	0.84	-0.90	0.61	-0.60
178 M	33	66	0.20	0.34	0.99	0.00	1.18	0.50
101 F	31	66	-0.04	0.34	0.90	-0.50	0.67	-0.40
172 M	31	66	-0.04	0.34	0.84	-0.90	0.59	-0.60
121 F	30	66	-0.16	0.35	0.69	-1.90	0.44	-1.00
125 F	30	66	-0.16	0.35	1.04	0.30	1.79	1.30
141 F	29	66	-0.27	0.35	0.90	-0.50	0.57	-0.70
157 F	27	66	-0.52	0.35	1.09	0.50	1.01	0.20
122 F	25	66	-0.77	0.36	0.63	-2.20	0.37	-1.30
199 F	25	66	-0.77	0.36	0.93	-0.30	0.55	-0.70
110 M	24	66	-0.90	0.36	1.13	0.70	0.84	-0.10
120 F	24	66	-0.90	0.36	0.90	-0.50	0.98	0.20
194 F	23	66	-1.03	0.36	1.25	1.30	1.56	1.00
140 M	22	66	-1.16	0.37	0.93	-0.30	0.71	-0.30
147 F	22	66	-1.16	0.37	1.39	1.80	2.79	2.20
202 M	22	66	-1.16	0.37	0.72	-1.50	0.42	-1.00
171 F	21	66	-1.30	0.37	0.85	-0.70	0.48	-0.80
180 M	21	66	-1.30	0.37	0.85	-0.70	0.94	0.10
159 F	20	66	-1.44	0.38	0.84	-0.80	0.48	-0.70
164 F	20	66	-1.44	0.38	1.32	1.50	1.29	0.60
129 M	19	66	-1.59	0.39	1.09	0.50	0.91	0.10
135 F	19	66	-1.59	0.39	1.02	0.10	0.73	-0.20
192 M	19	66	-1.59	0.39	1.03	0.20	0.73	-0.20
118 M	18	66	-1.74	0.39	0.93	-0.30	0.74	-0.10
139 F	18	66	-1.74	0.39	0.79	-0.90	0.48	-0.60
152 F	18	66	-1.74	0.39	0.62	-1.90	0.59	-0.40
176 M	18	66	-1.74	0.39	0.81	-0.80	0.45	-0.60
182 M	18	66	-1.74	0.39	0.73	-1.30	0.37	-0.80
189 M	18	66	-1.74	0.39	0.99	0.00	0.82	0.00
165 F	17	66	-1.90	0.40	0.95	-0.20	0.51	-0.50
115 M	16	66	-2.06	0.41	1.27	1.10	1.34	0.70

185 F	16	66	-2.06	0.41	0.76	-1.00	0.62	-0.30
169 F	15	66	-2.23	0.42	0.56	-2.00	0.28	-0.90
188 M	15	66	-2.23	0.42	1.25	1.00	1.46	0.80
191 F	15	66	-2.23	0.42	1.35	1.30	1.83	1.10
131 M	14	66	-2.42	0.43	0.88	-0.40	0.67	-0.10
143 M	14	66	-2.42	0.43	1.43	1.50	4.92	2.70
168 F	14	66	-2.42	0.43	0.96	-0.10	0.55	-0.30
160 F	13	66	-2.61	0.44	0.88	-0.40	0.52	-0.30
173 M	13	66	-2.61	0.44	0.82	-0.60	0.36	-0.60
184 M	12	66	-2.81	0.46	0.76	-0.80	0.51	-0.30
163 M	9	66	-3.50	0.50	1.00	0.10	0.49	-0.20
186 M	7	66	-4.05	0.55	0.67	-1.00	0.24	-0.60
Mean	24.6	66	-0.94	0.38	0.99	0.00	0.93	0.00
SD	10	0	1.35	0.04	0.23	1.10	0.72	0.80
r	0.95							

Table 4.19 reveals that three children's performances on the phase three DAW sub-test were detected as marginally misfitting. A faintly noisy response string was detected from case ID 196F (infit: Mn Sq. = 1.44, *t* = 2.3; outfit: Mn Sq. = 1.88, *t* = 1.4), and two slightly overfitting performances were reported by case IDs 122F (infit: Mn Sq. = .63, *t* -2.2; outfit: Mn Sq. = .37, *t* -1.3) and 169F (infit: Mn Sq. = .56, *t* = -2.0; outfit: Mn Sq. = .28, *t* = -.9). This is the first time that each of these children's performances have been detected as misfitting the Rasch model's expectations.

Examples of Children's Draw-a-Woman Drawings

The least successful, mean and most successful drawings of women from phases one, two and three are displayed in Figures 4.11, 4.12, and 4.13 respectively. Similar to the examples of the children's DAM drawings in Figures 4.4, 4.5 and 4.6, the DAW drawings show remarkable growth over the twelve month period.

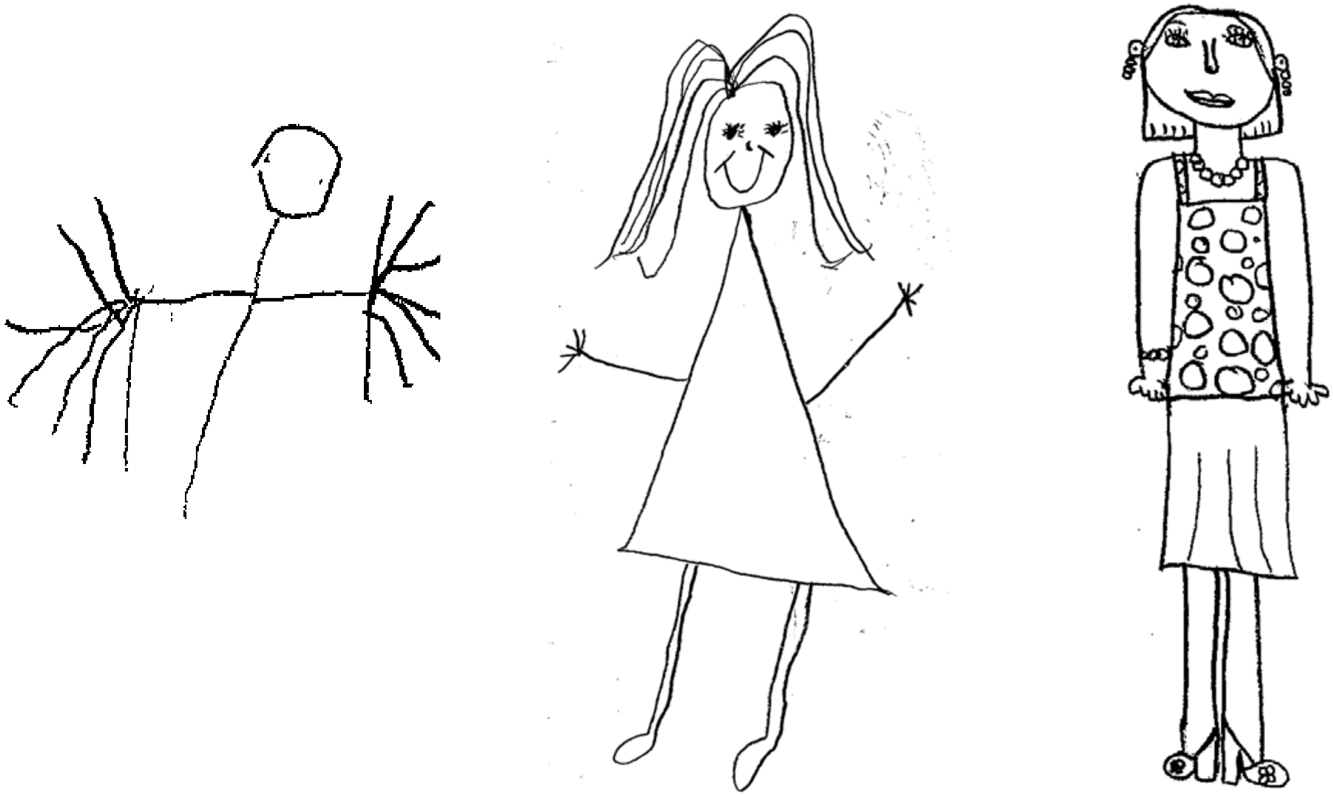
Figure 4.11

Phase One Draw-a-Woman Drawings: Examples of Most Successful, Mean, and Least Successful

Case ID: 130M Ability Estimate = -5.24; SE = .68	Case ID: 136F Ability Estimate = -1.26; SE = .36	Case ID: 196F Ability Estimate = 1.76; SE = .36

Figure 4.12

Phase Two Draw-a-Woman Drawings: Examples of Most Successful, Mean, and Least Successful



Case ID: 131M Ability Estimate = -5.86; SE = .76	Case ID: 136F Ability Estimate = 1.25; SE = .37	Case ID: 174F Ability Estimate = 2.21; SE = .37

Figure 4.13

Phase Three Draw-a-Woman Drawings: Examples of Most Successful, Mean, and Least Successful

Case ID: 186M	Case ID: 194F	Case ID:100F
Ability Estimate = -4.05; SE = .55	Ability Estimate = -1.03; SE = .36	Ability Estimate = 2.1; SE = 2.10

Draw-a-Woman Results from Analysis of Complete Data Set

For the same reasons that it was useful to stack the three phases of DAM data to further examine the sub-test's psychometrics, this was repeated for the three phases of DAW sub-test data. Furthermore, the results produced from this stacked DAW data analysis would be useful in the developmental graphs and common linking plots which are presented further on in this chapter.

Variable Map

Figure 4.14 presents the variable map produced by the Rasch analysis of the stacked DAW sub-test data for all three phases. As for the variable map for the stacked DAM sub-test data, the DAW variable map in Figure 4.14 spanned 14 logits from -8 to +6. The 71 DAW items, like the 73 DAM items, were well spread out along the logit scale. Resembling the phase one, two and three variable maps in Figures 4.8, 4.9, and 4.10, items 22 (hair IV), 27 (elbow joint shown), 50 (skirt "modeled" to indicate pleats or draping), 66 (directed lines and form: head outline), 67 (directed lines and form: breast), 68 (directed lines and form: hip contour), 70 (directed lines and form: calf of leg), and 71 (directed lines and form: facial features) were plotted towards the top of the scale indicating that they were amongst the most difficult items for these children to include in drawings of women. Items 1 (head present), 4 (eyes present), 13 (mouth present), 24 (arms present), 33 (legs present), and 55 (trunk present) were also consistently reported as some of the least difficult items in drawings of women.

Figure 4.14

Stacked Draw-a-Woman Sub-Test Data Person-Item Variable Map

```
          Persons MAP OF Items
              <more>|<rare>
   6                +
                    |
                    |T
                    |  70
   5                +
                    |
                    |  66
                    |  71
   4                +
                    |  22     68
                    |  27     49     50
                    |  47
   3                +  54     67     69
                    |S 45
                    |
                 .  |  30
   2             #  +  12     53     61     8
                 # T|  16     31     65
                 #  |  18     39
                ##  |  14     15     21     23     48
   1         #####  +  11     62
             .####  |  7
            .#####  |  52
            .##### S|  25     29     34     38     5      51
   0      .#######  +M 3
            ######  |  10     57
             .####  |  37
             .####  |  26     36     40     46     60
  -1        .#####  +  32     42     6      63
       .########## M|  44
           .######  |  2      64
         #########  |
  -2  ############  +  20     28     58     9
           .######  |  17
            ###### S|
            .#####  |S 35
  -3        ######  +  43
                ##  |
                 #  |  56
                .#  |  59
  -4               T+
                 .  |
                 .  |  19     41
                    |
  -5             .  +
                 #  |
                    |T 13
                    |  55
  -6                +  24
                    |  33
                    |
                    |
  -7                +
                    |
                    |
                    |  4
  -8                +  1
              <less>|<frequ>
EACH '#' IS 2.
```

Item Fit Statistics

Table 4.20 presents the item fit statistics generated from the Rasch analysis of the stacked DAW sub-test data. Six items were detected as underfitting. These erratic items were: 17 (chin and forehead shown), 28 (fingers present), 55 (trunk present), 57 (head-trunk proportion), 58 (head: proportion), and 64 (motor coordination). Most of these items yielded large infit and outfit *t* statistics suggesting that both low and high ability children responded unpredictably on these items. Review of Tables 4.13, 4.14, and 4.15 shows that items 17, 28, 57, 58, and 64 were also reported as underfitting in the separate analyses of phases one, two and three. The DAM equivalents of DAW items 57 (49), 58 (50), and 64 (63) were also reported as erratic in the stacked DAM sub-test data analysis.

Twelve items were detected as overfitting: items 3 (neck, two dimensions), 15 ("cosmetic lips"), 25 (shoulders), 26 (arms at side or engaged in activity or behind back), 29 (correct number of fingers shown), 34 (hip), 36 (feet II: proportion), 37 (feet III: detail), 40 (placement of feet appropriate to figure), 42 (attachment of arms and legs II), 43 (clothing indicated), 46 (neckline I), and 51 (no transparencies in the figure). Most of these items had large negative infit *t* statistics which are said to be "more sensitive to the pattern of responses to items targeted on the person, and vice-versa" (Linacre, 2002, p.878). Therefore, the children whose ability is well-targeted by the above items performed overly predictably on them – a little more variation was expected by the Rasch model. Items 3, 25, 26, 36, 40, 42, 46, and 51 were reported as overly predictable in the separate analyses of the DAM sub-test data. The DAM equivalents of items 25 (31), 26 (33), 34 (36), and 43 (55) were reported as overfitting in the stacked DAM sub-test data analysis.

Table 4.20

Stacked Draw-a-Woman Sub-Test Measure Order Item Fit Statistics

Item	Raw		Difficulty	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
70	1	246	5.26	1.01	1.01	0.30	0.40	-0.90
66	2	246	4.55	0.72	1.00	0.20	0.30	-1.10
71	3	246	4.13	0.59	0.93	0.00	0.20	-1.50
22	4	246	3.83	0.51	0.95	0.10	1.95	1.30
68	4	246	3.83	0.51	1.06	0.30	0.97	0.20
50	5	246	3.59	0.46	0.87	-0.20	0.19	-1.60
27	6	246	3.40	0.43	0.99	0.10	0.41	-0.90
49	6	246	3.40	0.43	0.86	-0.30	0.21	-1.60
47	7	246	3.23	0.40	0.94	-0.10	0.68	-0.30
67	8	246	3.08	0.37	1.06	0.30	1.19	0.50
69	8	246	3.08	0.37	0.96	0.00	0.49	-0.80
54	9	246	2.95	0.35	1.02	0.20	2.13	1.60
45	10	246	2.83	0.34	1.03	0.20	0.54	-0.70
30	17	246	2.20	0.27	0.96	-0.10	0.51	-1.10
61	20	246	2.00	0.25	0.97	-0.10	0.70	-0.60
8	21	246	1.94	0.25	1.06	0.40	1.10	0.40
12	22	246	1.88	0.24	0.98	0.00	1.05	0.30
53	22	246	1.88	0.24	0.84	-1.00	0.37	-1.90
16	23	246	1.82	0.24	0.86	-0.90	0.38	-1.90
65	23	246	1.82	0.24	1.02	0.20	1.10	0.40
31	25	246	1.71	0.23	1.10	0.70	1.33	0.90
18	29	246	1.51	0.22	1.21	1.50	1.76	1.90
39	30	246	1.46	0.22	0.88	-0.90	0.48	-1.80
15	33	246	1.33	0.21	0.85	-1.20	0.47	-2.00
21	33	246	1.33	0.21	0.81	-1.60	0.57	-1.50
23	33	246	1.33	0.21	1.00	0.10	1.22	0.80
14	34	246	1.28	0.21	0.87	-1.10	0.52	-1.70
48	35	246	1.24	0.20	0.99	0.00	0.74	-0.80
62	39	246	1.08	0.20	0.99	-0.10	0.77	-0.80
11	41	246	1.01	0.19	1.01	0.20	1.06	0.30
7	50	246	0.69	0.18	0.84	-1.60	0.65	-1.50
52	55	246	0.53	0.18	1.11	1.20	1.04	0.30
5	62	246	0.32	0.17	0.94	-0.60	1.38	1.70
38	63	246	0.29	0.17	0.89	-1.30	0.66	-1.80
51	64	246	0.27	0.17	0.73	-3.50	0.54	-2.60
29	66	246	0.21	0.17	1.34	3.60	1.55	2.40
34	68	246	0.15	0.17	1.67	6.70	2.44	5.40
25	69	246	0.13	0.17	0.72	-3.70	0.52	-3.00
3	72	246	0.04	0.16	0.78	-2.90	0.70	-1.80
57	82	246	-0.22	0.16	1.62	6.70	1.96	4.80
10	87	246	-0.34	0.16	1.02	0.30	0.95	-0.30
37	90	246	-0.42	0.16	0.73	-3.90	0.59	-3.30
60	99	246	-0.63	0.15	1.05	0.70	1.45	3.10
26	101	246	-0.68	0.15	0.70	-4.60	0.62	-3.40

40	105	246	-0.77	0.15	0.74	-3.90	0.65	-3.10
36	107	246	-0.82	0.15	0.79	-3.20	0.73	-2.40
46	107	246	-0.82	0.15	0.76	-3.70	0.65	-3.20
42	112	246	-0.94	0.15	0.92	-1.10	0.94	-0.50
32	113	246	-0.96	0.15	1.14	1.90	1.18	1.40
6	116	246	-1.03	0.15	1.02	0.40	1.16	1.30
63	117	246	-1.05	0.15	0.96	-0.60	0.92	-0.60
44	123	246	-1.19	0.15	0.90	-1.50	0.91	-0.80
2	140	246	-1.58	0.15	1.09	1.40	1.06	0.50
64	142	246	-1.62	0.15	1.19	2.70	1.39	2.80
20	154	246	-1.90	0.15	0.93	-1.10	0.82	-1.40
9	161	246	-2.07	0.16	1.20	2.70	1.31	2.00
28	162	246	-2.09	0.16	1.19	2.60	1.56	3.20
58	162	246	-2.09	0.16	1.61	7.50	2.65	7.70
17	166	246	-2.19	0.16	1.23	3.10	1.58	3.20
35	187	246	-2.75	0.17	0.83	-2.10	0.71	-1.40
43	193	246	-2.93	0.18	0.79	-2.40	0.55	-2.30
56	210	246	-3.52	0.20	0.89	-0.90	1.06	0.30
59	216	246	-3.78	0.22	1.12	0.90	1.59	1.50
19	228	246	-4.46	0.27	0.92	-0.40	0.51	-1.10
41	229	246	-4.53	0.27	0.89	-0.50	0.91	0.00
13	239	246	-5.59	0.40	1.00	0.10	0.65	-0.40
55	240	246	-5.77	0.43	0.96	0.00	2.76	2.10
24	241	246	-5.97	0.47	0.93	-0.10	0.88	0.00
33	242	246	-6.21	0.52	0.92	0.00	0.28	-1.30
4	245	246	-7.67	1.01	1.01	0.30	0.25	-1.30
1	246	246	-8.90	1.83	Minimum Estimated Measure			
Mean	88.5	246	-0.13	0.29	0.99	-0.10	0.95	-0.10
SD	77.6	0	2.95	0.25	0.19	2.20	0.57	2.10
r	0.99							

Self-Portrait-Man Sub-Test Results

The SPM sub-test comprised the drawing of a self-portrait by the male children in the sample (i.e., the female children completed a SPW drawing and the results are described in the following section). The SPM drawings were scored using the DAM 71 point scoring guide.

Identical to the sub-tests already discussed, the SPM sub-test data gathered from each of the three phases were Rasch analysed separately. The variable maps, summary statistics, and item and case fit statistics are described first, followed by examples of the male children's self-portrait drawings. This section closes with the presentation and description of the results from the stacked SPM sub-test data Rasch analysis.

Person-Item Variable Maps

Figures 4.15, 4.16 and 4.17 display the variable maps produced by the Rasch analysis of the phase one, two and three SPM sub-test data respectively. Whilst the variable maps produced by the analysis of the phase one and two data resulted in logits scales that spanned some 10 logits (both from -6 to +4), the logit scale produced by the phase three data was little smaller than this spanning 8 logits from -5 to +3. While the logit scales for phase one and two SPM data were similar to that produced for the DAM and DAW sub-test data (about 10 to 12 logits in length), the 8 point logit scale for phase three falls a little short in comparison.

Perhaps the most noticeable aspect of these SPM variable maps was the somewhat top-heavy item dispersion that was dissimilar to the variable maps produced by the DAM and DAW sub-test data thus far. Indeed, the total number of minimum extreme measure items in the SPM sub-test data was 17, 26 and 19 for phases one, two and three correspondingly. Items 12 (lips, two dimensions), 13 (both nose and lips in two dimensions), 21 (hair IV), 34 (elbow joint shown), 37 (hip II), 38 (knee joint shown), 58 (clothing IV), 59 (clothing V), 60 (profile I), 61 (profile II), 66 (directed lines and form: head outline), 69 (directed lines and form: facial features), 71 ('modeling' technique), 72 (arm movement), and 73 (leg movement) were not included in any self-portraits produced by males in all three phases. It is useful to note that items 59, 60, 61, 69, and 71 were also reported as extreme items in the DAM sub-tests.

Rasch analysis of the three phases of SPM data revealed that items 1 (head present), 4 (eyes present), 30 (arms present), 35 (legs present), 44 (attachment of arms and legs), and 46 (trunk present) were some of the most credited items. This aligns well with the results produced by the Rasch analysis of the DAM and DAW sub-test data collected across the three phases which indicated that these same items (or the equivalents of them in the case of the DAW sub-test) were reported as some of the most frequently included items in drawings of men and women.

Figures 4.15 and 4.16 show that case IDs 130M, 131M, 149M and 186M were reported as producing some of the least creditable self-portraits. This result is not unlike the results produced by the Rasch analysis of the DAM and DAW sub-test data where these cases were also reported as producing amongst the least creditable drawings of both men and women. Figure 4.17 reveals that the phase three SPM data analysis detected case ID 186M as the least successful. In this instance, case ID 131M seemed to have produced a slightly more creditable drawing, and case ID 130M was absent from this data collection phase.

Review of the top sections of the SPM variable maps reveals that case IDs 110M, 150M, 172M, 178M and 192M were amongst the most successful cases on the SPM sub-test. Case IDs 110M, 172M and 178M were similarly well located on the DAM sub-test variable maps (see Figures 4.1 to 4.3), and slightly less well located on the DAW sub-test variable maps (see Figures 4.8 to 4.10). Case ID 150M, on the other hand, has been in the top 5% performing children on each of the GHDT sub-tests thus far.

Figure 4.15

Phase One Self-Portrait-Man Person-Item Variable Map

```
                             Persons -MAP- Items
                                 <more>|<rare>
    4                                  +  12 13 15 21 32 34 37 38 58 59 60 61 66 69 71 72 73
                                       |
                                       |
                                       |  57 70
                                       |
                                       |
    3                                  +
                                       |
                                       |  23 5  56 62 68
                                       |
                                       |
                                       |S 20 26 27 41 52
    2                             150M  +
                                  100M  |  16 17
                                       |
                                      T|  65 7  8
                                       |  49
                                  110M  |
    1                             108M  +
                                  192M  |  29 3  31 67
                                       |  22 42
                                  172M  |  40 43
                                       |  10 19
                                  173M S|  36 6
    0           181M 197M 202M 206M   +M
                                       |  25 33
                             178M 180M  |  64
                                       |  45 48
     106M 111M 114M 117M 182M 203M    |  51
                                  119M  |  28
   -1                        118M 148M  +  2  63
                                  142M M|  54
                             103M 128M  |
                   132M 134M 140M 176M  |  14 53
                                       |  55 9
                   115M 143M 170M 198M  |  18
   -2                             137M  +
                                       |S 24 39
                                  129M  |
                        144M 184M 200M S|  47
                        163M 186M 188M  |
                                       |
   -3                   131M 149M 189M  +  50
                                       |
                                       |
                                       |
                                      T|  11
   -4                                  +
                                       |  44
                                       |
                                  130M  |T
                                       |
                                       |
   -5                                  +  46
                                       |
                                       |
                                       |
                                       |  30 4
                                       |
   -6                                  +  1  35
                                  <less>|<frequ>
```

Figure 4.16

Phase Two Self-Portrait-Man Person-Item Variable Map

```
               Persons -MAP- Items
                 <more>|<rare>
    4                  +  12 13 16 20 21 32 34 37 38 41 42 52 57 58 59 60 61 62 65 66 68 69 70 71 72 73
                       |T
                       |
                       |
                       |
                       |
    3                  +  15 56 7
                       |
                       |
                       |
                       |  23 44 49
                       |
    2                  +
                       |S 26 31 3
                       |
                       |  17 27 5 8
                      T|
                 172M  |
    1                  +
                 180M  |  33 45
                       |  25 36 40
     114M 170M 192M    |  29
                 150M  |  10 22
     110M 178M 202M S|    67
    0            181M  +M
                 173M  |  48 64
                 117M  |  19
                       |  51
                 103M  |
                       |  2 28 54 6 9
   -1       163M 176M M+
 115M 149M 182M 186M   |  63
            118M 132M  |  53
                 128M  |  39
                       |  14
 129 137 184 188 189   |S 55
   -2            148M  +  47
            119M 142M S|  24
                       |
                       |
                       |
                 143M  |  18 50
   -3                  +
                       |
                      T|
            130M 131M  |
                       |
                       |T
   -4                  +
                       |
                       |  11
                       |
                       |
                       |
   -5                  +
                       |  4
                       |
                       |
                       |
                       |
   -6                  +  1 30 35 44 46
                 <less>|<frequ>
```

Figure 4.17

Phase Three Self-Portrait-Man Person-Item Variable Map

```
        Persons -MAP- Items
             <more>|<rare>
    3              +  12 13 15 21 34 37 38 42 58 59 60 61 62 66 69 7  71 72 73
                   |
                   |  17 52 57 68 70
                   |
                   |
                   |
                   |
    2        192M  +
                   |S 26 3  32 41 5  56
                  T|
                   |
                   |
                   |  16 20 23 31 8
             150M  |
    1              +
                   |  45 49 65
                   |
        172M 178M  |
             202M  |  22 25 27 29 43
             180M S|
                   |  28
    0              +M
                   |  2  33 36 67
                   |
                   |  10 40 48
             140M  |
             176M  |
                   |
   -1        110M  +  24 51 54
                  M|
                   |
        118M 163M  |
        182M 189M  |  6  9
                   |
   115M 129M 173M  |S 19 53 63 64
   -2              +
             188M  |  55
                   |
             131M S|  18 39
                   |
             143M  |
                   |
   -3        184M  +  14 50
                   |
             186M  |
                   |
                   |
                   |T
                  T|
   -4              +  47
                   |
                   |
                   |
                   |
                   |  11
                   |
   -5              +  1  30 35 4  44 46
             <less>|<frequ>
```

Summary Statistics

Tables 4.21, 4.22 and 4.23 present the summary statistics produced by the Rasch analysis of the three phases of the SPM sub-test data. The lower halves of these tables display the information pertinent to the measured items. Across the three phases the number of measurable items varied slightly with 54, 42 and 48 deemed measurable for phases one, two, and three correspondingly. The item mean raw score decreased over time with 15.3, 11.7 and 7.3 reported for phases one to three respectively. The DAM sub-test reported 65, 61 and 63 measurable items for phases one, two and three respectively, and the item mean raw score also decreased over time with 36.4, 28.9 and 20.0 reported for phases one, two and three correspondingly. The SPM sub-test uses the exact same scoring criteria as the DAM sub-test, and these results show that the male children's self-portraits could be said to be less creditable against this criteria than their drawings of men.

The upper halves of Tables 4.21, 4.22 and 4.23 present the summary statistics for the cases in each of the phases of the SPM sub-test. The 45 measurable children in phase one yielded a mean raw score of 18.4, the 33 measurable children in phase two generated a mean raw score of 14.9, and the 21 children measured in phase three reported a mean raw score of 16.7. These were measurably less than those yielded by the DAM sub-test data analysis which were 21.2 (phase one), 22.1 (phase two), and 22.5 (phase three).

The person mean measures of -1.14, -1.00, and -1.08 for phases one to three respectively indicated that the SPM sub-test could be considered slightly better targeted to the sample than the DAM sub-test which yielded person means of -1.30, -1.21, and -1.14 for phases one, two and three. However, the SPM person mean raw scores mentioned above could suggest that it was a little more difficult to gain credit when the children were drawing themselves as opposed to men. So whilst asking male children to draw themselves (and scoring these drawings with the DAM scoring guide) was considered more suited to the sample, it was more difficult for the boys to gain credit for these drawings than when they draw a man which was then scored according to the DAM scoring guide. The issue is discussed in relation to the research questions in Chapter Five.

Table 4.21

Phase One Self-Portrait-Man Sub-Test Summary Statistics

SUMMARY OF 45 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	18.4	54.0	-1.14	.42	1.01	.0	.87	-.1
S.D.	8.6	.0	1.36	.07	.22	1.0	.55	.7
MAX.	41.0	54.0	2.07	.69	1.44	2.2	3.40	2.0
MIN.	4.0	54.0	-4.45	.36	.62	-2.2	.34	-1.2

REAL RMSE	.45	ADJ.SD	1.29	SEPARATION	2.88	Person RELIABILITY	.89
MODEL RMSE	.42	ADJ.SD	1.29	SEPARATION	3.04	Person RELIABILITY	.90
S.E. OF Person MEAN = .21							

LACKING RESPONSES: 62 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .91

SUMMARY OF 54 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	15.3	45.0	.00	.52	.97	.1	.87	.0
S.D.	12.7	.0	2.24	.20	.22	.9	.63	1.0
MAX.	44.0	45.0	3.51	1.04	1.45	2.3	3.38	2.5
MIN.	1.0	45.0	-5.70	.35	.61	-2.7	.10	-1.9

REAL RMSE	.57	ADJ.SD	2.17	SEPARATION	3.80	Item RELIABILITY	.94
MODEL RMSE	.56	ADJ.SD	2.17	SEPARATION	3.90	Item RELIABILITY	.94
S.E. OF Item MEAN = .31							

MAXIMUM EXTREME SCORE: 17 Items
MINIMUM EXTREME SCORE: 2 Items
UMEAN=.000 USCALE=1.000

Table 4.22

Phase Two Self-Portrait-Man Sub-Test Summary Statistics

SUMMARY OF 33 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	14.9	42.0	-1.00	.44	1.00	.0	.94	.0
S.D.	6.6	.0	1.20	.06	.28	1.2	.57	.8
MAX.	28.0	42.0	1.19	.65	1.72	2.7	2.63	2.0
MIN.	4.0	42.0	-3.48	.39	.59	-2.4	.31	-1.0

REAL RMSE .47	ADJ.SD 1.10	SEPARATION 2.33	Person RELIABILITY .84
MODEL RMSE .45	ADJ.SD 1.11	SEPARATION 2.49	Person RELIABILITY .86
S.E. OF Person MEAN = .21			

LACKING RESPONSES: 74 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .86

SUMMARY OF 42 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	11.7	33.0	.00	.55	.99	.1	.94	.1
S.D.	8.8	.0	1.91	.19	.19	.8	.61	.9
MAX.	32.0	33.0	3.06	1.04	1.49	2.6	3.77	3.1
MIN.	1.0	33.0	-5.09	.40	.63	-1.7	.18	-1.5

REAL RMSE .59	ADJ.SD 1.81	SEPARATION 3.05	Item RELIABILITY .90
MODEL RMSE .58	ADJ.SD 1.82	SEPARATION 3.13	Item RELIABILITY .91
S.E. OF Item MEAN = .30			

MAXIMUM EXTREME SCORE: 26 Items
MINIMUM EXTREME SCORE: 5 Items
UMEAN=.000 USCALE=1.000

Table 4.23

Phase Three Self-Portrait-Man Sub-Test Summary Statistics

SUMMARY OF 21 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	16.7	48.0	-1.08	.42	.97	-.1	.96	.0
S.D.	8.8	.0	1.38	.05	.25	1.1	.69	1.0
MAX.	37.0	48.0	1.94	.56	1.57	2.8	3.18	2.6
MIN.	5.0	48.0	-3.29	.37	.54	-2.0	.24	-1.4

REAL RMSE .44	ADJ.SD 1.31	SEPARATION 2.98	Person RELIABILITY .90
MODEL RMSE .42	ADJ.SD 1.32	SEPARATION 3.10	Person RELIABILITY .91
S.E. OF Person MEAN = .31			

LACKING RESPONSES: 86 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .91

SUMMARY OF 48 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	7.3	21.0	.00	.68	.98	.0	.96	.1
S.D.	5.3	.0	1.86	.18	.30	.9	.87	.9
MAX.	20.0	21.0	2.76	1.08	1.79	1.8	5.46	2.8
MIN.	1.0	21.0	-4.74	.51	.54	-1.6	.15	-1.3

REAL RMSE .73	ADJ.SD 1.70	SEPARATION 2.32	Item RELIABILITY .84
MODEL RMSE .70	ADJ.SD 1.72	SEPARATION 2.46	Item RELIABILITY .86
S.E. OF Item MEAN = .27			

MAXIMUM EXTREME SCORE: 19 Items
MINIMUM EXTREME SCORE: 6 Items
UMEAN=.000 USCALE=1.000

Item Fit Statistics

Tables 4.24, 4.25 and 4.26 display the item statistics produced from the Rasch analysis of the SPM sub-test data from phases one, two and three correspondingly. Table 4.24, which displays the item statistics for phase one, reveals two underfitting items: item 48 (infit Mn Sq. = 1.42, t = 2.3; outfit Mn Sq. = 1.74, t = 2.2), and item 50 (infit Mn Sq. = 1.38, t = 1.6; outfit Mn Sq. = 2.99, t = 2.5). Item 62 reported a large outfit mean square (infit Mn Sq. = .94, outfit Mn Sq. = 3.38); however, the corresponding infit and outfit t statistics of .1 and 1.6 respectively did not identify it as misfitting. Items 48 (proportion: head I) and 50 (Proportion: face) were also detected as underfitting in the DAM sub-test data analysis.

Only one SPM sub-test item was revealed to be overfitting: item 55 (infit Mn q. = .65, t = -2.7; outfit Mn Sq. = .54, t = -1.9). This item was not reported as misfitting in any of the DAM sub-test analyses.

Review of the DAW sub-test item fit statistics in Tables 4.14 to 4.16 reveals that the equivalents of items 50 (DAW item 58) and 55 (DAW item 43) were detected as misfitting.

Table 4.24

Phase One Self-Portrait-Man Sub-Test Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
12	0	45	4.78	1.84	Maximum Estimated Measure			
13	0	45	4.78	1.84	Maximum Estimated Measure			
15	0	45	4.78	1.84	Maximum Estimated Measure			
21	0	45	4.78	1.84	Maximum Estimated Measure			
32	0	45	4.78	1.84	Maximum Estimated Measure			
34	0	45	4.78	1.84	Maximum Estimated Measure			
37	0	45	4.78	1.84	Maximum Estimated Measure			
38	0	45	4.78	1.84	Maximum Estimated Measure			
58	0	45	4.78	1.84	Maximum Estimated Measure			
59	0	45	4.78	1.84	Maximum Estimated Measure			
60	0	45	4.78	1.84	Maximum Estimated Measure			
61	0	45	4.78	1.84	Maximum Estimated Measure			
66	0	45	4.78	1.84	Maximum Estimated Measure			
69	0	45	4.78	1.84	Maximum Estimated Measure			
71	0	45	4.78	1.84	Maximum Estimated Measure			
72	0	45	4.78	1.84	Maximum Estimated Measure			
73	0	45	4.78	1.84	Maximum Estimated Measure			
57	1	45	3.51	1.04	0.76	0.00	0.11	-0.40
70	1	45	3.51	1.04	1.01	0.30	0.30	-0.10
5	2	45	2.73	0.76	0.62	-0.60	0.13	-0.60
23	2	45	2.73	0.76	1.00	0.20	0.31	-0.30
56	2	45	2.73	0.76	0.62	-0.60	0.13	-0.60
62	2	45	2.73	0.76	0.94	0.10	3.38	1.60
68	2	45	2.73	0.76	0.88	0.00	0.45	-0.10
20	3	45	2.24	0.64	0.63	-0.80	0.20	-0.80
26	3	45	2.24	0.64	0.61	-0.80	0.19	-0.80
27	3	45	2.24	0.64	0.92	0.00	0.40	-0.40
41	3	45	2.24	0.64	0.92	0.00	0.40	-0.40
52	3	45	2.24	0.64	0.99	0.10	0.73	0.10
16	4	45	1.87	0.57	0.99	0.10	0.94	0.20
17	4	45	1.87	0.57	1.34	0.90	0.86	0.20
7	5	45	1.57	0.53	0.66	-1.00	0.38	-0.70
8	5	45	1.57	0.53	0.84	-0.40	1.05	0.30
65	5	45	1.57	0.53	1.11	0.40	1.01	0.30
49	6	45	1.31	0.49	1.45	1.40	1.35	0.70
3	8	45	0.88	0.44	0.96	-0.10	0.80	-0.20
29	8	45	0.88	0.44	0.80	-0.70	0.49	-0.90
31	8	45	0.88	0.44	0.64	-1.50	0.38	-1.20
67	8	45	0.88	0.44	0.87	-0.40	0.51	-0.80
22	9	45	0.69	0.43	1.20	0.90	1.06	0.30
42	9	45	0.69	0.43	1.27	1.10	1.05	0.30
40	10	45	0.51	0.41	0.79	-0.90	0.57	-0.90
43	10	45	0.51	0.41	0.92	-0.30	0.66	-0.60
10	11	45	0.35	0.40	1.24	1.10	1.30	0.80

19	11	45	0.35	0.40	1.10	0.50	0.98	0.10
6	12	45	0.19	0.39	1.24	1.20	1.05	0.30
36	12	45	0.19	0.39	0.84	-0.80	0.68	-0.70
25	14	45	-0.10	0.37	1.11	0.60	1.39	1.10
33	15	45	-0.23	0.37	0.96	-0.10	0.84	-0.40
64	16	45	-0.37	0.36	0.72	-1.80	0.58	-1.50
45	17	45	-0.49	0.36	0.94	-0.30	0.96	0.00
48	17	45	-0.49	0.36	1.42	2.30	1.74	2.20
51	19	45	-0.75	0.35	1.11	0.80	1.36	1.30
28	20	45	-0.87	0.35	1.11	0.80	1.04	0.20
2	21	45	-0.99	0.35	1.32	2.00	1.35	1.30
63	21	45	-0.99	0.35	1.19	1.30	1.31	1.20
54	23	45	-1.23	0.35	0.78	-1.60	0.68	-1.30
14	25	45	-1.47	0.35	0.99	0.00	0.93	-0.20
53	25	45	-1.47	0.35	1.08	0.60	1.76	2.40
55	26	45	-1.59	0.35	0.65	-2.70	0.54	-1.90
9	27	45	-1.71	0.35	1.25	1.60	1.83	2.40
18	28	45	-1.84	0.35	0.88	-0.70	1.04	0.30
39	30	45	-2.09	0.36	0.89	-0.60	0.77	-0.60
24	31	45	-2.23	0.37	0.98	-0.10	0.93	-0.10
47	33	45	-2.50	0.38	1.04	0.30	0.87	-0.20
50	36	45	-2.98	0.41	1.38	1.60	2.99	2.50
11	40	45	-3.81	0.51	1.03	0.20	0.55	-0.30
44	41	45	-4.10	0.56	0.79	-0.40	0.38	-0.60
46	43	45	-4.93	0.76	1.06	0.30	0.46	-0.10
4	44	45	-5.70	1.04	1.05	0.40	0.46	0.10
30	44	45	-5.70	1.04	0.68	-0.10	0.10	-0.60
1	45	45	-6.96	1.84	Minimum Estimated Measure			
35	45	45	-6.96	1.84	Minimum Estimated Measure			
Mean	12.6	45	0.92	0.86	0.97	0.10	0.87	0.00
SD	13.8	0	3.09	0.60	0.22	0.90	0.63	1.00
r	0.94							

Table 4.25

Phase Two Self-Portrait-Man Sub-Test Measure Order Item Statistics

Item	Raw		Difficulty	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
12	0	33	4.31	1.84	Maximum Estimated Measure			
13	0	33	4.31	1.84	Maximum Estimated Measure			
16	0	33	4.31	1.84	Maximum Estimated Measure			
20	0	33	4.31	1.84	Maximum Estimated Measure			
21	0	33	4.31	1.84	Maximum Estimated Measure			
32	0	33	4.31	1.84	Maximum Estimated Measure			
34	0	33	4.31	1.84	Maximum Estimated Measure			
37	0	33	4.31	1.84	Maximum Estimated Measure			
38	0	33	4.31	1.84	Maximum Estimated Measure			
41	0	33	4.31	1.84	Maximum Estimated Measure			
42	0	33	4.31	1.84	Maximum Estimated Measure			
52	0	33	4.31	1.84	Maximum Estimated Measure			
57	0	33	4.31	1.84	Maximum Estimated Measure			
58	0	33	4.31	1.84	Maximum Estimated Measure			
59	0	33	4.31	1.84	Maximum Estimated Measure			
60	0	33	4.31	1.84	Maximum Estimated Measure			
61	0	33	4.31	1.84	Maximum Estimated Measure			
62	0	33	4.31	1.84	Maximum Estimated Measure			
65	0	33	4.31	1.84	Maximum Estimated Measure			
66	0	33	4.31	1.84	Maximum Estimated Measure			
68	0	33	4.31	1.84	Maximum Estimated Measure			
69	0	33	4.31	1.84	Maximum Estimated Measure			
70	0	33	4.31	1.84	Maximum Estimated Measure			
71	0	33	4.31	1.84	Maximum Estimated Measure			
72	0	33	4.31	1.84	Maximum Estimated Measure			
73	0	33	4.31	1.84	Maximum Estimated Measure			
7	1	33	3.06	1.03	0.97	0.30	0.40	0.00
15	1	33	3.06	1.03	1.06	0.40	0.84	0.40
56	1	33	3.06	1.03	0.99	0.30	0.47	0.10
23	2	33	2.30	0.75	0.98	0.20	0.47	-0.10
43	2	33	2.30	0.75	1.12	0.40	1.52	0.80
49	2	33	2.30	0.75	1.13	0.40	2.10	1.10
3	3	33	1.82	0.63	0.89	-0.10	0.52	-0.30
26	3	33	1.82	0.63	0.91	0.00	0.43	-0.40
31	3	33	1.82	0.63	0.93	0.00	0.65	-0.10
5	4	33	1.47	0.57	0.84	-0.30	0.65	-0.20
8	4	33	1.47	0.57	0.74	-0.70	0.37	-0.80
17	4	33	1.47	0.57	0.96	0.00	0.73	-0.10
27	4	33	1.47	0.57	1.03	0.20	0.61	-0.30
33	6	33	0.91	0.49	1.17	0.70	1.14	0.40
45	6	33	0.91	0.49	0.89	-0.30	0.59	-0.60
25	7	33	0.68	0.47	1.26	1.10	1.15	0.40
36	7	33	0.68	0.47	0.88	-0.40	0.88	-0.10
40	7	33	0.68	0.47	0.97	0.00	0.76	-0.30

29	8	33	0.47	0.45	0.72	-1.30	0.54	-1.10
10	9	33	0.27	0.44	0.87	-0.60	0.69	-0.70
22	9	33	0.27	0.44	1.17	0.80	1.09	0.30
67	10	33	0.09	0.43	0.70	-1.60	0.55	-1.30
48	11	33	-0.09	0.42	0.91	-0.40	0.89	-0.20
64	11	33	-0.09	0.42	1.00	0.10	0.96	0.00
19	12	33	-0.26	0.41	0.87	-0.70	0.88	-0.30
51	13	33	-0.43	0.41	1.16	0.90	1.66	2.10
6	15	33	-0.75	0.40	1.07	0.50	1.03	0.20
9	15	33	-0.75	0.40	1.03	0.20	0.95	-0.10
2	16	33	-0.91	0.40	1.28	1.60	1.43	1.70
28	16	33	-0.91	0.40	0.93	-0.40	0.95	-0.10
54	16	33	-0.91	0.40	0.73	-1.70	0.67	-1.50
63	18	33	-1.23	0.40	1.16	1.00	1.11	0.50
53	19	33	-1.39	0.40	0.81	-1.20	0.73	-1.10
39	20	33	-1.55	0.40	0.96	-0.20	0.98	0.00
14	21	33	-1.72	0.41	1.49	2.60	1.96	2.70
55	22	33	-1.89	0.42	0.98	0.00	0.98	0.00
47	23	33	-2.06	0.43	0.77	-1.20	0.60	-1.20
24	24	33	-2.25	0.44	1.12	0.60	0.89	-0.10
18	27	33	-2.89	0.50	1.27	1.00	1.61	1.10
50	27	33	-2.89	0.50	1.38	1.30	3.77	3.10
11	31	33	-4.31	0.76	0.63	-0.50	0.19	-0.60
4	32	33	-5.09	1.04	0.81	0.10	0.18	-0.20
1	33	33	-6.35	1.84	Minimum Estimated Measure			
30	33	33	-6.35	1.84	Minimum Estimated Measure			
35	33	33	-6.35	1.84	Minimum Estimated Measure			
44	33	33	-6.35	1.84	Minimum Estimated Measure			
46	33	33	-6.35	1.84	Minimum Estimated Measure			
Mean	9.00	33.00	1.10	1.10	0.99	0.10	0.94	0.10
SD	10.80	0.00	3.20	0.65	0.19	0.80	0.61	0.90
r	0.90							

Table 4.25 reveals that two items were identified as misfitting in phase two according to the Rasch model's expectations. Item 14 (infit: Mn Sq. = 1.49, t = 2.6; outfit: Mn Sq. = 1.96, t = 2.7) and item 50 (infit: Mn Sq. = 1.38, t = 1.3; outfit: Mn Sq. = 3.77, t = 3.1) were reported to be underfitting. Whilst item 14 was underfitting in both the infit and outfit statistics, item 50 presented with only misfitting outfit statistics indicating that a child, or children, unexpectedly received credit on this item. Interestingly, item 50 presented with a similar fit issue in the phase one SPM analysis

(item 50 infit: Mn Sq. = 1.38, *t* = 1.6; outfit: Mn Sq. = 2.99, *t* = 2.5) as it did in this phase two analysis. None of the items in the phase two SPM analysis presented as overfitting. Item 50 has continually been identified as a problematic item, and while this was the first time that item 14 was detected as misfitting in the SPM sub-test data, it was reported as underfitting in the phase one DAM item statistics (see Table 4.4).

Rasch analysis of the phase three SPM data also detected two misfitting items; however they were not the same as those reported in phase two. As displayed in Table 4.24, items 16 (infit Mn Sq. = 1.79; outfit Mn Sq. = 2.56) and 49 (infit Mn Sq. = 1.43, outfit Mn Sq. = 1.90) both presented with rather high mean square statistics, whilst the *t* statistics for both of these items was acceptable (item 16: infit *t* = 1.7, outfit *t* = 1.4; item 49: infit *t* = 1.2, outfit *t* = 1.2). Neither of these items was detected as misfitting in the phase one nor phase two SPM data analyses, although both were reported in the DAM sub-test data analyses.

Table 4.26

Phase Three Self-Portrait-Man Sub-Test Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
7	0	21	4.09	1.86	Maximum Estimated Measure			
12	0	21	4.09	1.86	Maximum Estimated Measure			
13	0	21	4.09	1.86	Maximum Estimated Measure			
15	0	21	4.09	1.86	Maximum Estimated Measure			
21	0	21	4.09	1.86	Maximum Estimated Measure			
34	0	21	4.09	1.86	Maximum Estimated Measure			
37	0	21	4.09	1.86	Maximum Estimated Measure			
38	0	21	4.09	1.86	Maximum Estimated Measure			
42	0	21	4.09	1.86	Maximum Estimated Measure			
58	0	21	4.09	1.86	Maximum Estimated Measure			
59	0	21	4.09	1.86	Maximum Estimated Measure			
60	0	21	4.09	1.86	Maximum Estimated Measure			
61	0	21	4.09	1.86	Maximum Estimated Measure			
62	0	21	4.09	1.86	Maximum Estimated Measure			
66	0	21	4.09	1.86	Maximum Estimated Measure			
69	0	21	4.09	1.86	Maximum Estimated Measure			
71	0	21	4.09	1.86	Maximum Estimated Measure			
72	0	21	4.09	1.86	Maximum Estimated Measure			
73	0	21	4.09	1.86	Maximum Estimated Measure			
17	1	21	2.76	1.08	1.28	0.60	1.65	0.90
52	1	21	2.76	1.08	0.95	0.20	0.28	0.20
57	1	21	2.76	1.08	0.95	0.20	0.28	0.20
68	1	21	2.76	1.08	0.64	-0.20	0.15	0.00
70	1	21	2.76	1.08	0.64	-0.20	0.15	0.00
3	2	21	1.90	0.82	0.98	0.10	0.39	-0.10
5	2	21	1.90	0.82	1.18	0.50	0.53	0.00
26	2	21	1.90	0.82	0.74	-0.40	0.31	-0.20
32	2	21	1.90	0.82	0.56	-0.80	0.21	-0.40
41	2	21	1.90	0.82	1.33	0.70	0.98	0.40
56	2	21	1.90	0.82	1.01	0.20	0.42	-0.10
8	3	21	1.32	0.71	0.80	-0.40	0.40	-0.40
16	3	21	1.32	0.71	1.79	1.70	2.56	1.40
20	3	21	1.32	0.71	1.02	0.20	1.45	0.70
23	3	21	1.32	0.71	0.77	-0.50	0.51	-0.20
31	3	21	1.32	0.71	0.59	-1.00	0.28	-0.70
45	4	21	0.87	0.64	0.55	-1.40	0.31	-1.00
49	4	21	0.87	0.64	1.43	1.20	1.90	1.20
65	4	21	0.87	0.64	1.38	1.10	1.55	0.90
22	5	21	0.48	0.60	0.90	-0.20	0.60	-0.50
25	5	21	0.48	0.60	0.89	-0.30	0.64	-0.40
27	5	21	0.48	0.60	0.54	-1.60	0.34	-1.20
29	5	21	0.48	0.60	0.54	-1.60	0.34	-1.20
43	5	21	0.48	0.60	0.88	-0.30	0.68	-0.40
28	6	21	0.14	0.57	0.59	-1.50	0.41	-1.30

2	7	21	-0.17	0.55	1.35	1.20	1.28	0.70
33	7	21	-0.17	0.55	1.07	0.30	1.00	0.20
36	7	21	-0.17	0.55	0.74	-0.90	0.66	-0.70
67	7	21	-0.17	0.55	0.59	-1.60	0.56	-1.00
10	8	21	-0.47	0.54	1.09	0.40	0.96	0.00
40	8	21	-0.47	0.54	1.03	0.20	0.91	-0.10
48	8	21	-0.47	0.54	1.29	1.10	1.48	1.20
24	10	21	-1.02	0.52	0.67	-1.50	0.56	-1.30
51	10	21	-1.02	0.52	1.00	0.10	1.09	0.40
54	10	21	-1.02	0.52	0.84	-0.70	0.68	-0.90
6	12	21	-1.54	0.51	0.93	-0.20	0.88	-0.20
9	12	21	-1.54	0.51	1.15	0.80	1.69	1.50
19	13	21	-1.81	0.52	1.37	1.60	1.41	0.90
53	13	21	-1.81	0.52	0.95	-0.20	1.51	1.10
63	13	21	-1.81	0.52	1.24	1.10	1.31	0.80
64	13	21	-1.81	0.52	0.93	-0.30	0.73	-0.50
55	14	21	-2.08	0.53	0.82	-0.80	0.69	-0.50
18	15	21	-2.36	0.54	1.26	1.10	1.69	1.10
39	15	21	-2.36	0.54	0.82	-0.70	0.68	-0.40
14	17	21	-3.02	0.61	1.27	0.90	2.28	1.40
50	17	21	-3.02	0.61	1.62	1.80	5.46	2.80
47	19	21	-3.94	0.78	1.00	0.20	0.56	0.10
11	20	21	-4.74	1.05	1.09	0.40	0.88	0.50
1	21	21	-6.02	1.85	Minimum Estimated Measure			
4	21	21	-6.02	1.85	Minimum Estimated Measure			
30	21	21	-6.02	1.85	Minimum Estimated Measure			
35	21	21	-6.02	1.85	Minimum Estimated Measure			
44	21	21	-6.02	1.85	Minimum Estimated Measure			
46	21	21	-6.02	1.85	Minimum Estimated Measure			
Mean	6.50	21.00	0.57	1.08	0.98	0.00	0.96	0.10
SD	6.90	0.00	3.05	0.58	0.30	0.90	0.87	0.90
r	0.84							

Clearly, there are some common misfitting items across the DAM and SPM sub-tests. Although, this could be somewhat expected given that both use the exact same scoring criteria. On the other hand, however, the tests require slightly different responses with the DAM calling for drawings of male adults and the SPM sub-test involving drawings of male children. Harris (1963) asserted that drawings of human figures – whether adults or children – involve the same latent construct of intellectual maturity. However, this important, and somewhat assumed, aspect of the GHDT has not yet been empirically established from a modern test theory perspective.

Case Fit Statistics

Table 4.27 reveals that three cases were detected in phase one of the SPM data collection as having performances considered as underfitting according to the Rasch model's expectations. The unpredictable performances were by: case 108M (infit: Mn SQ. = 1.44, t = 2.2; outfit: Mn Sq. = 2.36, t = 1.9); case 109M (infit: Mn Sq. = 1.40, t = 2.0; outfit: Mn Sq. = 1.92, t = 1.3); and case 129M (infit: Mn Sq. = 1.19, t = .7; outfit: Mn Sq. = 3.40, t = 2.0). Case 203M was reported as the only child to have a response string detected as somewhat Guttman-like (infit: Mn Sq. = .62, t = -2.2; outfit: Mn Sq. = .40, t = -1.2).

Case ID 108M was also reported as having an overly erratic response string in phase one DAM sub-test analysis (see Table 4.7). Conversely, there were no common children reported as misfitting across the phase one SPM and phase one DAW sub-tests.

Table 4.27

Phase One Self-Portrait-Man Measure Order Person Statistics

Case	Raw		Ability	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
150 M	41	54	2.07	0.39	1.43	2.10	1.18	0.50
109 M	39	54	1.78	0.38	1.40	2.00	1.92	1.30
110 M	34	54	1.10	0.36	0.95	-0.20	0.74	-0.30
108 M	33	54	0.97	0.36	1.44	2.20	2.36	1.90
192 M	32	54	0.84	0.36	0.83	-0.90	0.66	-0.40
172 M	29	54	0.46	0.36	0.73	-1.60	0.58	-0.60
173 M	27	54	0.20	0.36	0.94	-0.30	0.80	-0.10
206 M	26	54	0.07	0.36	1.08	0.50	1.19	0.50
181 M	25	54	-0.06	0.36	1.14	0.80	1.20	0.50
197 M	25	54	-0.06	0.36	1.18	1.00	0.99	0.20
202 M	25	54	-0.06	0.36	0.95	-0.20	1.06	0.30
178 M	23	54	-0.33	0.37	1.24	1.30	1.01	0.20
180 M	23	54	-0.33	0.37	0.83	-0.90	0.56	-0.70
111 M	21	54	-0.61	0.38	0.77	-1.20	0.50	-0.90
114 M	21	54	-0.61	0.38	1.22	1.10	1.04	0.30
106 M	20	54	-0.75	0.38	0.86	-0.70	0.56	-0.80
117 M	20	54	-0.75	0.38	1.25	1.30	1.34	0.80
182 M	20	54	-0.75	0.38	1.05	0.30	0.90	0.00
203 M	20	54	-0.75	0.38	0.62	-2.20	0.40	-1.20
119 M	19	54	-0.89	0.38	0.80	-1.00	0.53	-0.90
118 M	18	54	-1.04	0.39	1.12	0.60	0.84	-0.10
148 M	18	54	-1.04	0.39	0.71	-1.50	0.44	-1.10
142 M	17	54	-1.20	0.40	0.74	-1.30	0.47	-0.90
103 M	16	54	-1.36	0.40	0.78	-1.00	0.54	-0.70
128 M	16	54	-1.36	0.40	0.97	-0.10	0.63	-0.50
132 M	15	54	-1.52	0.41	1.02	0.20	0.88	0.00
134 M	15	54	-1.52	0.41	1.21	1.00	1.10	0.40
140 M	15	54	-1.52	0.41	1.05	0.30	0.92	0.10
176 M	15	54	-1.52	0.41	1.29	1.30	1.20	0.50
115 M	13	54	-1.87	0.43	0.82	-0.70	0.49	-0.60
143 M	13	54	-1.87	0.43	0.92	-0.30	0.47	-0.60
170 M	13	54	-1.87	0.43	0.88	-0.40	0.99	0.20
198 M	13	54	-1.87	0.43	0.63	-1.70	0.34	-0.90
137 M	12	54	-2.07	0.45	0.86	-0.50	0.43	-0.60
129 M	11	54	-2.27	0.46	1.19	0.70	3.40	2.00
144 M	10	54	-2.49	0.48	0.81	-0.60	0.39	-0.50
184 M	10	54	-2.49	0.48	1.06	0.30	0.60	-0.20
200 M	10	54	-2.49	0.48	1.12	0.50	0.75	0.00
163 M	9	54	-2.73	0.50	1.31	1.00	0.97	0.30
186 M	9	54	-2.73	0.50	1.25	0.90	1.02	0.40
188 M	9	54	-2.73	0.50	0.72	-0.90	0.41	-0.40
131 M	8	54	-3.00	0.53	1.07	0.30	0.64	0.00
149 M	8	54	-3.00	0.53	0.75	-0.70	0.35	-0.40
189 M	8	54	-3.00	0.53	1.30	0.90	0.60	0.00

130 M	4	54	-4.45	0.69	1.16	0.50	0.52	0.00
Mean	18.4	54	-1.14	0.42	1.01	0.00	0.87	-0.10
SD	8.6	0	1.36	0.07	0.22	1.00	0.55	0.70
r	0.89							

Table 4.28 displays the phase two SPM measure order person statistics. Four children were reported as having performances considered misfitting by the Rasch model's expectations. The case IDs were: 110M (infit: Mn Sq. = 1.59, t = 2.7; outfit Mn Sq. = 199, t = 1.7); 150M (infit: Mn Sq. = 1.51, t = 2.4; outfit: Mn Sq. = 2.08, t = 1.7); and 172M (infit: Mn Sq. = 1.27, t 1.3; outfit: Mn Sq. = 2.63, t = 2.0). The single performance detected as overfitting was case ID 180M (infit: Mn Sq. = .59, t = -2.4; outfit: Mn Sq. = .42, t = -1.0). None of these children were reported as having misfitting performances in the phase one SPM analysis. Case ID 150M's phase one DAM sub-test drawing performance was also detected as misfitting. Whereas case ID 180M's drawings of men and women were both detected as misfitting in the phase one DAM and phase two DAW sub-test analyses respectively.

Table 4.28

Phase Two Self-Portrait-Man Measure Order Person Statistics

Case #	Raw Score	Count	Ability Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
172 M	28	42	1.19	0.40	1.27	1.30	2.63	2.00
180 M	26	42	0.86	0.40	0.59	-2.40	0.42	-1.00
114 M	24	42	0.55	0.39	1.00	0.00	0.96	0.10
170 M	24	42	0.55	0.39	0.90	-0.50	0.70	-0.40
192 M	24	42	0.55	0.39	0.78	-1.10	0.64	-0.50
150 M	23	42	0.40	0.39	1.51	2.40	2.08	1.70
110 M	22	42	0.25	0.39	1.59	2.70	1.99	1.70
178 M	22	42	0.25	0.39	0.98	-0.10	0.87	-0.10
202 M	22	42	0.25	0.39	1.10	0.60	1.38	0.80
181 M	20	42	-0.06	0.39	0.78	-1.20	0.61	-0.80
173 M	19	42	-0.22	0.40	1.08	0.50	1.32	0.80
117 M	18	42	-0.37	0.40	1.02	0.20	0.95	0.00
103 M	16	42	-0.70	0.41	1.45	2.00	1.42	1.00
163 M	14	42	-1.03	0.42	0.99	0.00	0.88	-0.10
176 M	14	42	-1.03	0.42	1.41	1.80	1.24	0.60
115 M	13	42	-1.21	0.42	0.73	-1.30	0.56	-0.80
149 M	13	42	-1.21	0.42	0.65	-1.80	0.48	-1.00
182 M	13	42	-1.21	0.42	0.91	-0.40	0.59	-0.70
186 M	13	42	-1.21	0.42	0.73	-1.30	0.53	-0.90
118 M	12	42	-1.39	0.43	1.17	0.80	1.73	1.20
132 M	12	42	-1.39	0.43	0.84	-0.70	0.62	-0.50
128 M	11	42	-1.58	0.44	1.01	0.10	0.79	-0.10
129 M	10	42	-1.78	0.46	0.86	-0.50	0.57	-0.50
137 M	10	42	-1.78	0.46	0.81	-0.70	0.49	-0.60
184 M	10	42	-1.78	0.46	1.11	0.50	1.98	1.30
188 M	10	42	-1.78	0.46	0.94	-0.10	0.61	-0.40
189 M	10	42	-1.78	0.46	0.72	-1.20	0.42	-0.80
148 M	9	42	-2.00	0.47	0.96	-0.10	0.85	0.10
119 M	8	42	-2.23	0.49	0.90	-0.30	0.54	-0.30
142 M	8	42	-2.23	0.49	0.68	-1.20	0.36	-0.70
143 M	6	42	-2.77	0.55	0.68	-0.90	0.31	-0.50
130 M	4	42	-3.48	0.65	1.72	1.50	0.82	0.30
131 M	4	42	-3.48	0.65	1.14	0.50	0.74	0.20
Mean	14.90	42.00	-1.00	0.44	1.00	0.00	0.94	0.00
SD	6.60	0.00	1.20	0.06	0.28	1.20	0.57	0.80
r	0.84							

Table 4.29

Phase Three Self-Portrait-Man Measure Order Person Statistics

Case #	Raw Score	Count	Ability Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
192 M	37	48	1.94	0.40	0.97	-0.10	0.68	-0.20
150 M	32	48	1.19	0.38	1.57	2.80	3.18	2.60
172 M	27	48	0.50	0.37	1.10	0.60	1.57	1.20
178 M	27	48	0.50	0.37	1.12	0.70	1.86	1.70
202 M	26	48	0.37	0.37	0.92	-0.40	0.78	-0.40
180 M	25	48	0.23	0.37	0.75	-1.40	0.54	-1.20
140 M	19	48	-0.60	0.38	1.02	0.20	0.87	-0.20
176 M	18	48	-0.75	0.38	1.26	1.30	1.78	1.70
110 M	16	48	-1.05	0.39	0.93	-0.30	0.65	-0.70
118 M	14	48	-1.36	0.41	1.09	0.50	0.75	-0.40
163 M	14	48	-1.36	0.41	0.78	-1.00	0.53	-0.90
182 M	13	48	-1.53	0.41	0.61	-2.00	0.36	-1.40
189 M	13	48	-1.53	0.41	0.78	-1.00	0.74	-0.30
115 M	11	48	-1.89	0.43	1.40	1.60	1.76	1.20
129 M	11	48	-1.89	0.43	0.91	-0.30	1.14	0.40
173 M	11	48	-1.89	0.43	1.14	0.70	0.75	-0.20
188 M	10	48	-2.08	0.44	0.99	0.00	0.87	0.10
131 M	8	48	-2.50	0.48	0.54	-2.00	0.24	-0.90
143 M	7	48	-2.73	0.50	0.86	-0.40	0.37	-0.50
184 M	6	48	-3.00	0.53	0.91	-0.20	0.55	-0.10
186 M	5	48	-3.29	0.56	0.62	-1.10	0.26	-0.50
Mean	16.70	48.00	-1.08	0.42	0.97	-0.10	0.96	0.00
SD	8.80	0.00	1.38	0.05	0.25	1.10	0.69	1.00
r	0.90							

Table 4.29 presents the phase three SPM measure order person statistics. Two cases had performances which were reported as underfitting: 131M (infit: Mn Sq. = .54, t = -2.0; outfit: Mn Sq. = .24, t = -.9) and 182M (infit: Mn Sq. = .61, t = -2.0; outfit: Mn Sq. = .36, t = -1.4). Only one case was detected as having a slightly overfitting performance: 150M (infit: Mn Sq. = 1.57, t = 2.8; outfit: Mn Sq. = 3.18, t = 2.6). Case ID 150M's drawings were detected as overfitting on several occasions across the DAM, DAW and SPM sub-tests. Whilst this was the first time case ID 131M was reported as a

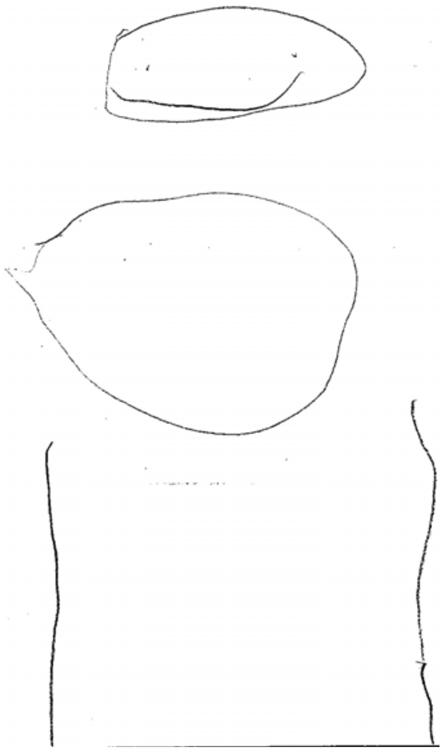
‘misfitter’ for SPM sub-test, his drawing of a woman in phase two was similarly detected.

Examples of the Children’s Self-Portrait-Man Drawings

Examples of the least successful, mean and most successful self-portraits produced by the male children from phases one, two and three are presented in Figures 4.18, 4.19 and 4.20. Similar to the other examples of the children’s drawings, these self-portraits produced by the male children show much advancement across the twelve month data collection period.

Figure 4.18

Phase One Self-Portrait-Man Drawings: Examples of Least Successful, Mean and Most Successful



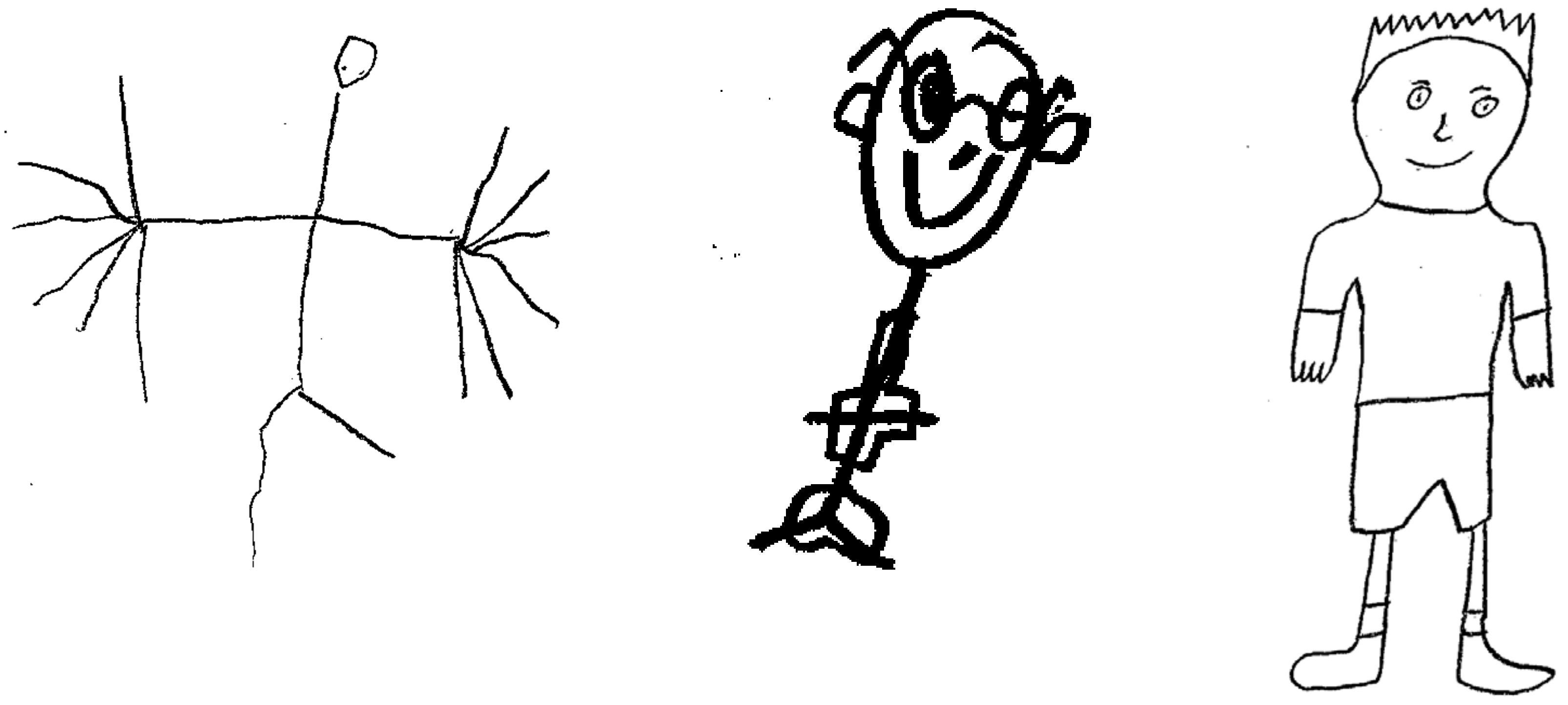
Case ID: 130M
Ability Estimate = -4.45; SE = .69

Case ID: 142M
Ability Estimate = -1.20; SE = .40

Case ID: 150M
Ability Estimate = 2.07; SE = .39

Figure 4.19

Phase Two Self-Portrait-Man Drawings: Examples of Least Successful, Mean and Most Successful



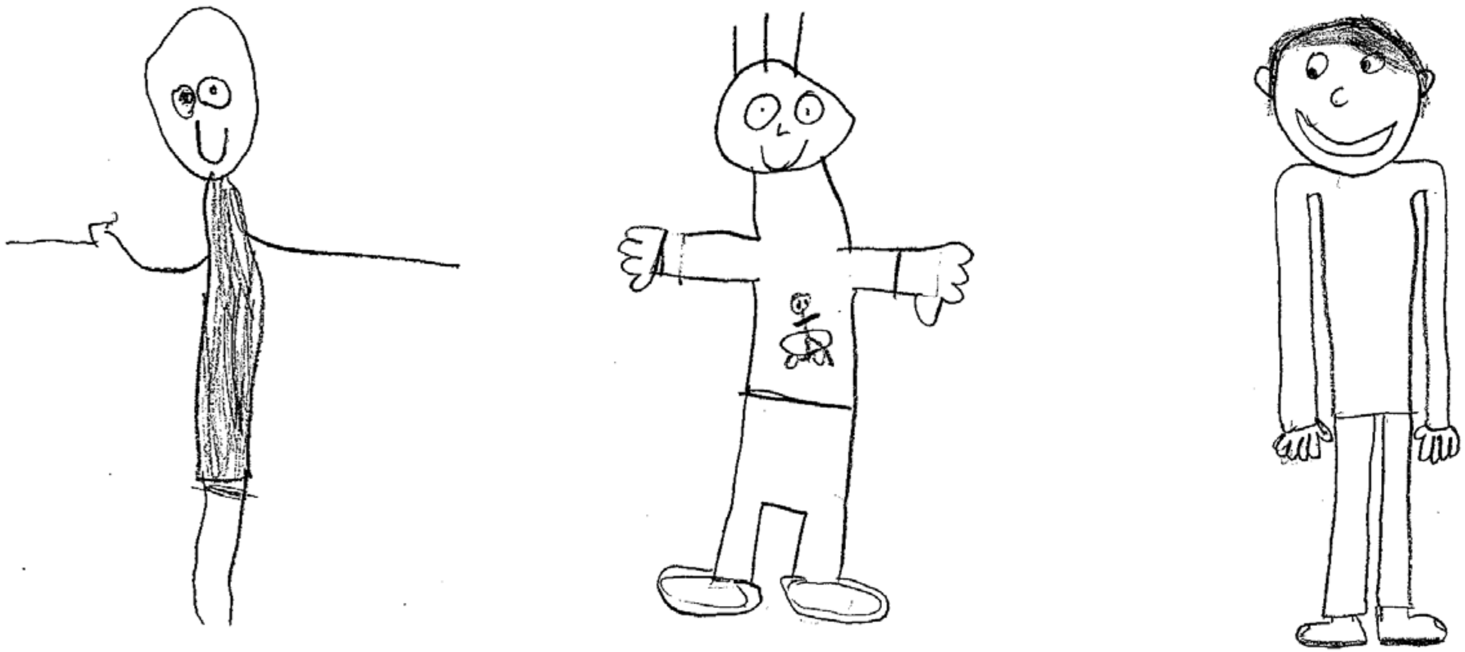
Case ID: 131M
Ability Estimate = -3.48; SE = .65

Case ID: 176M
Ability Estimate = -1.21; SE = .42

Case ID: 172M
Ability Estimate = 1.19; SE = .40

Figure 4.20

Phase Three Self-Portrait-Man Drawings: Examples of Least Successful, Mean and Most Successful



Case ID: 186M
Ability Estimate = -3.29; SE = .56

Case ID: 118M
Ability Estimate = -1.36; SE = .41

Case ID: 192M
Ability Estimate = 1.94; SE = .40

Self-Portrait-Man Results from Analysis of Complete Data Set

All three phases of the SPM sub-test data were combined and Rasch analysed to produce a more thorough indication of test unidimensionality. The results produced were also useful for the plotting of drawing development charts and common linking graphs. The variable map and item fit statistics generated by this stacked SPM data analysis are presented below.

Variable Map

Figure 4.21, the variable map produced by the Rasch analysis of the stacked SPM sub-test data, shows a logit scale spanning 12 logits from -7 to +5. This is two logits shorter than that generated by the Rasch analysis of the stacked DAM and DAW data which both spanned 14 logits from -8 to +6 (see Figures 4.7 and 4.14), and consistent with the DAM, DAW, and SPM variable maps from the analyses of the phases separately which have already been presented.

Review of the bottom of Figure 4.21 reveals that items 1 (head present), 4 (eyes present), 11 (mouth present), 35 (legs present), 44 (attachment of arms and legs I), and 46 (trunk present) were amongst the least difficult items to include in male children's self-portraits. These items have consistently been reported as amongst the least difficult to include in drawings of men, women and boys. In alignment with the variable maps produced by the Rasch analyses of each phase of the SPM data separately (see Figures 4.15, 4.16, and 4.17), this Rasch analysis of the stacked SPM data also reported a large number of items as being difficult to include in male children's self-portraits. Indeed, 13 items were considered maximum extreme measures by the Rasch model, which was in line with the other SPM sub-test results, but in contrast to the DAM sub-test results. The variable map produced by the stacked DAM sub-test data shows only two items, 59 (clothing V) and 61 (profile II), as being maximum extreme measures (see Figure 4.7).

Figure 4.21

Stacked Self-Portrait-Man Data Person-Item Variable Map

```
         Persons MAP OF Items
              <more>|<rare>
   5                +T 12  13  21  34  37  38  58  59  60  61  66  69  71
                    |
                    |
                    |
                    |  15
   4                +
                    |
                    |
                    |  32  57  62  70
                    |
   3                +  68
                    |
                    |  52
                    |S 41  56
                    |  20  7
   2             #  +  16  23
                    |  17  26  42  5   65
                 .  |
                   T|  27  49  8
                 .  |  3   31
   1             .  +
                 #  |  43
                .#  |
                ##  |  22  29
                    |  25  36  40  67
   0           ### S+M 10  33  45
               ###  |
                ##  |  19  48
                 #  |  6
              ####  |  28  51  64
  -1             #  +  2
                ## M|  54
               .##  |  63  9
          .#######  |  53
                 #  |
  -2         .####  +  14  24  39  55
                .#  |
                ##  |S 18
                ## S|
                    |  47
  -3           .##  +  50
                .#  |
                    |
                 #  |
                   T|
  -4                +
                    |  11
                    |
                    |
                 .  |
  -5                +T 44
                    |
                    |
                    |
                    |  4   46
  -6                +
                    |
                    |
                    |  30
                    |
  -7                +  1   35
              <less>|<frequ>
 EACH '#' IS 2.
```

Item Fit Statistics

Table 4.30 presents the item fit statistics generated from the Rasch analysis of the stacked SPM sub-test data. Five items were reported as underfitting, these were items: 2 (neck present), 9 (nose present), 50 (proportion: face), 53 (Proportion: legs), and 62 (full face). Most of these items had large outfit statistics indicating that some lower-ability children unexpectedly scored on these items, and some higher-ability children, perhaps, accidentally missed these items. None of these items were detected as underfitting in the analysis of the stacked DAM sub-test data. Although, items 9 (in phase one) and 50 (in phases two and three) were detected as overly erratic in the separate analyses of the DAM sub-test data.

Two items yielded fit statistics typical of an overfitting item. Items 54 (proportion: limbs in two dimensions) and 55 (clothing I) were reported as being too predictable according to the expectations of the Rasch model. Both of these items were detected as overfitting in the analysis of the stacked DAM sub-test data. Item 54 was detected as overly predictable in the Rasch output from the DAM sub-test in phases one and two.

Table 4.30

Stacked Self-Portrait-Man Measure Order Item Fit Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
12	0	99	5.43	1.83	Maximum Estimated Measure			
13	0	99	5.43	1.83	Maximum Estimated Measure			
21	0	99	5.43	1.83	Maximum Estimated Measure			
34	0	99	5.43	1.83	Maximum Estimated Measure			
37	0	99	5.43	1.83	Maximum Estimated Measure			
38	0	99	5.43	1.83	Maximum Estimated Measure			
58	0	99	5.43	1.83	Maximum Estimated Measure			
59	0	99	5.43	1.83	Maximum Estimated Measure			
60	0	99	5.43	1.83	Maximum Estimated Measure			
61	0	99	5.43	1.83	Maximum Estimated Measure			
66	0	99	5.43	1.83	Maximum Estimated Measure			
69	0	99	5.43	1.83	Maximum Estimated Measure			
71	0	99	5.43	1.83	Maximum Estimated Measure			
15	1	99	4.20	1.02	1.06	0.40	1.23	0.60
32	2	99	3.47	0.73	0.85	0.00	0.15	-0.90
57	2	99	3.47	0.73	0.85	0.00	0.15	-0.90
62	2	99	3.47	0.73	0.95	0.10	3.84	1.90
70	2	99	3.47	0.73	0.88	0.00	0.20	-0.80
68	3	99	3.02	0.61	0.81	-0.20	0.33	-0.60
52	4	99	2.70	0.53	0.94	0.00	0.62	-0.20
41	5	99	2.44	0.49	1.01	0.10	0.53	-0.40
56	5	99	2.44	0.49	0.82	-0.40	0.31	-0.90
7	6	99	2.22	0.45	0.86	-0.30	0.44	-0.70
20	6	99	2.22	0.45	0.77	-0.60	0.68	-0.20
16	7	99	2.03	0.42	1.14	0.60	1.43	0.80
23	7	99	2.03	0.42	0.92	-0.20	0.45	-0.80
5	8	99	1.86	0.40	0.91	-0.20	0.56	-0.60
26	8	99	1.86	0.40	0.77	-0.80	0.35	-1.20
17	9	99	1.71	0.38	1.28	1.10	0.92	0.10
42	9	99	1.71	0.38	1.21	0.90	1.24	0.60
65	9	99	1.71	0.38	1.10	0.50	1.15	0.40
8	12	99	1.33	0.34	0.83	-0.80	0.68	-0.60
27	12	99	1.33	0.34	0.87	-0.50	0.48	-1.20
49	12	99	1.33	0.34	1.33	1.40	1.56	1.10
3	13	99	1.21	0.33	0.95	-0.20	0.75	-0.40
31	14	99	1.11	0.32	0.73	-1.50	0.43	-1.50
43	17	99	0.82	0.30	0.96	-0.20	0.72	-0.60
29	21	99	0.49	0.28	0.74	-1.80	0.50	-1.70
22	23	99	0.34	0.27	1.13	0.90	1.00	0.10
40	25	99	0.19	0.26	0.90	-0.70	0.70	-1.10
67	25	99	0.19	0.26	0.76	-1.80	0.59	-1.60
25	26	99	0.13	0.26	1.10	0.70	1.31	1.10
36	26	99	0.13	0.26	0.82	-1.40	0.72	-1.10
45	27	99	0.06	0.26	0.88	-0.90	0.96	-0.10

10	28	99	-0.01	0.26	1.07	0.60	1.02	0.20
33	28	99	-0.01	0.26	1.04	0.40	0.95	-0.10
19	36	99	-0.50	0.24	1.09	0.80	1.06	0.40
48	36	99	-0.50	0.24	1.22	1.80	1.42	2.00
6	39	99	-0.67	0.24	1.11	1.10	1.07	0.40
64	40	99	-0.73	0.24	0.85	-1.40	0.76	-1.40
28	42	99	-0.84	0.24	0.99	-0.10	0.95	-0.20
51	42	99	-0.84	0.24	1.10	0.90	1.33	1.80
2	44	99	-0.95	0.23	1.29	2.70	1.52	2.80
54	49	99	-1.22	0.23	0.78	-2.40	0.68	-2.10
63	52	99	-1.38	0.23	1.19	1.90	1.20	1.20
9	54	99	-1.49	0.23	1.16	1.60	1.45	2.40
53	57	99	-1.65	0.23	0.96	-0.40	1.44	2.20
55	62	99	-1.93	0.24	0.79	-2.20	0.69	-1.70
14	63	99	-1.98	0.24	1.20	1.90	1.31	1.50
24	65	99	-2.10	0.24	1.00	0.00	0.89	-0.40
39	65	99	-2.10	0.24	0.90	-0.90	0.81	-0.90
18	70	99	-2.40	0.25	1.07	0.70	1.36	1.40
47	75	99	-2.72	0.26	0.94	-0.40	0.74	-0.80
50	80	99	-3.09	0.28	1.42	2.40	3.68	4.50
11	91	99	-4.25	0.39	0.95	-0.10	0.50	-0.80
44	95	99	-5.07	0.53	0.84	-0.20	0.33	-0.80
4	97	99	-5.84	0.73	1.00	0.20	0.30	-0.60
46	97	99	-5.84	0.73	1.03	0.30	0.39	-0.40
30	98	99	-6.57	1.02	0.75	0.00	0.06	-1.20
1	99	99	-7.81	1.83	Minimum Estimated Measure			
35	99	99	-7.81	1.83	Minimum Estimated Measure			
Mean	12.6	45	0.77	0.69	0.98	0.10	0.87	0.00
SD	13.8	0	3.37	0.61	0.16	0.90	0.63	1.00
r	0.94							

The following section describes the results from the Rasch analysis of the SPW sub-test data.

Self-Portrait-Woman Sub-Test Results

The SPW sub-test component of the GHDT was the third and final drawing made by female children in the sample. These drawings were scored using the 71 point DAW scoring guide. The results of the separate Rasch analysis of each of the three phases of SPW data are described in the sections below, with the variable maps being presented first, followed by the summary, item and person statistics, and some examples of the girls' self-portraits. This section concludes with the presentation and description of some key results from the Rasch analysis of the stacked SPW sub-test.

Person-Item Variable Maps

The variable maps displayed in Figures 4.22, 4.23 and 4.24 were produced by the Rasch analysis of the SPW sub-test data from phases one, two and three respectively. Each of these maps shows considerable similarity to those generated by the Rasch analyses of the DAW sub-test in particular, (i.e., both of these sub-tests used the same scoring guide), and the DAM and SPM sub-tests (which used DAM scoring guide) more generally.

The SPW logit scales spanned 11, 10 and 9 logits for phases one, two and three respectively. This aligns well with the logit scale ranges produced by the Rasch analyses of the three previous sub-tests (DAM sub-test: 12, 12 and 11; DAW sub-test: 13, 12, and 10; SPM sub-test: 10, 10, and 8 for phases one, two and three correspondingly). Resembling most other variable maps, the SPW variable map shows the item difficulty estimate locations were well-dispersed along the logit scale with a good number of items targeting the bottom, middle and top of the logit scales. In fact, it could be argued that these SPW items were, somewhat, more evenly spread along the scope of the scales than that presented in the other sub-tests' variable maps.

Review of the three variable maps shows that items 22 (hair IV), 66 (directed lines and form: head outline), 67 (directed line and form: hip contour), 68 (directed lines and form: hip contour), 69 (directed lines and form: arms taper), 70 (directed lines and form: calf of leg), and 71 (directed lines and form: facial features) were consistently reported as some of the most difficult to include in the female children's self-portraits. The DAW variable maps in Figures 4.8, 4.9, and 4.10 reveal that all of these items were reported as similarly difficult to include in drawings of women. The DAM/SPM equivalents of items 22 (DAM/SPM item 21), 66 (DAM/SPM item 66), and 71 (DAM/SPM item 69) were also reported as being amongst the most difficult items to include in drawings of men and male children's self-portraits. DAW/SPW items 67, 69, and 70 do not have an equivalent DAM/SPM item.

At the other end of the SPW logit scale is items 1 (head present), 4 (eyes present), 13 (mouth present), 19 (hair I), 24 (arms present), 33 (legs present), and 55 (trunk present). Not surprisingly, these items concerning generic human attributes were also amongst the least difficult to include in drawings of women (see Figures 4.8, 4.9, and 4.10), men (see Figures 4.1, 4.2, and 4.3), and self-portraits by male children (see Figures 4.15, 4.16 and 4.17).

Figures 4.22, 4.23, and 4.24 show that case IDs 100F, 112F, 174F, and 204F were consistent in producing self-portraits that were reported as being amongst the most successful in each of the three SPW data collection phases. Both the DAM and DAW variables maps displayed in Figures 4.1 to 4.3 and 4.8 to 4.10 respectively, show that these girls performed correspondingly well on those two sub-tests across each of the three phases.

The lower sections of the SPW sub-test variable maps show that case IDs 160F, 168F and 191F produced self-portraits which were reported as being amongst the least successful. The drawings of men and women produced by these girls yielded similar results on the DAM and DAW sub-test variable maps (see Figures 4.1 to 4.3 and 4.8 to 4.10).

Figure 4.22

Phase One Self-Portrait-Woman Person-Item Variable Map

```
                              Persons -MAP- Items
                                  <more>|<rare>
 5                                      +T 22    66    67    70    71
                                        |
                                        |
                                        |
                                        |  27    68    69
 4                                      +
                                        |
                                        |
                                        |  30    31    49    50
                                        |
 3                                      +  54    8
                                        |
                                        |S 12    61
                                        |  45    47    48
                                 100 F T|
 2                                      +
                                        |  21
                                 204 F  |  16    18    34
                                 112 F  |  23    29    39
                           113 F 146 F  |  11    53    65
 1                   174 F 177 F 196 F  +  14    15
                           107 F 124 F  |  62
                     122 F 193 F 195 F S|  38
               116 F 121 F 145 F 179 F  |  51    57
               125 F 156 F 166 F 205 F  |  7
 0                   102 F 175 F 201 F  +M 10    25    37    52
               101 F 120 F 154 F 155 F  |  36    40    46
                           105 F 167 F  |  3     5     6
                     123 F 127 F 151 F  |  44    60
                                       M|
-1                               136 F  +  26    32
                                        |  28    63    64
         104 F 147 F 153 F 164 F 194 F  |  2
                           133 F 138 F  |
                                 199 F  |  17    42
-2 126 F 135 F 141 F 158 F 161 F 187 F  +  20    58    9
         139 F 152 F 165 F 169 F 171 F  |  35
                                 183 F S|
                           157 F 159 F  |S
                     168 F 185 F 190 F  |
-3                                      +
                                        |  43    56
                                        |
                                 162 F  |
                                       T|  59
-4                                      +
                                        |  41
                           160 F 191 F  |
                                        |
                                        |  19
-5                                      +T
                                        |  24    55
                                        |
                                        |
                                        |
-6                                      +  1     13    33    4
                                  <less>|<frequ>
```

Figure 4.23

Phase Two Self-Portrait-Woman Person-Item Variable Map

```
                   Persons - MAP - Items
                       <more>|<rare>
 4                           +  68   70
                             |
                             |
                             |  22   66   67   69
                             |
                             |
 3                           +
                             |  47   50   54   71
                             |
                             |
                             |  31   45
                             |S
 2                           +  12   39
                            T|  21   30   49   61   62   65   8
                             |
                             |  11   16   27
                             |
                  174F 204F  |
 1                     112F  +
             101F 125F 175F  |  52
                  120F 156F  |  14   15   38   48   53
                       177F  |  18   29   57
                            S|  23
   122F 146F 166F 193F 205F  |  5    7
 0                124F 167F  +M
             116F 123F 126F  |  34
                             |  10   37   51
                             |  60
        113F 121F 155F 196F  |  25   3
                       154F  |  32   40   42
-1                     151F  +
                  179F 199F M|  36   63
                  136F 157F  |
                       141F  |  26   44   46
                       133F  |  6
                       153F  |  28
-2                     135F  +
                  147F 194F  |S 17   2
        139F 160F 171F 190F  |  58   64
                       187F  |  20
                  127F 165F S|  9
                             |  56
-3           152F 161F 185F  +
                             |
                  158F 191F  |
                       169F  |
                             |
                       159F  |  43
-4                          T+
                       168F  |  59
                             |  35
                             |T
                             |
                             |
-5                           +
                             |
                             |
                             |
                             |
                             |
-6                           +  1    13   19   24   33   4    41   55
                       <less>|<frequ>
```

Figure 4.24

Phase Three Self-Portrait-Woman Person-Item Variable Map

```
                                 Persons -MAP- Items
                                     <more>|<rare>
    4                                      +  45   66   68   69   70
                                           |
                                           |
                                           |
                                           |  22   49   50   61   62   67
                                           |
    3                                      +
                                           |
                                           |
                                           |  27   39   54   71
                                           |
                                           |S 16   47
    2                                      +
                                           |
                                    100 F T|  14   30   52
                                    174 F  |  15   23   31   48   53
                                           |
                                           |  12   18
    1                                      +  38   65   8
                                           |
                                           |
                              166 F 179 F  |  11   21   29   3
      112 F 123 F 167 F 175 F 177 F 193 F S|
                                           |  57
    0                   101 F 195 F 196 F  +M 51
                              121 F 125 F  |
                                    120 F  |  10   7
                                    157 F  |  34   42
                                    116 F  |  32   5
                                    139 F M|  25   37
   -1                         122 F 141 F  +  60
                                           |  36
                                    147 F  |  26   40   6    63
                                           |  2
                                    199 F  |  28   44   46
                  135 F 159 F 169 F 171 F  |  58
   -2                   165 F 185 F 194 F S+
                                    191 F  |S 64
                              152 F 164 F  |  9
                                           |
                                           |
                                    160 F  |  17
   -3                               168 F  +  20
                                          T|
                                           |  35
                                           |
                                           |  41   59
                                           |
   -4                                      +
                                           |  24   56
                                           |T
                                           |
                                           |
                                           |
   -5                                      +  1    13   19   33   4    43   55
                                     <less>|<frequ>
```

Summary Statistics

The lower sections of Table 4.31, 4.32, and 4.33 present information relevant to the SPW items. The number of measured items in each phase diminished slightly with 64, 63, and 60 measured in phases one, two, and three respectively. This was only slightly less than the numbers of measurable items in the DAW sub-test which analysed 67 items in phase one and 66 in phases two and three. The SPW item mean raw scores decreased a little from 24.2 in phase one, to 17.5 in phase two, to 13.7 in phase three. The latter result could be linked to the decreasing sample of females upon which the mean raw scores were calculated. Samples sizes decreased from 62 females in phase one, to 50 females in phase two, to just 35 in phase three. In any case, these item mean raw scores were less than those reported for DAW sub-test which were 36.0, 28.2 and 20.9 for phases one, two and three respectively. On the other hand, the standard deviations produced from the analyses of the phase one (SD = 2.52), two (SD = 2.24), and three SPW sub-test data (SD = 2.16) were relatively stable. So too were the item reliability indexes (phase one r = .96; phase two r = .94; phase three r = .92). Both of these sets of statistics from the SPW sub-test align well with that produced by the analysis of the DAW sub-test (phase one SD = 2.61, r = .97; phase two SD = 2.56, r = .96; phase three SD = 2.48, r = .95)

The upper sections of Tables 4.31, 4.32 and 4.33 display the summary statistics for the female children who produced a self-portrait. Unlike the item mean raw scores, which decreased over the three phases, the person mean raw scores for phases one (25), two (22.1), and three (23.6) showed a little more stability. These person mean raw scores were also strikingly similar to those produced by the Rasch analysis of the DAW sub-test data from phases one (22.6), two (22.4) and three (24.6) correspondingly. The SPW sub-test standard deviations were also relatively consistent across the phases (phase one SD = 1.50; phase two SD = 1.47, phase three SD = 1.21), and showed quite remarkable uniformity with the DAW sub-test standard deviations (phase one SD = 1.49, phase two SD = 1.45, phase three SD = 1.35). The SPW person reliability index also showed stability across the three phases with .92 for phases one and two, and .90 for phase three, and were highly comparable to the DAW sub-test person reliability indexes of .97, .96, and .95.

Table 4.31

Phase One Self-Portrait-Woman Summary Statistics

SUMMARY OF 62 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	25.0	64.0	-.83	.39	1.00	.0	.88	.0
S.D.	10.7	.0	1.50	.06	.22	1.0	.42	.6
MAX.	48.0	64.0	2.14	.60	1.56	1.9	2.09	1.5
MIN.	6.0	64.0	-4.34	.34	.55	-2.4	.23	-1.1

REAL RMSE .41 ADJ.SD 1.45 SEPARATION 3.50 Person RELIABILITY .92
MODEL RMSE .39 ADJ.SD 1.45 SEPARATION 3.70 Person RELIABILITY .93
S.E. OF Person MEAN = .19

LACKING RESPONSES: 45 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .93

SUMMARY OF 64 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	24.2	62.0	.00	.46	.99	.0	.88	-.1
S.D.	18.5	.0	2.52	.21	.26	1.5	.59	1.2
MAX.	61.0	62.0	4.15	1.03	1.83	4.7	3.34	4.4
MIN.	1.0	62.0	-5.99	.31	.61	-3.0	.10	-2.0

REAL RMSE .52 ADJ.SD 2.46 SEPARATION 4.74 Item RELIABILITY .96
MODEL RMSE .51 ADJ.SD 2.46 SEPARATION 4.86 Item RELIABILITY .96
S.E. OF Item MEAN = .32

MAXIMUM EXTREME SCORE: 5 Items
MINIMUM EXTREME SCORE: 2 Items
UMEAN=.000 USCALE=1.000

Table 4.32

Phase Two Self-Portrait-Woman Summary Statistics

```
SUMMARY OF 50 MEASURED Persons
+-----------------------------------------------------------------------------+
|          RAW                          MODEL         INFIT        OUTFIT     |
|         SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      22.1      63.0       -1.15     .39      1.01     .0    .89     .0 |
| S.D.      10.5        .0        1.47     .07       .26    1.1    .48     .7 |
| MAX.      40.0      63.0        1.11     .61      2.18    2.8   2.90    1.9 |
| MIN.       5.0      63.0       -4.20     .34       .62   -2.5    .26   -1.2 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .42  ADJ.SD    1.40  SEPARATION  3.37  Person RELIABILITY  .92 |
|MODEL RMSE    .39  ADJ.SD    1.41  SEPARATION  3.58  Person RELIABILITY  .93 |
| S.E. OF Person MEAN = .21                                                   |
+-----------------------------------------------------------------------------+
LACKING RESPONSES:     57 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .93

SUMMARY OF 63 MEASURED Items
+-----------------------------------------------------------------------------+
|          RAW                          MODEL         INFIT        OUTFIT     |
|         SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      17.5      50.0         .00     .49       .98     .0    .89    -.1 |
| S.D.      13.9        .0        2.24     .20       .28    1.5    .73    1.4 |
| MAX.      49.0      50.0        3.56    1.03      2.01    4.9   3.69    4.8 |
| MIN.       1.0      50.0       -5.98     .34       .49   -3.8    .19   -2.8 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .54  ADJ.SD    2.18  SEPARATION  4.01  Item   RELIABILITY  .94 |
|MODEL RMSE    .53  ADJ.SD    2.18  SEPARATION  4.11  Item   RELIABILITY  .94 |
| S.E. OF Item MEAN = .29                                                     |
+-----------------------------------------------------------------------------+
MAXIMUM EXTREME SCORE:      2 Items
MINIMUM EXTREME SCORE:      6 Items
UMEAN=.000 USCALE=1.000
```

Table 4.33

Phase Three Self-Portrait-Woman Summary Statistics

```
SUMMARY OF 35 MEASURED Persons
+-----------------------------------------------------------------------------+
|           RAW                          MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      23.6      60.0        -.79     .37      1.00     .0    .87    -.1 |
| S.D.       9.0        .0        1.21     .03       .21    1.0    .35     .6 |
| MAX.      43.0      60.0        1.71     .46      1.54    2.0   1.72     .9 |
| MIN.       9.0      60.0       -2.97     .35       .67   -1.9    .37   -1.2 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .39  ADJ.SD    1.14  SEPARATION  2.92  Person RELIABILITY  .90 |
|MODEL RMSE    .38  ADJ.SD    1.15  SEPARATION  3.06  Person RELIABILITY  .90 |
| S.E. OF Person MEAN = .21                                                   |
+-----------------------------------------------------------------------------+
LACKING RESPONSES:     72 Persons
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .90

SUMMARY OF 60 MEASURED Items
+-----------------------------------------------------------------------------+
|           RAW                          MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      13.7      35.0         .00     .56      1.00     .1    .87     .0 |
| S.D.      10.4        .0        2.16     .20       .21    1.0    .48    1.0 |
| MAX.      34.0      35.0        3.35    1.03      1.72    3.1   2.84    3.8 |
| MIN.       1.0      35.0       -4.93     .39       .70   -1.7    .17   -1.7 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .61  ADJ.SD    2.08  SEPARATION  3.42  Item   RELIABILITY  .92 |
|MODEL RMSE    .59  ADJ.SD    2.08  SEPARATION  3.51  Item   RELIABILITY  .92 |
| S.E. OF Item MEAN = .28                                                     |
+-----------------------------------------------------------------------------+
MAXIMUM EXTREME SCORE:      5 Items
MINIMUM EXTREME SCORE:      6 Items
UMEAN=.000 USCALE=1.000
```

Item Fit Statistics

Table 4.34 displays the measure order item statistics for phase one of the SPW sub-test. Six items were reported as having more variation than expected by the Rasch model, items: 7 (infit: Mn Sq. = .62, t = -3.0; outfit: Mn Sq. = .44, t = -1.8), 9 (infit: Mn Sq. = 1.60, t = 3.2; outfit: Mn Sq. = 1.73, t = 1.9), 17 (infit: Mn Sq. = 1.66, t = 3.6; outfit: Mn Sq. = 1.94, t = 2.6), 34 (infit: Mn Sq. = 1.37, t = 1.6; outfit: Mn Sq. = 2.27, t = 1.6), 57 (infit: Mn Sq. = 1.83, t = 4.7; outfit: Mn Sq. = 2.46, t = 2.9), 58 (infit: Mn Sq. = 1.81, t = 4.2; outfit: Mn Sq. = 3.34, t = 4.4) and 62 (infit: Mn Sq. = 1.51, t = 2.8; outfit: Mn Sq. = 1.57, t = 1.2).

Table 4.35, the phase two SPW item statistics, shows that items 9 (infit: Mn Sq. = 1.71, t = 3.2; outfit: Mn Sq. = 2.37, t = 2.5), 17 (infit: Mn Sq. = 1.43, t = 2.2; outfit: Mn Sq. = 2.05, t = 2.5), 28 (infit: Mn Sq. = 1.44, t = 2.3; outfit: Mn Sq. = 1.58, t = 1.8), 34 (infit: Mn Sq. = 2.01, t = 4.9; outfit Mn Sq. = 3.69, t = 4.8), 52 (infit: Mn Sq. = .99, t = .0; outfit: Mn Sq. = 2.62, t = 2.1), and 58 (infit: Mn Sq. = 1.88, t = 3.9; outfit: Mn Sq. = 3.52, t = 4.4) were detected as underfitting.

The phase three Rasch output in Table 4.36 reveals that items 34 (infit: Mn Sq. = 1.51, t = 2.6; outfit: Mn Sq. = 1.69, t = 2.5), 57 (infit: Mn Sq. = 1.57, t = 2.8; outfit: Mn Sq. = 2.84, t = 3.8) and 58 (infit: Mn Sq. = 1.72, t = 3.1; outfit: Mn Sq. = 1.86, t = 2.1) were similarly overly erratic.

Items 6, 9, 17, 34/36 (DAW/DAM), 57/49, and 58/50 were also described as overly erratic in both the DAM and DAW sub-tests; whereas the common misfitting items across the SPW and SPM sub-tests were items 9, 58/50 (DAW/SPM), and 57/49.

With regard to overfitting items, Table 4.34 reveals that items 25, 40, 44, 46 and 51 were reported in the phase one SPW sub-test data analysis. Table 4.35 reveals that items 25 and 51 were also detected in phase two, together with items 3, 26, and 42. Table 4.36 indicates that no items were reported as being too predictable in the phase three SPW analysis. Review of the DAW item fit statistics in Tables 4.14, 4.15, and 4.16 shows that all of the items reported as overfitting in the SPW sub-test analyses, bar item 44, were also detected in the DAW sub-test analyses. The equivalent of the SPW item 26, DAM item 33, was detected as overly predictable in phases one and two of the DAM sub-test analyses.

There were no common overfitting items amongst the SPW and SPM sub-test results.

Table 4.34

Phase One Self-Portrait-Woman Measure Order Item Statistics

Item	Raw		Difficulty	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
22	0	62	5.39	1.83	Maximum Estimated Measure			
66	0	62	5.39	1.83	Maximum Estimated Measure			
67	0	62	5.39	1.83	Maximum Estimated Measure			
70	0	62	5.39	1.83	Maximum Estimated Measure			
71	0	62	5.39	1.83	Maximum Estimated Measure			
27	1	62	4.15	1.02	1.03	0.30	0.56	0.10
68	1	62	4.15	1.02	0.93	0.20	0.22	-0.40
69	1	62	4.15	1.02	1.03	0.30	0.56	0.10
30	2	62	3.41	0.74	1.05	0.30	0.62	0.10
31	2	62	3.41	0.74	1.00	0.20	0.43	-0.20
49	2	62	3.41	0.74	0.78	-0.20	0.18	-0.60
50	2	62	3.41	0.74	0.99	0.20	0.51	-0.10
8	3	62	2.96	0.61	1.02	0.20	0.61	0.00
54	3	62	2.96	0.61	0.99	0.10	0.50	-0.20
12	4	62	2.63	0.54	1.05	0.30	0.59	-0.20
61	4	62	2.63	0.54	0.88	-0.20	0.60	-0.10
45	5	62	2.37	0.49	1.07	0.30	0.77	0.00
47	5	62	2.37	0.49	0.97	0.00	0.84	0.10
48	5	62	2.37	0.49	0.97	0.00	1.69	0.90
21	8	62	1.76	0.41	1.04	0.20	0.68	-0.30
16	9	62	1.60	0.39	0.84	-0.70	0.48	-0.70
18	9	62	1.60	0.39	1.36	1.50	1.69	1.10
34	9	62	1.60	0.39	1.37	1.60	2.27	1.60
29	10	62	1.45	0.38	1.09	0.50	1.05	0.30
39	10	62	1.45	0.38	1.05	0.30	0.73	-0.20
23	11	62	1.31	0.37	1.08	0.50	1.86	1.30
11	12	62	1.18	0.36	1.02	0.20	1.59	1.10
53	12	62	1.18	0.36	0.75	-1.50	0.46	-1.00
65	12	62	1.18	0.36	0.92	-0.40	0.73	-0.30
14	13	62	1.05	0.35	0.79	-1.20	0.51	-0.90
15	14	62	0.93	0.34	0.80	-1.30	0.52	-1.00
62	15	62	0.81	0.34	1.51	2.80	1.57	1.20
38	16	62	0.70	0.33	1.12	0.80	0.84	-0.20
51	19	62	0.38	0.32	0.61	-3.00	0.43	-1.80
57	19	62	0.38	0.32	1.83	4.70	2.46	2.90
7	20	62	0.27	0.32	0.62	-3.00	0.44	-1.80
10	22	62	0.07	0.31	0.90	-0.70	0.96	0.00
37	22	62	0.07	0.31	0.78	-1.60	0.64	-1.10
52	22	62	0.07	0.31	1.00	0.00	1.04	0.20
25	23	62	-0.02	0.31	0.66	-2.60	0.49	-1.90
36	24	62	-0.12	0.31	0.87	-0.90	0.68	-1.10
40	25	62	-0.22	0.31	0.67	-2.50	0.52	-1.90
46	25	62	-0.22	0.31	0.65	-2.70	0.51	-2.00
5	26	62	-0.31	0.31	1.12	0.90	1.19	0.80

3	27	62	-0.41	0.31	0.79	-1.50	0.67	-1.30
6	27	62	-0.41	0.31	0.94	-0.40	1.60	2.00
44	29	62	-0.60	0.31	0.68	-2.40	0.78	-0.90
60	29	62	-0.60	0.31	1.34	2.10	1.23	0.90
26	34	62	-1.08	0.31	0.74	-1.80	0.73	-1.20
32	34	62	-1.08	0.31	1.00	0.00	0.88	-0.40
28	35	62	-1.17	0.31	0.99	0.00	0.95	-0.10
63	35	62	-1.17	0.31	0.98	-0.10	0.90	-0.30
64	36	62	-1.27	0.31	1.28	1.70	1.37	1.40
2	37	62	-1.37	0.31	0.86	-0.90	0.74	-1.00
17	41	62	-1.77	0.32	1.66	3.60	1.94	2.60
42	42	62	-1.87	0.33	0.84	-1.00	0.69	-1.00
9	43	62	-1.98	0.33	1.60	3.20	1.73	1.90
20	44	62	-2.09	0.33	0.88	-0.70	0.69	-0.90
58	44	62	-2.09	0.33	1.81	4.20	3.34	4.40
35	45	62	-2.20	0.34	0.89	-0.70	0.77	-0.60
43	52	62	-3.12	0.39	0.85	-0.60	0.47	-0.90
56	52	62	-3.12	0.39	0.91	-0.30	1.15	0.40
59	56	62	-3.86	0.48	0.85	-0.40	0.38	-0.70
41	57	62	-4.11	0.51	0.79	-0.50	0.37	-0.50
19	59	62	-4.75	0.63	0.84	-0.20	0.39	-0.30
24	60	62	-5.23	0.75	0.91	0.10	0.51	0.00
55	60	62	-5.23	0.75	0.62	-0.50	0.10	-0.70
13	61	62	-5.99	1.03	0.95	0.20	0.18	-0.50
33	61	62	-5.99	1.03	1.10	0.40	0.81	0.30
1	62	62	-7.25	1.83	Minimum Estimated Measure			
4	62	62	-7.25	1.83	Minimum Estimated Measure			
Mean	23.5	62	0.18	0.60	0.99	0.00	0.88	-0.10
SD	19.7	0	3.03	0.46	0.26	1.50	0.59	1.20
r	.92							

Table 4.35

Phase Two Self-Portrait-Woman Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
68	0	50	4.80	1.83	Maximum Estimated Measure			
70	0	50	4.80	1.83	Maximum Estimated Measure			
22	1	50	3.56	1.02	0.97	0.97	0.30	0.35
66	1	50	3.56	1.02	0.94	0.94	0.20	0.28
67	1	50	3.56	1.02	1.03	1.03	0.30	0.67
69	1	50	3.56	1.02	0.94	0.94	0.20	0.28
47	2	50	2.82	0.74	1.00	1.00	0.20	0.54
50	2	50	2.82	0.74	1.03	1.03	0.30	0.90
54	2	50	2.82	0.74	1.00	1.00	0.20	0.54
71	2	50	2.82	0.74	0.87	0.87	0.00	0.27
31	3	50	2.37	0.62	0.96	0.96	0.10	0.46
45	3	50	2.37	0.62	0.85	0.85	-0.20	0.31
12	4	50	2.03	0.55	0.91	0.91	-0.10	0.71
39	4	50	2.03	0.55	0.88	0.88	-0.20	0.42
8	5	50	1.76	0.50	1.05	1.05	0.30	0.88
21	5	50	1.76	0.50	0.81	0.81	-0.50	0.38
30	5	50	1.76	0.50	0.91	0.91	-0.20	0.47
49	5	50	1.76	0.50	0.88	0.88	-0.30	0.44
61	5	50	1.76	0.50	0.91	0.91	-0.20	0.45
62	5	50	1.76	0.50	0.93	0.93	-0.10	0.46
65	5	50	1.76	0.50	0.99	0.99	0.10	0.91
11	6	50	1.52	0.47	0.99	0.99	0.10	1.45
16	6	50	1.52	0.47	0.95	0.95	-0.10	0.65
27	6	50	1.52	0.47	1.17	1.17	0.70	2.07
52	10	50	0.80	0.40	0.99	0.99	0.00	2.62
14	11	50	0.64	0.39	1.04	1.04	0.30	0.80
15	11	50	0.64	0.39	1.04	1.04	0.30	0.80
38	11	50	0.64	0.39	1.06	1.06	0.40	0.83
48	11	50	0.64	0.39	0.94	0.94	-0.30	0.59
53	11	50	0.64	0.39	0.80	0.80	-1.10	0.49
18	12	50	0.50	0.38	1.02	1.02	0.20	0.74
29	12	50	0.50	0.38	1.08	1.08	0.50	1.91
57	12	50	0.50	0.38	1.47	1.47	2.40	1.68
23	13	50	0.36	0.37	1.16	1.16	1.00	1.42
5	15	50	0.09	0.36	1.03	1.03	0.20	0.85
7	15	50	0.09	0.36	0.92	0.92	-0.40	0.69
34	17	50	-0.16	0.35	2.01	2.01	4.90	3.69
37	18	50	-0.28	0.35	0.70	0.70	-2.00	0.55
10	19	50	-0.40	0.35	1.09	1.09	0.60	1.00
51	19	50	-0.40	0.35	0.65	0.65	-2.40	0.52
60	20	50	-0.52	0.35	0.73	0.73	-1.80	0.58
3	21	50	-0.64	0.34	0.49	0.49	-3.80	0.39
25	21	50	-0.64	0.34	0.62	0.62	-2.60	0.49
42	22	50	-0.76	0.34	0.57	0.57	-3.10	0.45

32	23	50	-0.88	0.34	1.34	1.34	1.90	1.45
40	23	50	-0.88	0.34	0.90	0.90	-0.60	0.76
63	25	50	-1.11	0.34	1.09	1.09	0.60	1.02
36	26	50	-1.23	0.34	1.08	1.08	0.50	1.01
26	28	50	-1.47	0.35	0.65	0.65	-2.30	0.55
44	28	50	-1.47	0.35	0.70	0.70	-1.90	0.69
46	28	50	-1.47	0.35	0.71	0.71	-1.80	0.60
6	30	50	-1.71	0.35	0.96	0.96	-0.20	0.96
28	31	50	-1.84	0.35	1.44	1.44	2.30	1.58
2	34	50	-2.22	0.36	1.16	1.16	0.90	0.88
17	34	50	-2.22	0.36	1.43	1.43	2.20	2.05
58	35	50	-2.35	0.37	1.88	1.88	3.90	3.52
64	35	50	-2.35	0.37	1.14	1.14	0.80	1.47
20	36	50	-2.49	0.37	0.76	0.76	-1.40	0.69
9	37	50	-2.63	0.38	1.71	1.71	3.20	2.37
56	38	50	-2.78	0.39	0.62	0.62	-2.20	0.38
43	44	50	-3.86	0.48	0.90	0.90	-0.30	0.45
59	45	50	-4.11	0.51	0.74	0.74	-0.80	0.32
35	46	50	-4.40	0.56	0.66	0.66	-0.90	0.23
24	49	50	-5.98	1.03	0.90	0.90	0.20	0.19
41	49	50	-5.98	1.03	0.90	0.90	0.20	0.19
1	50	50	-7.23	1.83	Minimum Estimated Measure			
4	50	50	-7.23	1.83	Minimum Estimated Measure			
13	50	50	-7.23	1.83	Minimum Estimated Measure			
19	50	50	-7.23	1.83	Minimum Estimated Measure			
33	50	50	-7.23	1.83	Minimum Estimated Measure			
55	50	50	-7.23	1.83	Minimum Estimated Measure			
Mean	19.8	50	-0.48	0.64	0.98	0.00	0.89	-0.10
SD	16.3	0	3.05	0.47	0.28	1.50	0.73	1.40
r	0.94							

Table 4.36

Phase Three Self-Portrait-Woman Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
45	0	35	4.60	1.84	Maximum Estimated Measure			
66	0	35	4.60	1.84	Maximum Estimated Measure			
68	0	35	4.60	1.84	Maximum Estimated Measure			
69	0	35	4.60	1.84	Maximum Estimated Measure			
70	0	35	4.60	1.84	Maximum Estimated Measure			
22	1	35	3.35	1.03	0.86	0.10	0.22	-0.10
49	1	35	3.35	1.03	0.79	0.10	0.17	-0.20
50	1	35	3.35	1.03	1.04	0.30	0.62	0.30
61	1	35	3.35	1.03	1.04	0.30	0.62	0.30
62	1	35	3.35	1.03	1.04	0.30	0.62	0.30
67	1	35	3.35	1.03	1.04	0.30	0.62	0.30
27	2	35	2.58	0.75	1.13	0.40	1.07	0.50
39	2	35	2.58	0.75	0.87	0.00	0.52	-0.10
54	2	35	2.58	0.75	0.90	0.00	0.42	-0.20
71	2	35	2.58	0.75	0.83	-0.10	0.36	-0.30
16	3	35	2.10	0.63	0.84	-0.20	0.42	-0.40
47	3	35	2.10	0.63	0.86	-0.20	0.46	-0.40
14	4	35	1.75	0.56	0.88	-0.20	0.52	-0.50
30	4	35	1.75	0.56	1.14	0.50	0.85	0.10
52	4	35	1.75	0.56	0.99	0.10	0.85	0.10
15	5	35	1.45	0.52	1.00	0.10	1.41	0.80
23	5	35	1.45	0.52	1.20	0.70	0.83	0.00
31	5	35	1.45	0.52	1.10	0.40	0.71	-0.20
48	5	35	1.45	0.52	0.82	-0.50	0.66	-0.30
53	5	35	1.45	0.52	0.91	-0.20	0.52	-0.60
12	6	35	1.20	0.49	1.00	0.10	0.85	-0.10
18	6	35	1.20	0.49	1.02	0.20	2.18	1.70
8	7	35	0.98	0.46	1.23	1.00	1.57	1.10
38	7	35	0.98	0.46	1.12	0.60	1.02	0.20
65	7	35	0.98	0.46	1.18	0.80	1.58	1.10
3	9	35	0.58	0.43	0.76	-1.20	0.53	-1.20
11	9	35	0.58	0.43	0.90	-0.40	0.67	-0.80
21	9	35	0.58	0.43	0.76	-1.20	0.53	-1.20
29	9	35	0.58	0.43	1.31	1.50	1.49	1.20
57	11	35	0.23	0.41	1.57	2.80	2.84	3.80
51	12	35	0.07	0.40	0.79	-1.30	0.61	-1.30
10	14	35	-0.25	0.39	0.81	-1.10	0.81	-0.70
7	15	35	-0.41	0.39	0.73	-1.70	0.62	-1.70
34	16	35	-0.56	0.39	1.51	2.60	1.69	2.50
42	16	35	-0.56	0.39	0.78	-1.30	0.70	-1.30
5	17	35	-0.71	0.39	0.86	-0.70	0.84	-0.60
32	17	35	-0.71	0.39	0.92	-0.40	0.86	-0.60
25	18	35	-0.86	0.39	0.92	-0.40	0.88	-0.40
37	18	35	-0.86	0.39	1.01	0.10	0.95	-0.10

60	19	35	-1.02	0.39	1.04	0.30	0.96	-0.10
36	20	35	-1.17	0.39	1.01	0.10	0.96	-0.10
6	21	35	-1.33	0.40	0.72	-1.60	0.62	-1.60
26	21	35	-1.33	0.40	0.80	-1.10	0.77	-0.80
40	21	35	-1.33	0.40	0.92	-0.40	0.86	-0.50
63	21	35	-1.33	0.40	1.17	0.90	1.12	0.50
2	22	35	-1.49	0.40	0.86	-0.70	0.86	-0.40
28	23	35	-1.65	0.41	1.25	1.30	1.35	1.10
44	23	35	-1.65	0.41	0.80	-1.10	0.79	-0.60
46	23	35	-1.65	0.41	0.70	-1.70	0.66	-1.10
58	24	35	-1.82	0.41	1.72	3.10	1.86	2.10
64	26	35	-2.17	0.43	1.34	1.60	1.61	1.40
9	27	35	-2.37	0.45	1.12	0.60	1.09	0.40
17	29	35	-2.80	0.49	0.97	0.00	0.89	0.00
20	30	35	-3.05	0.52	1.10	0.40	0.75	-0.20
35	31	35	-3.34	0.56	0.86	-0.30	0.44	-0.60
41	32	35	-3.70	0.63	0.99	0.10	0.60	-0.10
59	32	35	-3.70	0.63	0.99	0.10	0.60	-0.10
24	33	35	-4.17	0.75	1.03	0.30	0.72	0.20
56	33	35	-4.17	0.75	1.00	0.20	0.55	0.00
43	34	35	-4.93	1.03	0.86	0.10	0.23	-0.20
1	35	35	-6.18	1.83	Minimum Estimated Measure			
4	35	35	-6.18	1.83	Minimum Estimated Measure			
13	35	35	-6.18	1.83	Minimum Estimated Measure			
19	35	35	-6.18	1.83	Minimum Estimated Measure			
33	35	35	-6.18	1.83	Minimum Estimated Measure			
55	35	35	-6.18	1.83	Minimum Estimated Measure			
Mean	14.6	35	-0.20	0.76	1.00	0.10	0.87	0.00
SD	11.9	0	2.94	0.50	0.21	1.00	0.48	1.00
r	0.92							

Case Fit Statistics

The case fit statistics for phases one, two and three of the SPW sub-test are displayed in Tables 4.37, 4.38 and 4.39 respectively. Table 4.37 indicates that only one child presented as slightly underfitting, and another as slightly overfitting in phase one of the SPW sub-test according to the Rasch model. Case ID 187F (infit: Mn Sq. = .55, t = -2.2; outfit: Mn Sq. = .25, t = -1.1) was identified as having a response string with less variation than expected by the Rasch model; whereas case ID 193F (infit: Mn Sq. = 1.35, t = 1.9; outfit: Mn Sq. = 2.09, t = 1.5) was identified as having more variation. Case ID 193F was reported as having an unpredictable response string on the DAM sub-test as well.

Table 4.38 presents the phase two SPW measure order person statistics. Three children had performances which were detected as slightly unpredictable, these were: case ID 112F (infit: Mn Sq. = 1.55, t = 2.8; outfit: Mn Sq. = 2.23, t = 1.9), 125F (infit: Mn Sq. = 1.50, t = 2.6; outfit: Mn Sq. = 1.44, t = 1.0), and 159F (infit: Mn Sq. = 2.18, t = 2.5; outfit: Mn Sq. = 2.90, t = 1.5). On the other hand, Case ID 123F (infit: Mn Sq. = .62, t = -2.5; outfit: Mn Sq. = .44, t = -1.2) presented as having a response string which was considered too predictable by the Rasch model. None of these children presented as having performances detected as misfitting on the phase one SPW sub-test. However, case ID 112F was detected as misfitting in both the DAM and DAW sub-test analyses for phase two.

Table 4.37

Phase One Self-Portrait-Woman sub-Test Measure Order Person Statistics

Case	Raw		Ability	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
100 F	48	64	2.14	0.37	0.96	-0.10	0.75	-0.20
204 F	44	64	1.61	0.36	1.17	0.90	1.45	0.80
112 F	43	64	1.48	0.35	0.99	0.00	1.17	0.50
113 F	41	64	1.24	0.35	0.79	-1.20	0.60	-0.60
146 F	40	64	1.12	0.35	0.86	-0.80	0.80	-0.10
174 F	39	64	1.00	0.34	0.92	-0.40	0.81	-0.10
177 F	39	64	1.00	0.34	1.04	0.30	0.84	-0.10
196 F	39	64	1.00	0.34	1.06	0.40	1.12	0.40
124 F	38	64	0.88	0.34	0.63	-2.40	0.40	-1.10
107 F	37	64	0.76	0.34	0.75	-1.50	0.51	-0.80
193 F	36	64	0.64	0.34	1.35	1.90	2.09	1.50
195 F	36	64	0.64	0.34	1.24	1.40	1.25	0.60
122 F	35	64	0.53	0.34	1.06	0.40	1.41	0.80
116 F	34	64	0.41	0.34	0.85	-0.90	0.68	-0.40
121 F	34	64	0.41	0.34	0.97	-0.10	0.78	-0.20
145 F	34	64	0.41	0.34	1.24	1.40	1.09	0.30
179 F	34	64	0.41	0.34	1.03	0.20	0.76	-0.20
125 F	33	64	0.30	0.34	0.80	-1.20	0.56	-0.60
156 F	33	64	0.30	0.34	1.31	1.70	1.33	0.70
166 F	32	64	0.18	0.34	1.03	0.20	1.28	0.60
205 F	32	64	0.18	0.34	1.06	0.40	0.77	-0.20
102 F	31	64	0.06	0.34	1.13	0.80	0.89	0.00
175 F	30	64	-0.05	0.34	0.83	-1.00	0.93	0.10
201 F	30	64	-0.05	0.34	1.15	0.90	0.98	0.20
120 F	29	64	-0.17	0.34	0.85	-0.90	0.58	-0.60
155 F	29	64	-0.17	0.34	0.96	-0.20	0.68	-0.40
101 F	28	64	-0.29	0.35	1.26	1.50	1.03	0.30
154 F	28	64	-0.29	0.35	0.79	-1.20	0.54	-0.70
105 F	27	64	-0.41	0.35	1.34	1.80	1.12	0.40
167 F	27	64	-0.41	0.35	1.17	1.00	1.26	0.60
127 F	26	64	-0.53	0.35	1.02	0.20	0.75	-0.20
151 F	26	64	-0.53	0.35	0.85	-0.80	0.66	-0.40
123 F	25	64	-0.66	0.35	1.20	1.10	1.11	0.40
136 F	22	64	-1.04	0.37	1.00	0.10	1.03	0.30
104 F	20	64	-1.32	0.38	0.93	-0.30	0.76	-0.10
147 F	19	64	-1.46	0.38	0.94	-0.20	0.76	-0.10
153 F	19	64	-1.46	0.38	0.92	-0.30	0.87	0.10
164 F	19	64	-1.46	0.38	0.90	-0.40	0.66	-0.30
194 F	19	64	-1.46	0.38	0.80	-1.00	0.41	-0.80
133 F	18	64	-1.61	0.39	0.92	-0.30	0.54	-0.40
138 F	18	64	-1.61	0.39	0.82	-0.80	0.43	-0.70
199 F	17	64	-1.77	0.40	0.84	-0.70	0.46	-0.60
126 F	16	64	-1.93	0.41	1.05	0.30	0.96	0.20
135 F	16	64	-1.93	0.41	1.31	1.30	1.91	1.20

141 F	16	64	-1.93	0.41	1.43	1.70	1.51	0.80
158 F	16	64	-1.93	0.41	0.74	-1.20	1.47	0.80
161 F	16	64	-1.93	0.41	0.72	-1.20	0.44	-0.60
187 F	16	64	-1.93	0.41	0.55	-2.20	0.25	-1.10
139 F	15	64	-2.10	0.42	1.18	0.80	1.23	0.50
152 F	15	64	-2.10	0.42	0.87	-0.50	0.64	-0.20
165 F	15	64	-2.10	0.42	0.74	-1.10	0.33	-0.80
169 F	14	64	-2.28	0.43	0.82	-0.70	1.03	0.30
171 F	14	64	-2.28	0.43	0.84	-0.60	0.59	-0.20
183 F	13	64	-2.47	0.45	0.83	-0.60	0.41	-0.50
157 F	12	64	-2.68	0.46	1.09	0.40	0.61	-0.20
159 F	12	64	-2.68	0.46	1.50	1.60	1.66	0.90
168 F	11	64	-2.90	0.48	0.62	-1.30	0.23	-0.90
185 F	11	64	-2.90	0.48	1.03	0.20	1.82	1.00
190 F	11	64	-2.90	0.48	0.85	-0.40	0.35	-0.60
162 F	8	64	-3.68	0.55	1.42	1.10	0.57	-0.10
160 F	6	64	-4.34	0.60	1.56	1.40	1.23	0.60
191 F	6	64	-4.34	0.60	1.23	0.70	0.49	-0.20
Mean	25	64	-0.83	0.39	1.00	0.00	0.88	0.00
SD	10.7	0	1.50	0.06	0.22	1.00	0.42	0.60
r	.96							

Table 4.38

Phase Two Self-Portrait-Woman Sub-Test Measure Order Person Statistics

Case	Raw		Ability	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
174 F	40	63	1.11	0.34	1.03	0.20	1.15	0.40
204 F	40	63	1.11	0.34	0.92	-0.40	0.62	-0.60
112 F	39	63	0.99	0.34	1.55	2.80	2.23	1.90
101 F	38	63	0.88	0.34	1.10	0.60	1.40	0.90
125 F	37	63	0.77	0.34	1.50	2.60	1.44	1.00
175 F	37	63	0.77	0.34	1.09	0.50	1.36	0.80
120 F	36	63	0.65	0.34	1.19	1.10	1.31	0.80
156 F	36	63	0.65	0.34	0.87	-0.70	0.80	-0.30
177 F	35	63	0.54	0.34	0.85	-0.80	0.63	-0.70
166 F	32	63	0.20	0.34	0.65	-2.20	0.48	-1.10
193 F	32	63	0.20	0.34	1.09	0.60	1.07	0.30
122 F	31	63	0.09	0.34	1.05	0.30	0.91	0.00
146 F	31	63	0.09	0.34	1.20	1.20	1.15	0.50
205 F	31	63	0.09	0.34	1.13	0.80	1.43	0.90
124 F	30	63	-0.03	0.34	1.08	0.50	1.27	0.70
167 F	30	63	-0.03	0.34	1.06	0.40	0.95	0.10
116 F	29	63	-0.14	0.34	0.86	-0.80	0.73	-0.40
123 F	29	63	-0.14	0.34	0.62	-2.50	0.44	-1.20
126 F	29	63	-0.14	0.34	0.85	-0.80	0.57	-0.80
113 F	25	63	-0.61	0.35	0.79	-1.20	0.62	-0.70
121 F	24	63	-0.73	0.35	1.04	0.30	1.20	0.50
155 F	24	63	-0.73	0.35	0.98	0.00	0.70	-0.50
196 F	24	63	-0.73	0.35	1.07	0.40	0.77	-0.30
154 F	23	63	-0.86	0.35	0.79	-1.20	0.51	-0.90
151 F	22	63	-0.98	0.36	0.82	-1.00	0.65	-0.60
179 F	21	63	-1.11	0.36	1.22	1.20	1.40	0.80
199 F	20	63	-1.24	0.36	1.05	0.30	0.64	-0.50
136 F	19	63	-1.37	0.37	1.19	1.00	1.37	0.80
157 F	19	63	-1.37	0.37	1.29	1.40	1.00	0.20
141 F	18	63	-1.51	0.37	0.68	-1.70	0.46	-0.90
133 F	17	63	-1.65	0.38	1.06	0.40	0.94	0.10
153 F	16	63	-1.80	0.39	0.63	-2.00	0.32	-1.10
135 F	15	63	-1.95	0.39	0.84	-0.70	0.53	-0.50
147 F	14	63	-2.11	0.40	0.77	-1.00	0.44	-0.60
194 F	14	63	-2.11	0.40	0.66	-1.70	0.47	-0.50
139 F	13	63	-2.27	0.41	1.06	0.30	1.25	0.60
160 F	13	63	-2.27	0.41	1.04	0.20	0.76	0.00
171 F	13	63	-2.27	0.41	0.85	-0.60	0.44	-0.50
190 F	13	63	-2.27	0.41	0.84	-0.70	0.57	-0.30
187 F	12	63	-2.45	0.43	0.75	-1.00	1.11	0.40
127 F	11	63	-2.64	0.44	1.05	0.30	0.59	-0.20
165 F	11	63	-2.64	0.44	1.00	0.10	0.85	0.10
152 F	9	63	-3.05	0.48	0.83	-0.50	0.31	-0.60
161 F	9	63	-3.05	0.48	1.13	0.50	1.00	0.30

185 F	9	63	-3.05	0.48	1.02	0.20	0.84	0.20
158 F	8	63	-3.29	0.50	0.78	-0.60	0.26	-0.70
191 F	8	63	-3.29	0.50	1.15	0.50	0.68	0.00
169 F	7	63	-3.56	0.53	0.92	-0.10	0.68	0.00
159 F	6	63	-3.86	0.57	2.18	2.50	2.90	1.50
168 F	5	63	-4.20	0.61	1.18	0.60	0.45	-0.20
Mean	22.1	63	-1.15	0.39	1.01	0.00	0.89	0.00
SD	10.5	0	1.47	0.07	0.26	1.10	0.48	0.70
r	0.94							

Table 4.39 shows that only one female drew a self-portrait in phase three that could be considered misfitting to the Rasch model's expectations. Case ID 164F yielded slightly high mean square statistics (infit Mn Sq. = 1.54, outfit Mn Sq. = 1.72), although the *t* statistics were acceptable (infit $t = 2.0$, outfit $t = .9$). This was the first time that she had produced a drawing that was reported as misfitting to the Rasch model's expectations.

Table 4.39

Phase Three Self-Portrait-Woman Sub-Test Measure Order Person Statistics

Case	Raw		Ability	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
100 F	43	60	1.71	0.37	1.13	0.70	0.71	-0.20
174 F	41	60	1.45	0.36	0.77	-1.30	0.52	-0.60
166 F	33	60	0.45	0.35	1.14	0.80	1.00	0.20
179 F	33	60	0.45	0.35	1.31	1.60	1.11	0.40
112 F	32	60	0.32	0.35	1.26	1.40	1.38	0.90
123 F	32	60	0.32	0.35	1.12	0.70	1.41	0.90
167 F	32	60	0.32	0.35	0.89	-0.50	0.67	-0.60
175 F	32	60	0.32	0.35	1.27	1.40	1.03	0.20
177 F	32	60	0.32	0.35	0.78	-1.20	0.58	-0.90
193 F	32	60	0.32	0.35	1.24	1.30	1.39	0.90
195 F	30	60	0.08	0.35	0.96	-0.20	0.77	-0.40
101 F	29	60	-0.05	0.35	1.15	0.80	1.32	0.80
196 F	29	60	-0.05	0.35	1.05	0.30	0.96	0.10
121 F	28	60	-0.17	0.35	0.67	-1.90	0.51	-1.20
125 F	28	60	-0.17	0.35	0.89	-0.50	0.62	-0.90
120 F	27	60	-0.29	0.35	0.82	-0.90	0.61	-0.90
157 F	26	60	-0.42	0.35	0.96	-0.10	0.73	-0.50
116 F	24	60	-0.67	0.36	1.01	0.10	1.05	0.30
139 F	23	60	-0.80	0.36	0.92	-0.40	0.73	-0.50
122 F	22	60	-0.93	0.36	0.78	-1.20	0.52	-1.00
141 F	22	60	-0.93	0.36	1.01	0.10	1.02	0.20
147 F	19	60	-1.33	0.37	1.26	1.40	1.40	0.80
199 F	17	60	-1.61	0.38	0.94	-0.20	0.53	-0.70
171 F	16	60	-1.75	0.39	1.05	0.30	0.75	-0.20
135 F	15	60	-1.91	0.39	0.83	-0.80	0.57	-0.40
159 F	15	60	-1.91	0.39	0.87	-0.60	0.47	-0.60
169 F	15	60	-1.91	0.39	0.76	-1.20	0.41	-0.80
165 F	14	60	-2.06	0.40	1.23	1.00	1.31	0.60
185 F	14	60	-2.06	0.40	0.92	-0.30	0.92	0.20
194 F	14	60	-2.06	0.40	0.88	-0.50	0.47	-0.50
191 F	13	60	-2.23	0.41	0.69	-1.40	0.37	-0.70
152 F	12	60	-2.40	0.42	1.21	0.90	1.32	0.60
164 F	12	60	-2.40	0.42	1.54	2.00	1.72	0.90
160 F	10	60	-2.77	0.44	0.70	-1.20	0.56	-0.20
168 F	9	60	-2.97	0.46	0.98	0.00	0.90	0.20
Mean	23.6	60	-0.79	0.37	1.00	0.00	0.87	-0.10
SD	9	0	1.21	0.03	0.21	1.00	0.35	0.60
r	0.92							

Examples of Children's Self-Portrait-Woman Drawings

Figures 4.25, 4.26 and 4.27 present examples of the least successful, mean and most successful self-portraits drawn by the females in phase one, two and three of the data collections. In correspondence with the examples of children's drawings for the other GHDT sub-tests, these self-portraits made by the female children in the sample reveal much development over the three phases of data collection.

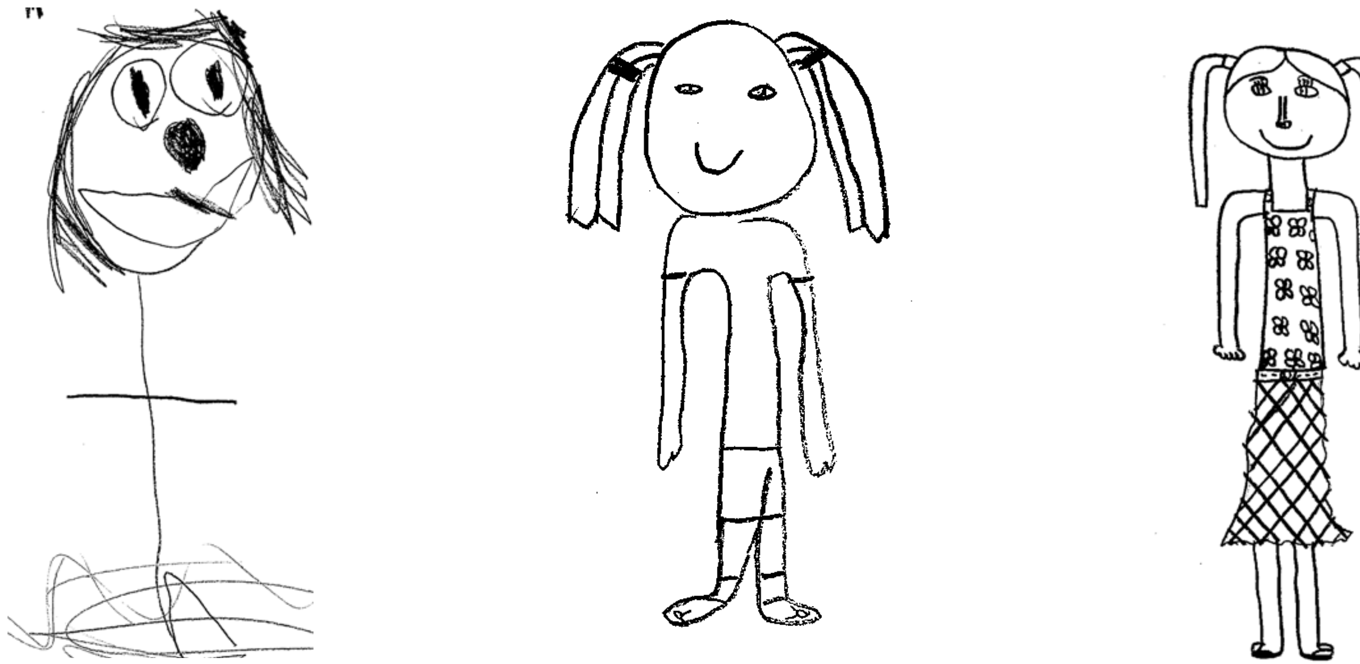
Figure 4.25

Phase One Self-Portrait-Woman Drawings: Examples of Least successful, Mean, and Most Successful

Case ID: 160F Ability Estimate = -4.34; SE = .60	Case ID: 136F Ability Estimate = -1.04; SE = .37	Case ID: 100F Ability Estimate = 2.14; SE = .37

Figure 4.26

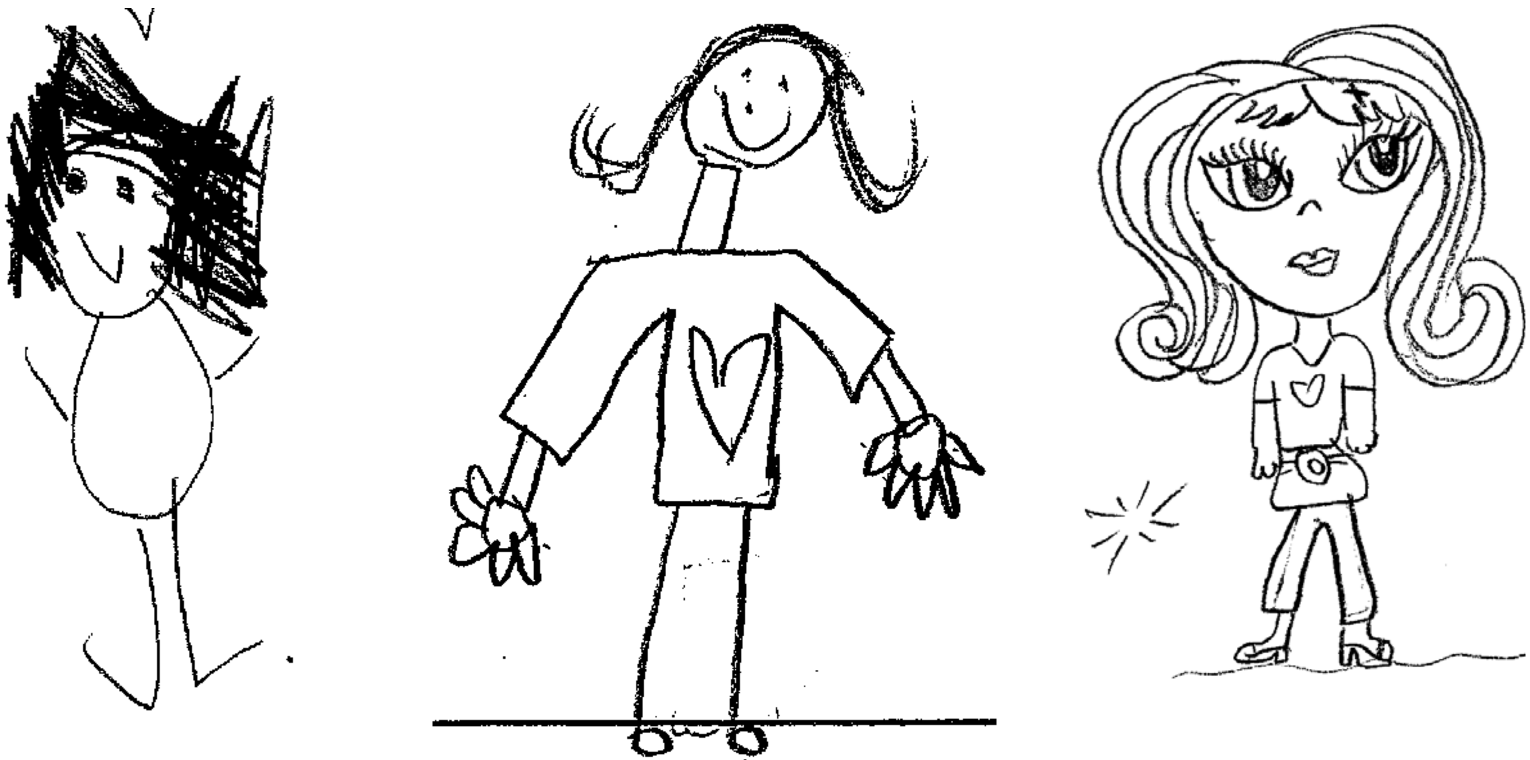
Phase Two Self-Portrait-Woman Drawings: Examples of Least successful, Mean, and Most Successful



Case ID: 168F	Case ID: 199F	Case ID: 174F
Ability Estimate = -4.20; SE = .61	Ability Estimate = -1.24; SE = .36	Ability Estimate = 1.11; SE = .34

Figure 4.27

Phase Three Self-Portrait-Woman Drawings: Examples of Least successful, Mean, and Most Successful



Case ID: 168F Ability Estimate = -2.97; SE = .46	Case ID: 139F Ability Estimate = -0.80; SE = .36	Case ID: 100F Ability Estimate = 1.71; SE = .37

Self-Portrait-Woman Results from Analysis of Complete Data Set

In order to test more stringently the unidimensionality of the SPW sub-test, and to produce person fit statistics for the drawing development and common linking graphs, all three phases of SPW sub-test data were stacked in one data file and Rasch analysed.

Variable Map

Figure 4.28 reveals that the SPW logit scale spans 12 logits from -7 to +5. This result is exactly the same as the variable map produced by the analysis of the stacked SPM sub-test data, and only two logits shorter than both of the variable maps generated by the analysis of the stacked DAM and DAW sub-test data.

Consistent with the other variable maps produced by the Rasch analysis of the stacked sub-test data, the items reported as some of the least difficult for girls to include in their self-portraits are: 1 (head present), 4 (eyes present), 13 (mouth present), 19 (hair I), 24 (arms present), 33 (legs present) and 55 (trunk present).

The Rasch analysis of the stacked SPW sub-test data shows that items 22 (hair IV), 50 (skirt 'modeled' to indicate pleats or draping), 66 (directed lines and form: head outline), 67 (directed lines and form: breast), 68 (directed lines and form: hip contour), 69 (directed lines and form: arms taper), 70 (directed lines and form: calf of leg), and 71 (directed lines and form: facial features) were some of the more difficult items for female children to include in their self-portraits. Most of these items are exclusive to the DAW/SPW sub-tests, therefore they do not have equivalent items that can be scored in the DAM or SPM sub-tests.

Figure 4.28

Stacked Self-Portrait-Woman Sub-Test Person-Item Variable Map

```
          Persons MAP OF Items
              <more>|<rare>
   5                +  70
                    |  66     68
                    |
                    |
                    |  22     67     69
   4                +
                    |
                    |
                    |  71
                    |  50
   3                +
                    |S 54
                    |  27     45     49
                    |  30     31     47     61
                    |
   2               T+  12     8
                 #  |  16     39
                    |  48     62
                ##  |  21     65
                .#  |  11     18
   1            .#  +  14     15     23     29     53
              .###  |  38
                ## S|  52
           #######  |  34     57
           #######  |
   0            ##  +M 51     7
               .##  |  10     3      37     5
             .####  |  25
                ##  |  40     60
              .### M|  32     36
  -1             .  +  42     44     46     6
               .##  |  26     63
                 #  |  28
            .#####  |  2
              .###  |  64
  -2           ###  +  17     58
             .#### S|  9
               .##  |  20
              ####  |
                    |S
  -3            ##  +  35     56
                ##  |
                 .  |
                 .  |  43
                   T|  59
  -4             .  +
                    |
                    |  41
                    |
                 #  |
  -5                +
                    |  24
                    |T
                    |  19
                    |
  -6                +
                    |  55
                    |
                    |
                    |  13     33
  -7                +  1      4
              <less>|<frequ>
EACH '#' IS 2.
```

Item Fit Statistics

Table 4.40 shows that twelve items were reported as misfitting according to the Rasch model's expectations. The underfitting items were: 9 (nose present), 17 (Both chin and forehead shown), 34 (hip), 57 (head-trunk proportion), and 58 (head: proportion). All of these items were reported as overly erratic in the separate analyses of the three phases of SPW and DAW sub-test data. Also, items 17, 57, and 58, in particular, were detected as underfitting in the stacked DAW sub-test data analysis.

Seven items reported as being overly predictable in the stacked SPW sub-test data analysis. These were items: 3 (neck, two dimensions), 7 (eye detail: proportion), 25 (shoulders), 26 (arms at side or engaged in activity behind back), 40 (placement of feet appropriate to figure), 46 (neckline I), and 51 (no transparencies in the figure). All of these items, bar item 7, were reported as overfitting in the Rasch analysis of the stacked DAW sub-test data. Similarly, review of the item fit statistics produced by the separate Rasch analyses of the SPW and DAW sub-test data shows that all items, except for item 7, were also reported as overly predictable.

Table 4.40

Stacked Self-Portrait-Woman Sub-Test Measure Order Item Statistics

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
70	0	147	6.11	1.83	Maximum Extreme Measure			
66	1	147	4.89	1.01	0.98	0.30	0.27	-0.90
68	1	147	4.89	1.01	0.97	0.30	0.24	-1.00
22	2	147	4.18	0.72	0.96	0.20	0.28	-0.90
67	2	147	4.18	0.72	1.02	0.30	0.62	-0.20
69	2	147	4.18	0.72	0.99	0.20	0.43	-0.50
71	4	147	3.46	0.52	0.91	0.00	0.31	-0.90
50	5	147	3.22	0.47	1.02	0.20	0.68	-0.20
54	7	147	2.85	0.40	0.97	0.00	0.49	-0.70
45	8	147	2.70	0.38	0.98	0.00	0.64	-0.40
49	8	147	2.70	0.38	0.84	-0.50	0.35	-1.10
27	9	147	2.56	0.36	1.09	0.40	1.71	1.10
31	10	147	2.44	0.34	1.02	0.20	0.56	-0.70
47	10	147	2.44	0.34	0.96	-0.10	0.69	-0.40
61	10	147	2.44	0.34	0.93	-0.20	0.53	-0.70
30	11	147	2.33	0.33	1.02	0.20	0.62	-0.60
12	14	147	2.03	0.30	0.99	0.00	0.72	-0.40
8	15	147	1.95	0.29	1.09	0.50	1.12	0.40
39	16	147	1.87	0.28	0.96	-0.10	0.67	-0.60
16	18	147	1.71	0.27	0.88	-0.70	0.54	-1.10
48	21	147	1.51	0.26	0.94	-0.30	0.84	-0.30
62	21	147	1.51	0.26	1.22	1.40	1.50	1.20
21	22	147	1.44	0.25	0.91	-0.60	0.56	-1.20
65	24	147	1.32	0.24	1.02	0.20	0.99	0.10
11	27	147	1.15	0.23	1.00	0.10	1.28	0.80
18	27	147	1.15	0.23	1.15	1.20	1.30	0.90
14	28	147	1.09	0.23	0.90	-0.80	0.64	-1.10
53	28	147	1.09	0.23	0.80	-1.70	0.49	-1.70
23	29	147	1.04	0.23	1.12	1.00	1.54	1.50
15	30	147	0.99	0.23	0.93	-0.60	0.74	-0.70
29	31	147	0.94	0.23	1.12	1.10	1.46	1.30
38	34	147	0.79	0.22	1.10	0.90	0.87	-0.30
52	36	147	0.69	0.22	1.03	0.30	1.47	1.50
34	42	147	0.42	0.21	1.59	5.40	2.58	4.60
57	42	147	0.42	0.21	1.63	5.70	2.33	4.10
7	50	147	0.09	0.20	0.75	-3.00	0.57	-2.40
51	50	147	0.09	0.20	0.68	-4.00	0.51	-2.90
10	55	147	-0.12	0.20	0.93	-0.70	0.93	-0.30
3	57	147	-0.20	0.20	0.72	-3.50	0.61	-2.50
5	58	147	-0.24	0.20	1.03	0.30	1.07	0.50
37	58	147	-0.24	0.20	0.82	-2.10	0.68	-2.00
25	62	147	-0.39	0.20	0.71	-3.50	0.58	-3.00
60	68	147	-0.63	0.20	1.05	0.60	0.97	-0.10
40	69	147	-0.67	0.20	0.81	-2.20	0.69	-2.30

36	70	147	-0.70	0.20	0.97	-0.30	0.87	-0.90
32	74	147	-0.86	0.20	1.11	1.10	1.06	0.50
46	76	147	-0.94	0.20	0.69	-3.60	0.60	-3.20
6	78	147	-1.02	0.20	0.90	-1.00	1.09	0.70
42	80	147	-1.09	0.20	0.82	-1.90	0.76	-1.80
44	80	147	-1.09	0.20	0.71	-3.30	0.77	-1.70
63	81	147	-1.13	0.20	1.06	0.70	0.99	0.00
26	83	147	-1.21	0.20	0.72	-3.20	0.68	-2.50
28	89	147	-1.45	0.20	1.19	1.80	1.22	1.40
2	93	147	-1.62	0.20	0.95	-0.50	0.84	-1.00
64	97	147	-1.78	0.21	1.23	2.20	1.42	2.10
58	103	147	-2.05	0.21	1.81	6.40	2.90	6.40
17	104	147	-2.09	0.21	1.40	3.50	1.78	3.20
9	107	147	-2.23	0.22	1.51	4.20	1.75	2.80
20	110	147	-2.37	0.22	0.87	-1.20	0.70	-1.30
35	122	147	-3.03	0.25	0.86	-1.00	0.61	-1.20
56	123	147	-3.09	0.25	0.83	-1.20	0.67	-0.90
43	130	147	-3.59	0.29	0.84	-0.80	0.42	-1.50
59	133	147	-3.85	0.31	0.87	-0.60	0.46	-1.10
41	138	147	-4.42	0.37	0.90	-0.30	0.53	-0.60
24	142	147	-5.12	0.48	0.99	0.10	0.73	-0.10
19	144	147	-5.69	0.60	0.86	-0.10	0.31	-0.70
55	145	147	-6.13	0.73	0.72	-0.30	0.07	-1.70
13	146	147	-6.86	1.01	0.96	0.30	0.13	-1.40
33	146	147	-6.86	1.01	1.04	0.40	0.65	-0.20
1	147	147	-8.09	1.83	Minimum Extreme Measure			
4	147	147	-8.09	1.83	Minimum Extreme Measure			
Mean	57.9	147	-0.14	0.40	0.99	0.00	0.86	-0.30
SD	47.4	0	3.06	0.37	0.21	2.00	0.54	1.70
r	0.98							

Summary

The Rasch analyses described in the previous sections were undertaken in response to the research questions outlined in Chapter Two. The separate Rasch analyses of each phase of each sub-test were conducted to inform the principle research questions which concerned the unidimensionality of each of the sub-tests, and the adherence of the resultant measurement scales to the Rasch model's expectations. Furthermore, these separate analyses offered an inside 'look' at each sub-test's performance in each phase, and each person's performance on each sub-test in each phase. Given that it was possible to Rasch analyse the sub-tests of the GHDT, and produce meaningful statistics about each item and person in each phase, additional Rasch analyses were performed in response to the subsequent questions of this research.

Therefore, in addition to the separate analyses performed, the data from all phases for each sub-test were also Rasch analysed (i.e., the stacked data analyses). These analyses were a more stringent test of unidimensionality and they produced person fit statistics which were useful for two reasons in particular. First, the person fit statistics produced by the Rasch analyses of the stacked sub-test data could be used to examine the growth of the children's drawings across the data collection phases. It was useful to plot the person fit statistics produced by the stacked data analyses in to Excel® graphs as it enabled the development of the drawings for the DAM and DAW sub-tests to be explored. Furthermore, the graphs offered information about which drawings might have developed the most or least, which was useful for inferring which items and/or sub-tests would be most useful to the development of the HFDC.

Second, the person and item fit statistics could be used in common linking plots to examine more closely if the children and items performed comparably across the sub-tests. This latter investigation was important especially given that the sub-tests were said to investigate the same latent construct; therefore, one would expect that the children and items would perform comparably in these graphs.

The drawing development graphs produced by the plotting of the person fit statistics from the stacked sub-test data analyses are presented and described in the following section. The common linking graphs follow this.

Children's Drawing Development across the Phases

Exploring the development of the children's drawings over time was useful as it enabled the researcher to explore how growth in young children's human figure drawings occurred across each age range, or school year level, and across each sub-test too. Using some of the results produced by the Rasch analysis of the stacked sub-test data, graphs were produced to represent the average development of the children's DAM and DAW drawings over the twelve month data collection period.

Draw-a-Man Sub-Test

Figure 4.29 presents the average drawing ability estimates of the children in prep, year one, two, three and four on the DAM sub-test. The data collection phases are plotted on the y axis and the ability estimates in logits are plotted on the x axis.

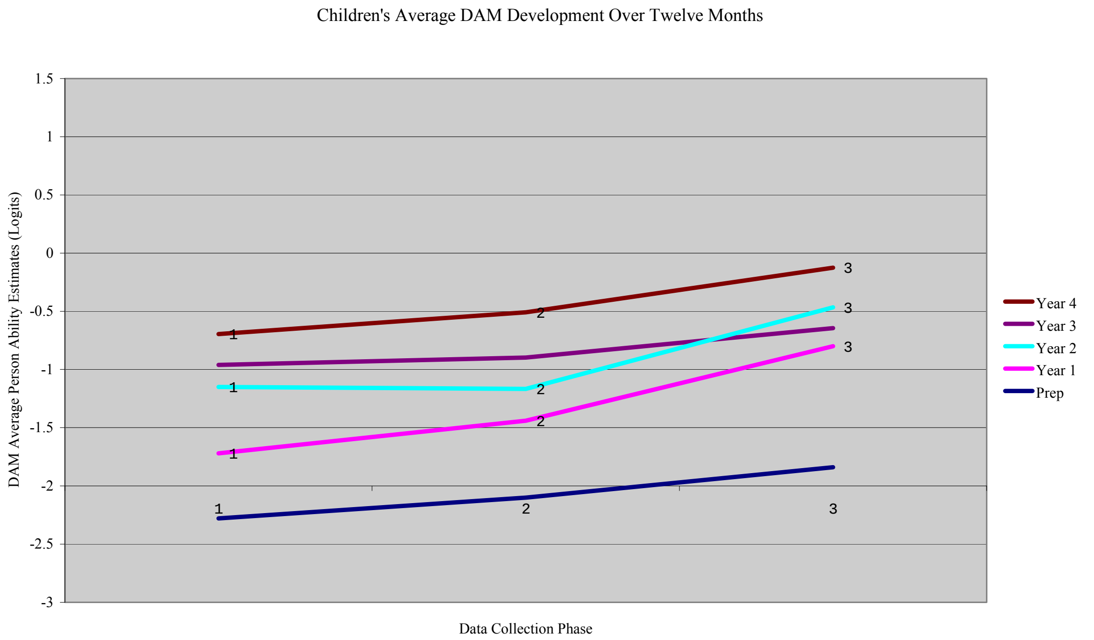
The graph shows that children in year one experienced, on average, the largest amount of development in their drawings of men across the twelve month data collection period. Indeed, the graph reveals about one logit of development on average, from -1.75 to +0.25 logits. Conversely, children in years three and four seem to have experienced the least amount of development in their drawings of men. The graph indicates less than

+0.5 logits of development for these children's drawings over the twelve-month period. Figure 4.29 shows that children in prep and year two experienced, on average, slightly more than +0.5 logits of development in their DAM drawings over the twelve month period.

Whilst the results plotted in Figure 4.29 indicate that the year two children out-performed the year three children in data collection phase three, this was actually a distorted result. Phase three data collection included a disproportionate number of children for year two. The result was based upon the ability estimates of only three children (two girls and one boy), one of which was an extremely high achiever yielding an ability estimate of +1.62 logits in phase three. Should a more representative number of children have been available one could reasonably expect that the average drawing development for children in year two would have been alignment with that of the other year levels.

Figure 4.29

Development of Children's Draw-a-Man Drawings over the Three Phases



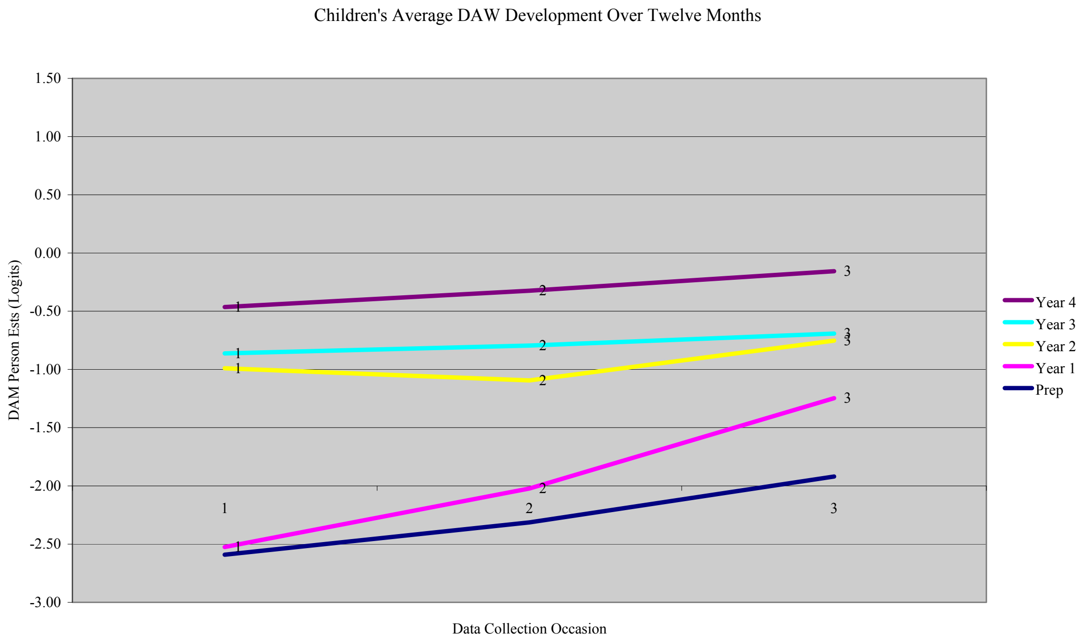
Draw-a-Woman Sub-Test

Figure 4.30 presents the average drawing ability estimates of the children in prep, year one, two, three and four on the DAW sub-test. Once again, the data collection occasion is plotted on the y axis and the ability estimates on the x axis.

The graph in Figure 4.30 reveals substantially more variation in the children's drawing development than that seen in the DAM drawing development graph in Figure 4.29. While children in year two, on average, seem to have experienced a slight decline in their development of drawings of women in phase two, the drawings of women produced by children in years three and four appear to barely change over the twelve month data collection period. On the other hand, the children in prep experienced a little more than +0.5 logits of development in their drawings of women between phase one and phase three. The children in year one, once again, experienced the most amount of development in their human figure drawings; Figure 4.30 shows nearly +1.25 logits of development in their drawings of women over the twelve-month data collection period. However, the large amount of development in the year one children's DAW drawings could attributed to the atypically low starting level of achievement in phase one (estimated at around -2.50 logits). The results indicate that the children in prep and year one did not draw women as well as might be expected from the DAM results presented in Figure 4.29.

Figure 4.30

Development of Children's Draw-a-Woman Drawings over the Three Phases



Due to the unexpectedly large attrition of the sample drawing development graphs for the SPM and SPW sub-tests were not included in this thesis. These graphs presented ambiguous information about the development of the boys and girls self-portraits as, in some instances, only one or two children were representative of a sub-sample such as 'male children in year two' or 'female children in year three'.

Common Linking Graphs

The GHDT comprises four sub-tests, the DAM, DAW, SPM, and SPW, of which children generally complete up to three (as only male children complete the SPM sub-test, and only female children complete the SPW sub-test). Given that the GHDT is held to be a test which investigates the latent construct of intellectually maturity, it could be inferred then that each of the sub-tests investigates the latent construct of intellectual maturity (or intellectual development). In order to examine whether children performed comparably on the different sub-tests, as one would expect, the person fit statistics produced by the Rasch analyses of the stacked data were used to produce common person linking invariance graphs. Clearly, a common person linking graph could not be constructed for comparing the SPM and SPW sub-tests as there were no common persons across these sub-tests (as only female children completed a SPW drawing and only male children completed a SPM drawing).

Person Measure Invariance: DAM vs. DAW

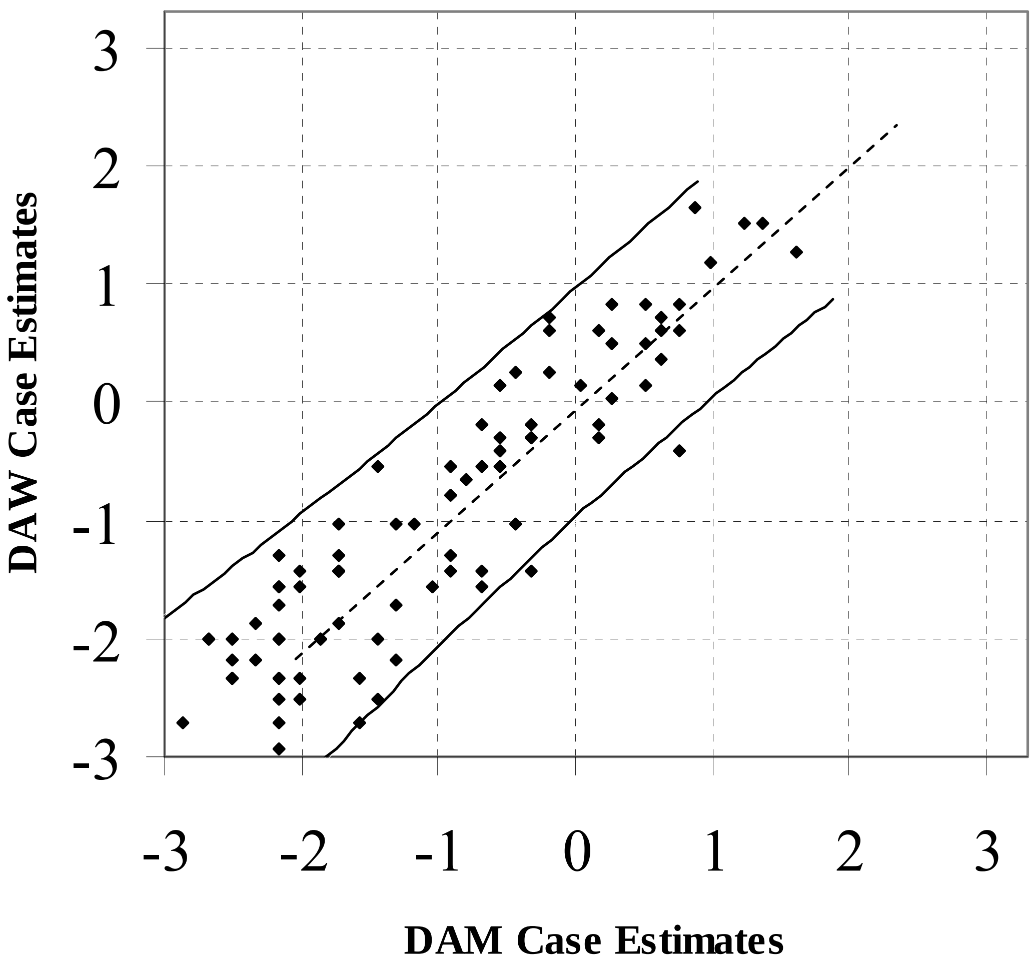
Figure 4.31 displays the common person linking graph created using the person estimate and fit statistics produced by the Rasch analysis of the stacked DAM and DAW sub-test data. The DAW person ability estimates are plotted on the y axis and the DAM person ability estimates are plotted on the x axis. Both logit scales span 6 logits from -3 to +3, although the person ability estimates only span about 5 logits from -3 to +2.

The error bands reveal that person measures remain invariant (within error) across the DAM and DAW sub-tests, even though just two children were more successful on the DAM sub-test than the DAW sub-test. Case ID 192M's ability estimates were -0.32 (DAM; SE = .34) and -1.34 (DAW; SE = .37), whereas case ID 108M was plotted at +0.75 (DAM; SE = .35) and -.34 (DAW; SE = .34).

The detection of only two ‘outliers’ from N = 246 total data points on the DAM/DAW common person linking graph is consistent with the claim that both tests are said to investigate the exact same latent construct.

Figure 4.31

Common Person Linking DAM vs. DAW



Person Measure Invariance: DAM vs. DAW 50 Common Items Only

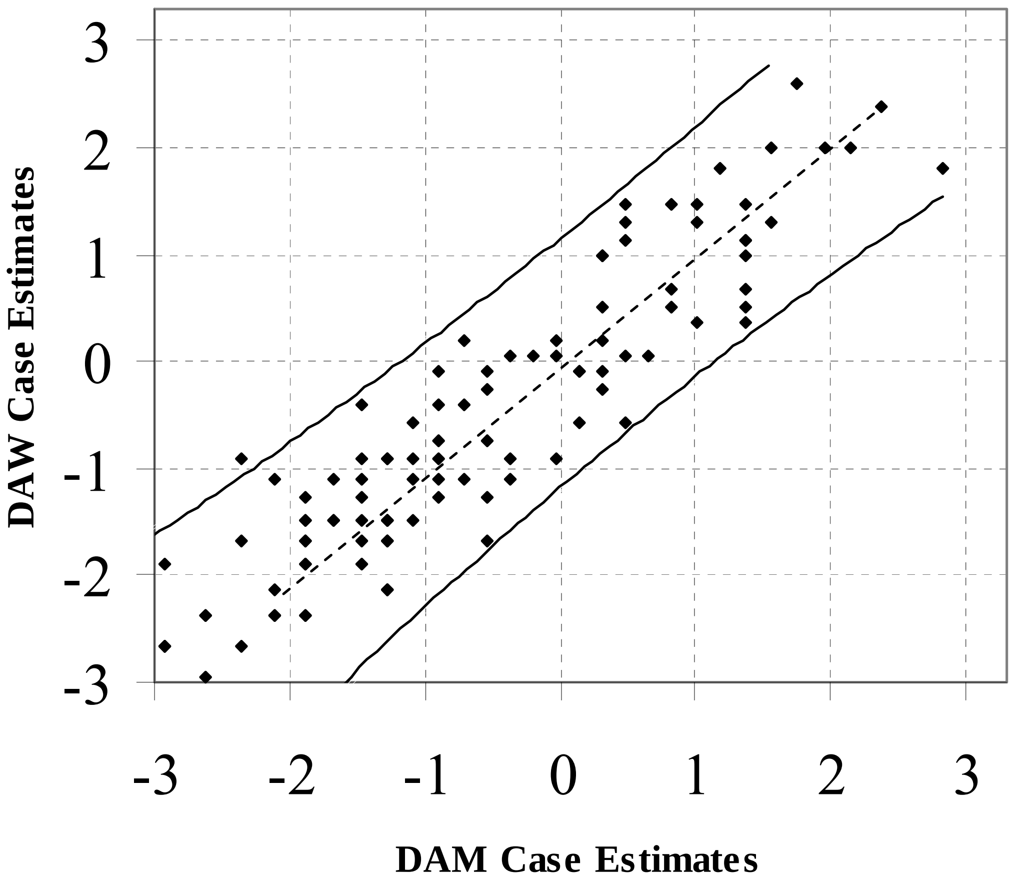
In an effort to investigate the GHDT thoroughly, another common person linking graph was constructed to compare the children's performances on the 50 common items across the DAM and DAW sub-tests.

Figure 4.32 displays the DAW ability estimates on the y axis and the DAM ability estimates on the x axis. Both logits scales span six logits from -3 to +3 and, unlike the common person linking graph in figure 4.31, the ability estimates were plotted across the entire length of the logit scales. Only one child (from N = 246 data points) had more success on one test than the other: case ID 187F was more successful on the 50 item DAW sub-test (estimate = -.85 logits; SE = .42) than she was on the 50 item DAM sub-test (estimate = -2.35 logits; SE = .51).

The results suggest that the 50 common items could provide a more parsimonious assessment of children's human figure drawings than would two separate 70+ item sub-tests. The variable maps, summary statistics, together with the item and person statistics produced by the Rasch analysis of the 50 common items in the DAM and DAW sub-tests are presented in Appendix D.

Figure 4.32

Common Person Linking DAM vs. DAW 50 Common Items Only



Person Measure Invariance: Drawings of Adults vs. Drawings of Children

The next step investigated whether the children performed more successfully when drawing children or adults (i.e. the self-portrait sub-test versus the appropriate DAM or DAW sub-test). Therefore, the male children's person measures from the DAM sub-test were used together with their person measures from the SPM sub-test. Then, the female children's person measures from the DAW sub-test were used together with those from the SPW sub-test.

The DAM versus the SPM common person linking plot (Figure 4.33) reveals two boys (from a total of N = 99): case ID 197M (DAM estimate = -2.16, SE = .40; SPM estimate = -0.18, SE = .37) and case ID 110M (DAM estimate = -1.72, SE = .38; SPM estimate = +1.01, SE = .36), who found it measurably easier to produce drawings of themselves than drawings of men.

Figure 4.33

Common Person Linking DAM vs. SPM

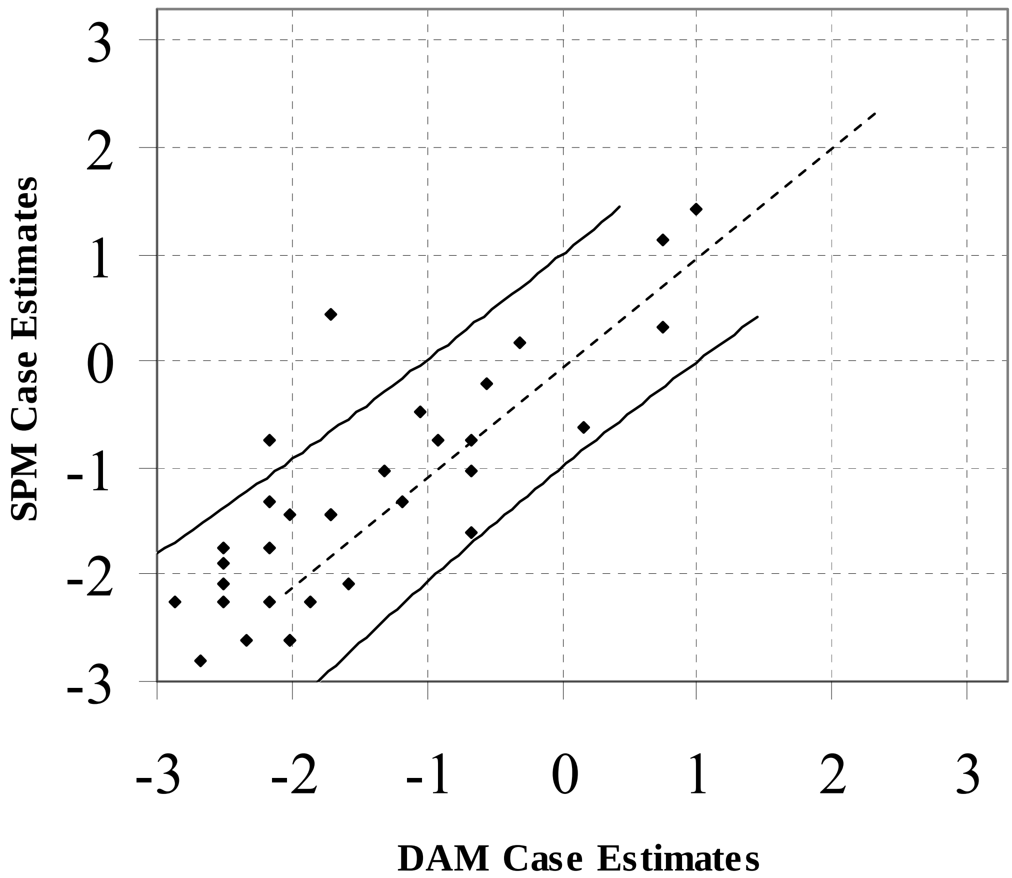


Figure 4.34

Common Person Linking DAW vs. SPW

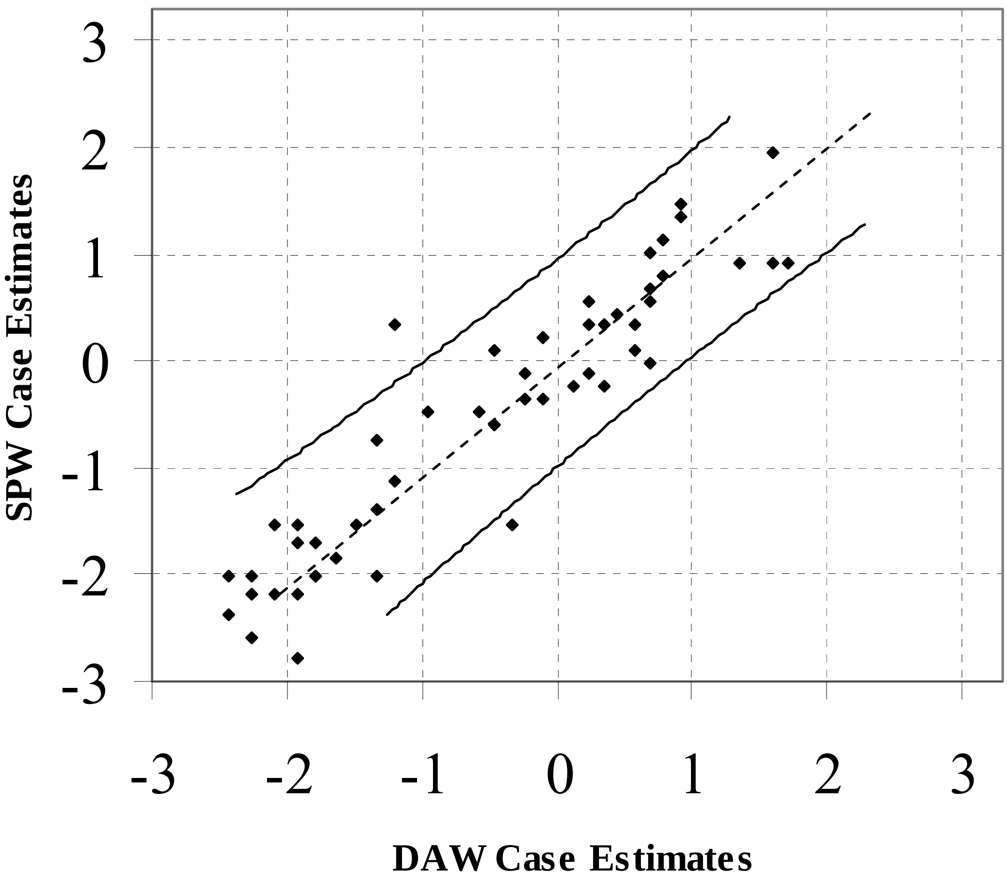


Figure 4.34 reveals two girls: Case ID 145F (DAW estimates = -1.26, SE = .36; SPW estimate = +0.41, SE = .34) and ID 153F (DAW estimate = -0.40, SE = .34; SPW estimate = -1.46, SE = .39) who found it measurably easier to produce drawings of themselves than drawings of women.

The Construction of the Human Figure Drawing Continuum

The remaining analytical task was aimed at supporting the development of a more general scoring guide for scoring all children's human figure drawings. The starting point for the development of the HFDC was the 50 common items across the DAM and DAW sub-tests. These items were not only useful because they could be applied to drawings of males and females, but also because examination using the common linking graphs indicated that these 50 items were generally just as effective for the 70 plus item DAM and DAW sub-tests. As there were no children that continually presented with misfitting performances across the sub-tests, but some items were continually reported as being misfitting the Rasch model's expectations, the removal of misfitting items was the first step undertaken towards the construction of a general scoring guide.

To thoroughly investigate the draft Human Figure Drawing Continuum (HFDC), the children's DAM *and* DAW drawings from all three phases of data collection were included in the Rasch analysis of the 45 item HFDC. This meant that the total sample of 246 children's drawings was doubled, becoming a sample of 492 children's drawings (DAM + DAW drawings from all three phases). The DAM and DAW drawings were used as they yielded more reliable results according to the expectations of the Rasch model than did the SPM and SPW sub-test drawings.

The variable map in Figure 4.35 shows an excellent spread of items on the logit scale that ranges an expansive 14 logits from -8 to +5 logits. The summary statistics in Table 4.41 show that the draft HFDC has an item standard deviation of 3.04 and item reliability index of .99. The upper section of Table 4.41 reveals a person mean raw score of 19.4, a mean person measure of -.59, and a person reliability index of .89.

Figure 4.35

Human Figure Drawing Continuum Person-Item Variable Map

```
          Persons - MAP - Items
               <more>|<rare>
    6                +T
                     |
                     |
                     |  45
    5                +  18
                     |
                     |  40
                     |
    4                +  44
                  .  |  26
                     |
                  .  |  12     39
    3             #  +S 20
                  .  |  11
                .## T|  21
                 ##  |  13     17     43
    2           .##  +
          .########  |  38
             ######  |  14     7
             .#####  |  30     5
    1          #### S+
         .#########  |  24     3
            .######  |  31
             .#####  |  9
    0         #####  +M 25     29
       ############  |  36
              .#### M|  22     33     42     6
          #########  |  16
   -1    ##########  +
         .#########  |  2      41
       .###########  |  8
       .###########  |  19
   -2    ##########  +
          .######## S|  28     37
             .#####  |
                ###  |  35
   -3                +S
              .####  |
                     |  15
                 .# T|
   -4             .  +
                     |
                     |  32
                 .#  |  10
   -5                +
                  .  |
                     |  23     34
                  .  |
   -6                +T
                     |  27
                     |
                     |
   -7                +
                     |  4
                     |
                     |
   -8                +  1
               <less>|<frequ>
 EACH '#' IS 3.
```

Table 4.41

Human Figure Drawing Continuum Summary Statistics

SUMMARY OF 492 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	19.4	45.0	-.59	.49	1.01	.0	1.01	.1
S.D.	7.1	.0	1.60	.06	.30	1.0	1.44	1.0
MAX.	39.0	45.0	3.82	.81	2.51	3.5	9.90	9.9
MIN.	4.0	45.0	-5.80	.44	.43	-2.5	.09	-1.0

REAL RMSE .52 ADJ.SD 1.52 SEPARATION 2.89 Person RELIABILITY .89
MODEL RMSE .49 ADJ.SD 1.53 SEPARATION 3.10 Person RELIABILITY .91
S.E. OF Person MEAN = .07

VALID RESPONSES: 99.9%
Person RAW SCORE-TO-MEASURE CORRELATION = .99 (approximate due to missing data)
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .90 (approximate due to missing data)

SUMMARY OF 44 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	206.2	492.0	.00	.19	.99	.0	1.44	.4
S.D.	160.1	.0	3.04	.12	.14	2.2	1.76	2.9
MAX.	490.0	492.0	5.15	.72	1.31	5.1	9.90	9.9
MIN.	5.0	492.0	-7.33	.11	.72	-5.6	.15	-3.9

REAL RMSE .23 ADJ.SD 3.03 SEPARATION 13.14 Item RELIABILITY .99
MODEL RMSE .23 ADJ.SD 3.03 SEPARATION 13.27 Item RELIABILITY .99
S.E. OF Item MEAN = .46

MINIMUM EXTREME SCORE: 1 Items
UMEAN=.000 USCALE=1.000

Table 4.42

Human Figure Drawing Continuum Measure Order Item Statistics

Item	Raw		Difficulty	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
45	5	492	5.15	0.46	0.90	-0.10	0.15	-2.70
18	6	492	4.96	0.42	0.97	0.10	6.72	5.60
40	9	492	4.52	0.35	1.00	0.10	5.67	4.90
44	14	492	4.04	0.28	1.07	0.40	2.47	2.40
26	17	492	3.82	0.26	1.11	0.60	9.90	9.90
39	26	492	3.31	0.22	0.91	-0.50	0.36	-2.00
12	28	492	3.22	0.21	0.86	-0.90	0.30	-2.40
20	36	492	2.90	0.19	1.00	0.00	0.56	-1.40
11	41	492	2.73	0.18	0.89	-0.90	1.01	0.20
21	52	492	2.41	0.16	1.11	1.10	1.12	0.50
13	54	492	2.36	0.16	1.31	2.90	2.11	3.00
43	56	492	2.30	0.16	1.03	0.40	1.43	1.40
17	57	492	2.28	0.16	0.89	-1.10	0.67	-1.20
38	79	492	1.79	0.14	0.85	-1.90	0.56	-2.20
14	89	492	1.59	0.14	1.14	1.80	1.13	0.60
7	93	492	1.52	0.14	0.84	-2.20	0.58	-2.30
30	105	492	1.31	0.13	0.96	-0.50	0.66	-2.00
5	114	492	1.16	0.13	1.03	0.50	1.63	3.00
3	145	492	0.69	0.12	0.84	-2.70	0.73	-2.00
24	146	492	0.67	0.12	0.78	-3.80	0.62	-3.00
31	161	492	0.46	0.12	0.82	-3.30	0.63	-3.10
9	181	492	0.20	0.11	1.03	0.60	0.97	-0.20
29	196	492	0.00	0.11	0.80	-3.90	0.65	-3.50
25	203	492	-0.08	0.11	0.72	-5.60	0.63	-3.90
36	222	492	-0.32	0.11	1.16	2.90	1.32	2.80
42	233	492	-0.45	0.11	1.12	2.30	1.17	1.60
6	239	492	-0.53	0.11	1.10	1.80	1.24	2.30
22	239	492	-0.53	0.11	1.15	2.80	1.18	1.70
33	245	492	-0.60	0.11	0.99	-0.10	1.05	0.50
16	266	492	-0.85	0.11	1.03	0.60	0.93	-0.60
41	289	492	-1.13	0.11	1.27	5.10	1.53	4.30
2	303	492	-1.30	0.11	1.21	4.00	1.48	3.70
8	323	492	-1.55	0.11	1.20	3.70	1.43	3.00
19	330	492	-1.64	0.11	1.18	3.40	1.71	4.40
28	375	492	-2.27	0.12	0.81	-3.20	0.67	-2.00
37	380	492	-2.34	0.12	0.85	-2.50	0.58	-2.60
35	404	492	-2.74	0.13	0.90	-1.30	1.01	0.10
15	435	492	-3.40	0.16	0.93	-0.70	0.65	-1.30
32	468	492	-4.54	0.23	0.88	-0.70	1.34	0.90
10	472	492	-4.77	0.25	0.92	-0.40	1.01	0.20
34	480	492	-5.38	0.31	0.92	-0.20	1.74	1.50
23	481	492	-5.48	0.32	0.90	-0.30	1.40	0.90
27	486	492	-6.16	0.43	0.99	0.10	0.36	-1.70
4	490	492	-7.33	0.72	0.99	0.20	0.16	-2.70

1	492	492	-9.26	1.83	Minimum Estimated Measure			
Mean	212.6	492	-0.21	0.23	0.99	0.00	1.44	0.40
SD	163.8	0	3.30	0.27	0.14	2.20	1.76	2.90
r	0.99							

Review of Table 4.42 reveals that a number of the HFDC were reported as misfitting to the Rasch model's expectations. However, it should be noted that most of the overly erratic items 18 (hair IV), 40 (clothing V), 44 (directed lines and form: head outline), and 26 (elbow joint shown – see HFDC Item Conversion Table in Appendix E) were reported as being quite difficult for this sample, and thus, only a comparatively small number of children included these items in their human figure drawings. The largish *t* outfit statistics indicate that some children with an ability estimated to be far from the difficulty levels of these items unexpectedly received credit for these items. Linacre (2002) states that theses response patterns are amongst the less difficult to diagnose and remedy with further research and examination of the items. And given that this is the very first Rasch analysis of the 45 item HFDC, better fit statistics could be expected with further research across more diverse samples.

On the other hand, items 45 (directed lines and form: facial features), 24 (shoulders), 31 (feet V: detail), 29 (feet II: proportion), and 25 (arms at side or engaged in activity), which yielded *t* statistics higher than -2.0 and Mean Squares of around 0.75, were reported as overly predictable. However, as Bond and Fox (2007) suggest, removing such items might deprive the instrument of some of its best criteria.

Just below the mean person ability and mean item difficulty levels of the instrument were items 41 (motor coordination: lines), 2 (neck present), 8 (nose present), and 19 (fingers present). These items were reported as yielding underfitting infit and outfit *t* statistics and mean square values. Linacre (2002) states that underfitting infit statistics can be more difficult to diagnose and remedy as they can suggest indifference to the underlying latent construct or idiosyncratic members of a sample, as well as lucky guesses and careless mistakes.

HFDC Comparisons with the DAM and DAW Sub-tests

To verify the effectiveness of the HFDC as a scoring mechanism in comparison to the DAM and DAW sub-tests of the GHDT, common person linking graphs were constructed using the person measures derived from each test.

Figure 4.36 presents the HFDC versus DAM common person linking plot and Figure 4.37 the HFDC versus DAW common person linking graph. Apart from one child on the borderline (case ID 128M: DAW estimate = -2.66, err = .44; HFDC estimate = -1.43; err = .49) all other person measures remain invariant (within measurement error) even when the shortened 45 items general scoring guide is used.

Figure 4.36

Human Figure Drawing Continuum versus the Draw-a-Man Sub-Test

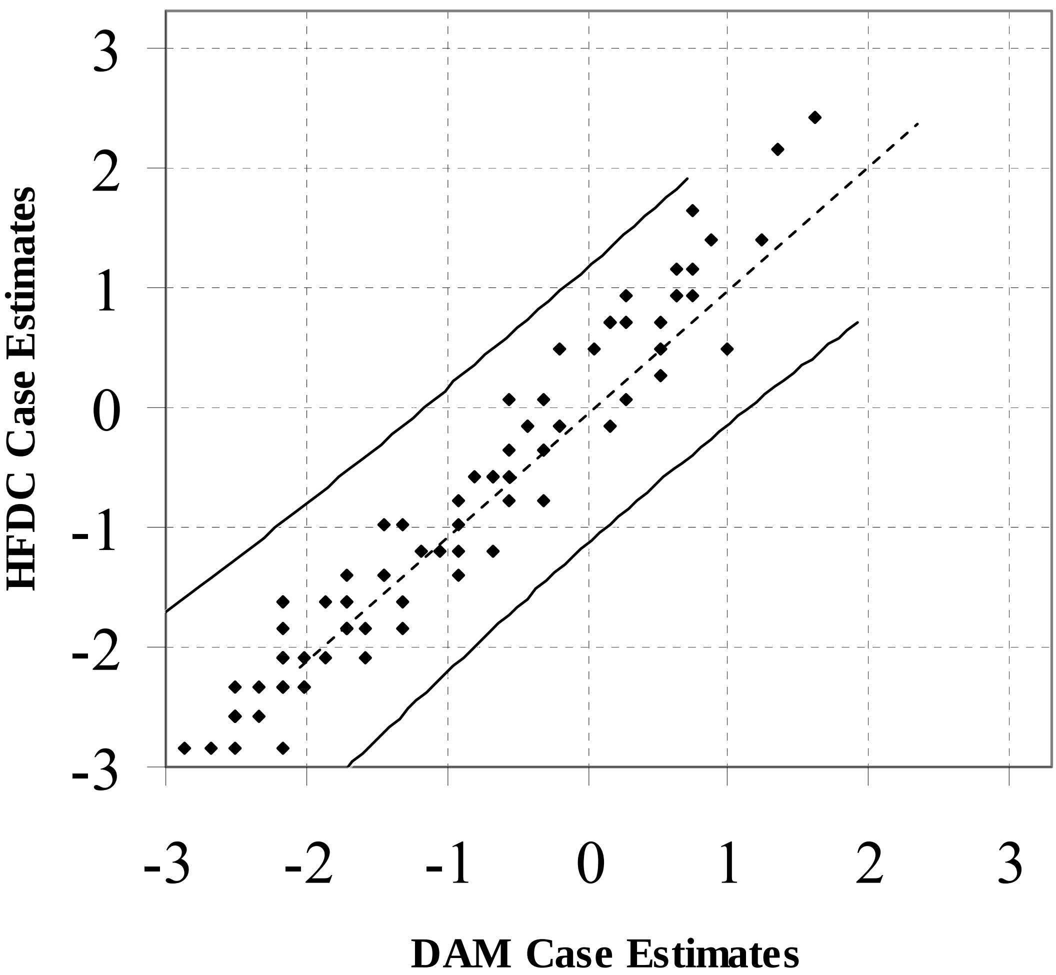
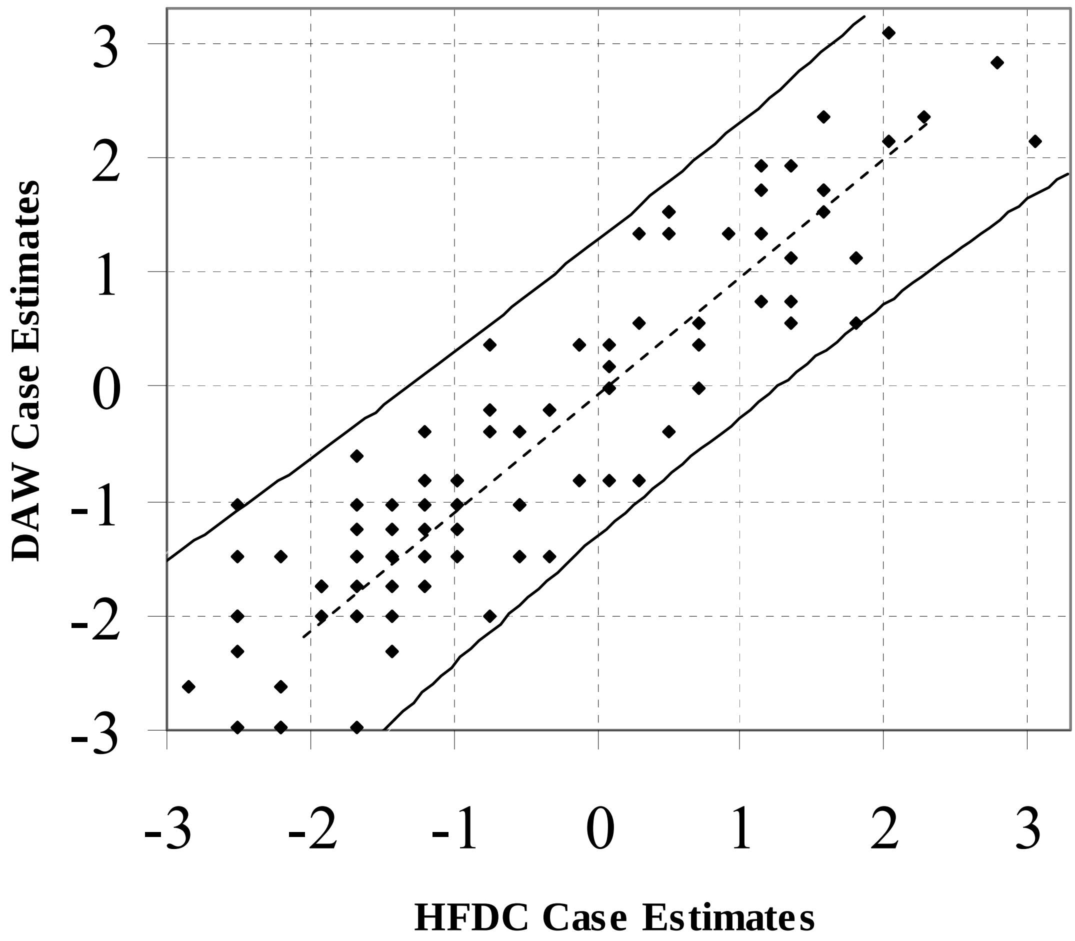


Figure 4.37

Human Figure Drawing Continuum versus the Draw-a-Woman Sub-Test



HFDC Comparisons with the SPM and SPW Sub-tests

Figure 4.38 presents the HFDC versus SPM common person linking graph and Figure 4.39 presents the HFDC versus the SPW common person linking graph. Similar to the HFDC versus DAM and HFDC versus DAW common person linking plots, the HFDC versus SPM and HFDC versus SPW plots show remarkable person measure invariance (within measurement error) in the children's performances. In fact, none of the children's performances resulted in any 'outliers' on these graphs.

Figure 4.38

Human Figure Drawing Continuum versus the Self-Portrait-Man Test

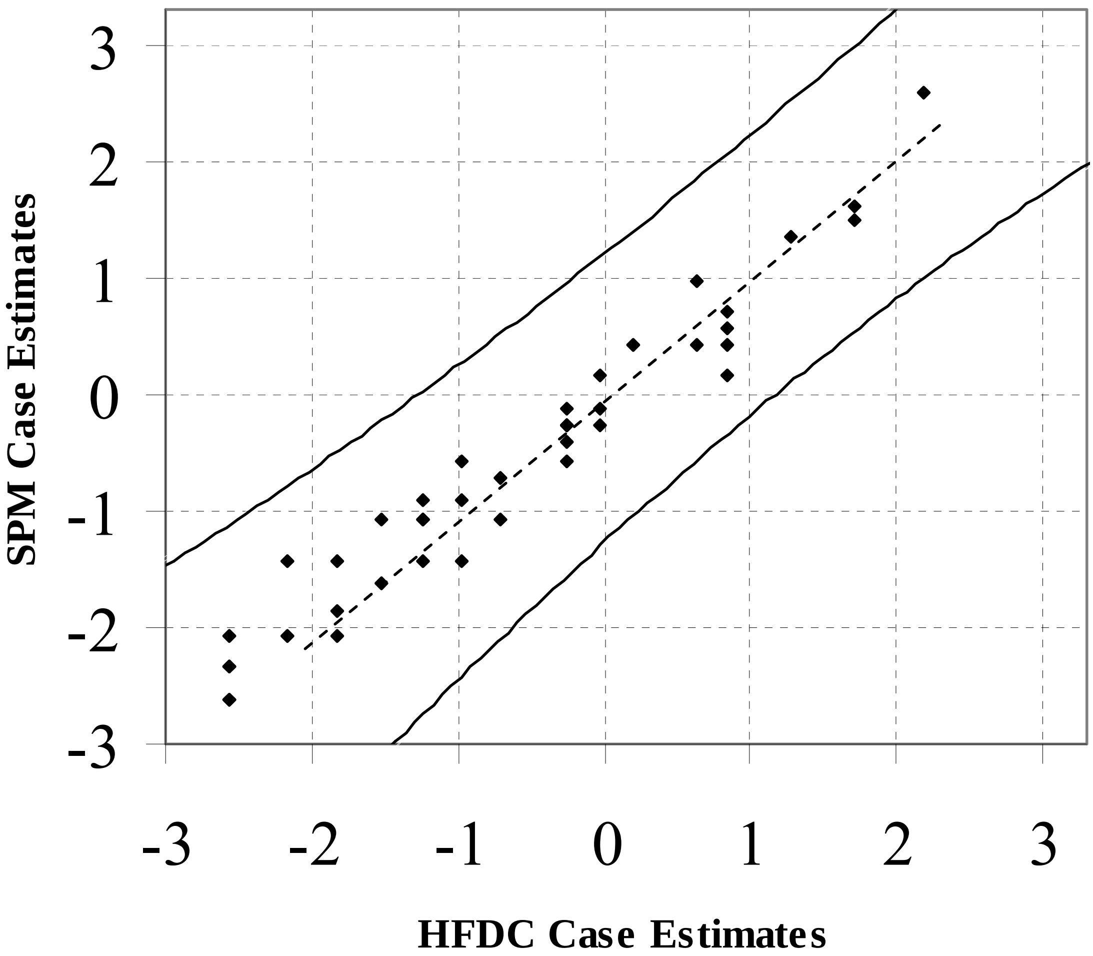
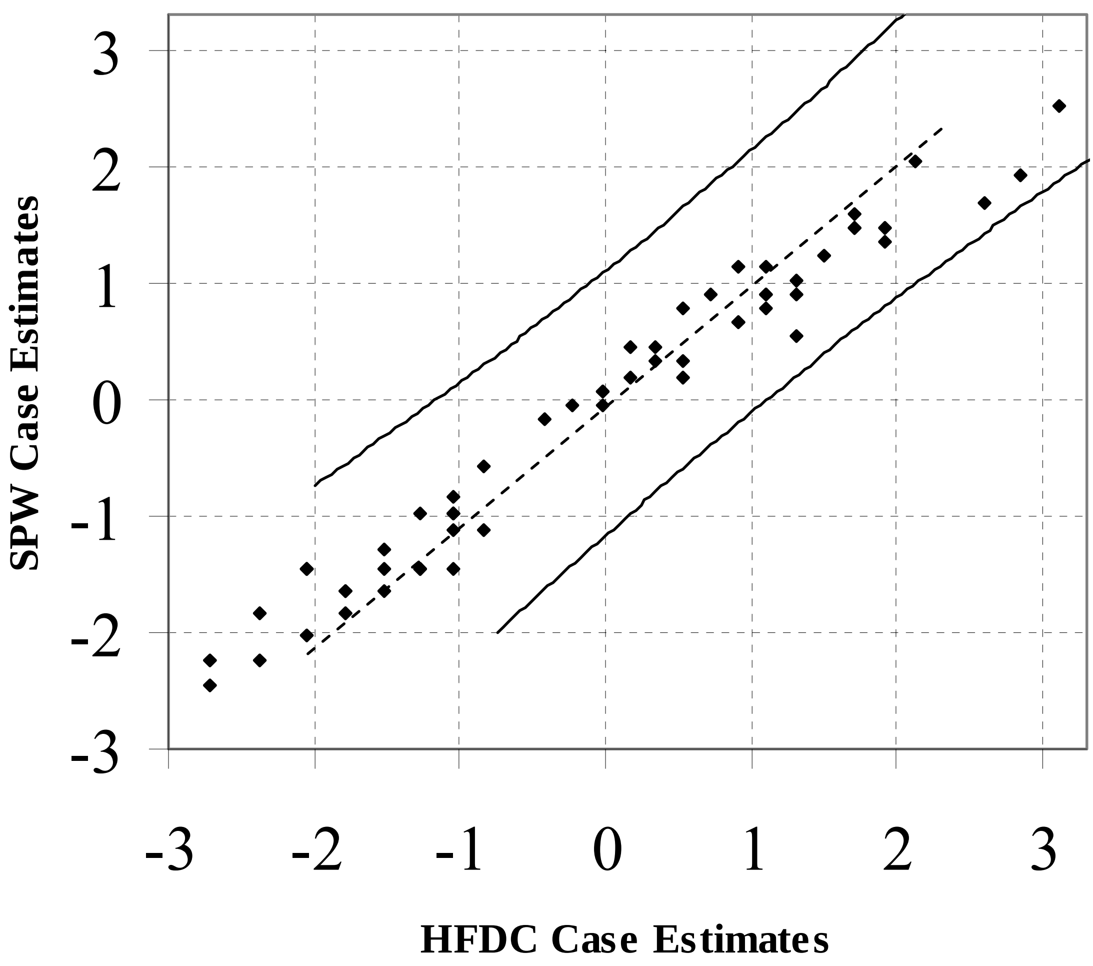


Figure 4.39

Human Figure Drawing Continuum versus the Self-Portrait-Woman Sub-Test



The results are discussed in response to the research questions, in Chapter Five.

Chapter Five

Discussion

Overview

This research investigated young children's human figure drawings through the Rasch analysis of responses to the Goodenough-Harris Drawing Test (GHDT). The GHDT was administered over three data collection phases to a sample of 107 school children in prep to year five. A total of 246 drawings were collected, examined, and scored using the appropriate scoring guide. The drawing data were converted into a format suitable for Rasch analysis with the 246 drawings producing a data file which contained some 53,000 data points. The results described in Chapter Four are discussed in relation to the individual research questions below.

Responses to the Research Questions

Question One

The first of the two principal research questions was: Is it possible to Rasch analyse each of the four sub-tests (DAM, DAW, SPM and SPW) of the Goodenough-Harris Drawing Test to produce Rasch measurement scales?

The findings presented in the separate DAM, DAW, SPM and SPW sections of Chapter Four reveal unequivocally that it was possible to apply the Rasch model to the data from the sub-tests of the GHDT and produce Rasch measurement scales: the variable maps in Figures 4.1 to 4.3 (DAM subtest), 4.8 to 4.10 (DAW sub-test), 4.15 to 4.16 (SPM sub-test), and 4.22 to 4.24 (SPW sub-test) – together with the variable maps produced by the Rasch analyses of the stacked data sets in Figures 4.7, 4.14, 4.21 and 4.28 and their accompanying results tables – are evidence of the successful application of Rasch model analysis to the GHDT.

Question Two

The second of the two principle research questions was: To what extent do the generated measurement scales fit the Rasch model's unidimensionality requirement, and its other expectations?

The summary, item and person fit statistics produced by the separate Rasch analyses of the DAM, DAW, SPM and SPW sub-tests reveal fit statistics sufficient to indicate unidimensionality. Indeed, very few items and persons yielded fit statistics that fell outside of the accepted ranges of -2 to +2 for *t* statistics, and +0.75 and +1.3 for Mean Squares. None of the person or item reliability indices were below .84, with most being in the mid to high .90s.

The output produced from the separate Rasch analyses of the sub-tests had already indicated unidimensionality, and the analyses of stacked data for each sub-test further substantiated this. By stacking the data collected from each of the three phases for each sub-test, the data set was nearly tripled. Rasch analysis of the large data set was a more rigorous test of unidimensionality (Bond & Fox, 2007). Relatively few items yielded fit statistics that indicated their misfit to the latent construct under investigation. Furthermore, most items from the DAM/SPM and DAW/SPW sub-tests yielded difficulty estimates which were well-dispersed locations along the logit scale. As Linacre (2009b) espoused, items need to be similar enough to investigate a single latent construct, yet be as different as possible in order to inform researchers about that construct: the DAM/SPM and DAW/SPW items adhered quite closely to that principle.

Question Three

Given that the two principal research questions were answered in the positive, a number of subsequent research questions arose. The first of these was: To what extent do the generated scales adequately summarize/quantify the human figure drawings made by young children in the study, and to what extent do these scales differ from the original hierarchy?

The measurement scales displayed in the twelve variable maps produced by the Rasch analysis of the DAM, DAW, SPM and SPW sub-test of the GHDT satisfactorily quantified the young children's human figure drawings. This Rasch output has provided quantitative verification of what had been suggested qualitatively: that the concepts included in a drawing of a human figure contribute meaningfully to the investigation of the latent construct of intellectual maturity; and that there is a continuum of concepts, or attributes, that can help to differentiate the responses of less-developed children from those provided by more-developed children. Moreover, these variable maps provide much more information than can usually be offered qualitatively as they show not only who is more able (or developed) than whom, or which item is more difficult than another, but also by *how much* – as revealed by their locations on the equal-interval measurement continuum.

The variable maps presented in Chapter Four show not only that the 73 DAM/SPM and 71 DAW/SPW items were well-dispersed along the logit measurement scales, but that the items common to both tests were located comparably. Most logit scales spanned an impressive 10 to 12 logits in length indicating that the scales summarized an expansive range of human figure drawings – from basic tadpoles to more detailed human figures. None of the children's drawings 'topped out' on any of the sub-tests, nor did any drawings receive zero credit.

Whilst the scales adequately described the young children's human figure drawings, there were some discrepancies from the posited item hierarchy. Goodenough (1926a) and Harris (1963) both stated that, generally speaking, the items were ordered in the scoring guide from least to most difficult to include in drawings of men and women. However, the Rasch variable maps for each of the sub-tests revealed some variations in item hierarchies compared to those in the scoring guides. Review of the variable maps in Chapter Four reveals that DAM/DAW items 1 (head present), 4 (eyes present), 11/13 (mouth present), 18/19 (hair I), 30/24 (arms present), 35/33 (legs present) and 46/55 (trunk present) were continually reported as some of the least difficult items to include in human figure drawings. Most drawing theorists would assert that the attributes credited in the above-described items are amongst the first features of humans that young children draw (Cox, 1992, 1993, 1997; Kellogg, 1967, 1970; Luquet, 1923, 1930; Piaget & Inhelder, 1956, 1971). Therefore, Rasch model output from the analysis of the GHDT has provided empirical quantitative evidence which supports this long-standing qualitative claim; although, the item numbers attributed to some of these attributes – for example, item 30/24, 35/33 and 46/55 – indicate that Goodenough and Harris considered them to be more difficult for children to include in their drawings.

Also, DAM/DAW items 2 (neck present), 3 (neck present two dimensions), 6 (eye detail: pupil), and 7 (eye detail proportion) all yielded difficulty estimates that indicated that these items were more challenging than their GHDT item numbers suggest. Whilst Goodenough and Harris might have considered these items to be less challenging, the literature on young children's drawings does not refer directly to the relative complexity of drawing these particular human attributes.

Features of some of the DAM items numbered between 60 and 73 and DAW items 60 and 71 were rarely or not at all included in the drawings of children in this research. Most of these items were included by Harris in an attempt to raise the ceiling of the test so that children aged over eleven or twelve years of age could be administered the test. Harris's own research (1977), and that of others (Scott, 1981), revealed that children over the age of about ten or twelve years cease to show significant growth on the GHDT. This study adhered to the principle espoused earlier by Goodenough and, therefore, included children within the recommended age range (of around four to 10 years) only. None of the drawings by children in this research showed evidence of a ceiling effect. Despite Rasch analysis detecting a large number of maximum extreme items in each of the sub-tests, the logit scales produced still managed to describe satisfactorily, and to differentiate between, the young children's human figure drawings.

Question Four

An important aspect of this research was the investigation of Harris's extension of the original Goodenough Draw-a-Man Test to the four sub-test Goodenough-Harris Drawing Test. Harris's extension effectively tripled the data collection load on young children as it required three drawings instead of one; furthermore, this extension was recommended in the absence of any empirical evidence suggesting that a single drawing of a man was actually insufficient for the task of inferring children's levels of intellectual development. Therefore, research question four was: What additional information beyond that provided by the DAM sub-test, if any, is revealed by the DAW and SPM/SPW sub-tests?

The results presented in Chapter Four show that little additional information is revealed by the DAW, and SP sub-tests over that already produced by the DAM sub-test. The person ability estimates plotted in the variable maps (see Figures 4.1 to 4.3 DAM subtest; 4.8 to 4.10 DAW sub-test; 4.15 to 4.16 SPM sub-test; and 4.22 to 4.24 SPW sub-test) and stated statistically in the person fit statistics tables (see Tables 4.7 to 4.9 DAM; 4.17 to 4.19 DAW; 4.27 to 4.29 SPM; and 4.37 to 4.39 SPW) show that most children performed almost identically on all GHDT sub-tests. This finding is further substantiated visually in the common linking graphs displayed in Figures 4.31 (DAM versus DAW), 4.32 (50 common item DAM versus DAW), 4.33 (DAM versus SPM) and 4.34 (DAW versus SPW). Each of these common linking graphs reveals that very few children had drawing performances which varied beyond measurement error across the sub-tests.

The output produced from the Rasch analysis of the SPM and SPW data indicated that, in general, children received less credit for their self-portraits than they did for their drawings of men and women, despite the SP sub-test being better targeted to the abilities of this sample. This could be linked to the fact that the DAM and DAW scoring guides which are used to evaluate the children's self-portraits were designed specifically to judge drawings of adults. They were not modified to score drawings of children; surely some would hold that certain criteria are quite inappropriate for the evaluation of young children's self-portraits (e.g., DAW items 15 cosmetic lips, 67 directed lines and form: breast, 68 directed lines and form: hip contour).

Indeed, these findings support Goodenough's (1926a) original contention that a single drawing of a man could be adequate for inferring young children's intellectual development levels. Whilst some might posit that the collection of three analogous human figure drawings from children could provide more comprehensive evidence of their levels of intellectual development – in that the second and third drawings further validate the first – the redundancy of the other sub-test results and the usual practice that high-stakes judgements about young children are based on a range of information and assessments generally negate the need for the tripled drawing data collection load.

Question Five

Investigating how the children's human figure drawings changed over the three data collection phases was important for two reasons: it enabled the researcher to examine how well the measurement scales produced by each sub-test quantified the children's drawings in each phase, and it also offered insight into which sub-test was most effective in revealing drawing development over time. Therefore, question five was: How do the children's GHDT drawings change over time?

Figures 4.4 to 4.6, 4.11 to 4.13, 4.18 to 4.20, and 4.25 to 4.27 each display examples of the least successful, mean and most successful drawings for the DAM, DAW, SPM and SPW sub-tests respectively. Review of each of these figures reveals that the children's human figure drawings develop from the early tadpoles, which consist of "not what the child actually sees", but more so the essentials of the human form for that child (Piaget & Inhelder, 1956, p. 50), through to the more distinguishable and sophisticated human figure drawings which "endeavour to take perspective, proportions and distance into account all at once" (Piaget & Inhelder, 1956, p. 52). The examples of the children's drawings reveal that regardless of whether children drew a man, a woman, or themselves, a similar drawing developmental progression was evident. The drawings showcased in Chapter Four, together with the indication of test unidimensionality in the Rasch model output, are empirical evidence of the developmental sequence of young children's human figure drawings.

The graphs displayed in Figures 4.29 and 4.30 revealed how the children's drawings developed, on average, for the year level cohorts. Figure 4.29 showed that the average drawing development of the children's DAM drawings was progressive, and minimally variant, over the three phases of data collection. Whilst there were small variations in the amount of average growth across the year levels – for example children in year one experienced around 1.5 logits of growth on average, whilst children in years three and four experienced less than .5 logits of growth. Piaget's (1956, 1971) theory of cognitive development suggests that children whose thinking is in transition between the pre-operational and concrete operational stages of thought (which typically occurs around the ages of about 5 to 7 years, or year one) might experience significant intellectual development as their mental structures become more organized and operational. It is likely that this growth would be accompanied by more detailed human figure drawings. As the concrete operational mental structures become consolidated and more organized, however, children's drawings might show less significant annual development; the measurably smaller progression of the year three and four children's human figure drawings could be attributed to this.

The graph of the DAW drawing development in Figure 4.30, however, showed considerably more variation than did the DAM drawings in Figure 4.29. Again, the drawings made by children in year one showed the most development – around 1.5

logits. On the other hand, the drawings made by children in year two seemed to regress slightly in phase two; review of the DAW person ability estimates in Tables 4.17, 4.18 and 4.19 reveals that a number of these children's drawings yielded ability estimates that were lower in phase two than in phase one. This is particularly interesting given that the exact same number of children were available for phase two data collection as for phase one data collection; therefore, mere sample attrition could not be the cause of this result.

However, some key drawing theorists offer a practical explanation. Luquet (1913), Piaget (1956, 1971) and Di Leo (1970, 1973) assert that drawing development is not a rigid, abrupt nor a completely orderly process, but rather that children's transitions include periods of regression and progression before they finally accomplish the next drawing style or system. Whilst the variation of development in the year two children's drawing development over the three phases might be somewhat expected given the explanation above; the overall variation revealed by Figure 4.30 suggests that the DAW sub-test is less well-suited to revealing drawing development than is the DAM sub-test. Indeed, Goodenough's (1926a) original selection of a man as the subject of her drawing instrument was due to her contention that drawings of women, in particular, were less useful for assessment purposes as women had a much more varied appearance than did men. Further research including the draft HFDC and drawings of women would be helpful in further examining this claim.

Drawing development graphs were not produced for the SPM and SPW sub-tests as the sample sizes were too small to produce graphs yielding meaningful results. In some instances, for example, only one child was representative of an entire year level. Therefore, an investigation into which of the four GHDT sub-tests was most effective in revealing the development of young children's human figure drawings could not completed at this time. However, the results presented in Chapter Four suggest that the DAM sub-test could be considered to be more apt than the DAW sub-test in revealing the development of these children's human figure drawings.

Question Six

The original Goodenough Draw-a-Man Test (GDAMT) was created in 1926, and Harris's revision and extension culminated in the Goodenough-Harris Drawing Test (GHDT) in 1963. In the 43 years since Harris's revision and extension of the drawing test to this research which began in 2006, there has been considerable change in societal, cultural and educational expectations of young children. Therefore, it was necessary to ask: What modifications might update the GHDT to better align it with expectations of current educational users?

One of the most prominent aspects of the GHDT that required updating, to this researcher, included the somewhat sexist elements of the test booklet and scoring guide. Review of the test booklet in Appendix A shows that the cover page calls for the occupations of children's fathers amongst more usual requirements such as name, age, grade and school. The draft test booklet produced as a by-product of this research (see Appendix E) does not include this question for the simple reason that it does not relate to any substantive aspect of the test. For example, the original GDAMT and GHDT scoring guides do not require children to be grouped or compared according to their fathers' occupations, nor is this information conducive to evaluating objectively young children's human figure drawings. On the other hand, if it was deemed necessary to ask about children's parents or caregivers in order to produce a more detailed evaluation of the child, then current societal and cultural expectations would suggest that a test booklet require information about the male and/or female parents or caregivers of the child. Given that the GDAMT, GHDT, and the HFDC are all assessments which should be used as part of a suite of information when making judgements about young children, information about children's parents or caregivers occupations is hardly relevant to evaluating drawing development using the fore mentioned assessments.

The results also suggested that the scoring guides required updating as it was revealed that the DAW and DAW scoring guides were less sensitive, or less effective, in evaluating the children's self-portraits. This was evidenced by the fact that the mean person measures from the DAM and DAW sub-tests indicated that these tests were less well-targeted to the sample (DAM: -1.30, -1.21, and -1.14; DAW: -1.37, -1.25 and -.94), than were self-portraits, and children generally received more credit for drawings of adults (person mean raw score DAM: 22.1, 21.2 and 22.5; DAW: 22.6, 22.4 and 24.6). On the other hand, the SPM and SPW sub-tests were reported to be slightly more well-suited to the sample (mean person measures SPM: -1.14, -1.00 and -1.08; SPW: -.83, -1.15, -.79), although, children generally received less credit for these drawings (person mean raw score: SPM: 18.4, 14.9 and 16.7; SPW: 25.0, 22.1, and 23.6). One might assume that children would receive more credit on assessments that are better suited to them; however, Rasch analysis output indicates otherwise in regards to the GHDT sub-tests. Therefore, the DAM and DAW scoring guides were deemed to be less-appropriate in evaluating effectively the self-portrait drawings made by children in this sample. Thus, another key modification should be the construction of a scoring guide that is verified as more sensitive in evaluating the drawings of children. The draft HFDC scoring guide contains some modified criteria in an effort to be more sensitive to the features of drawings of children.

Furthermore, the DAW scoring guide contained a number of items that could be considered sexist, unsuitable or inappropriate given current societal, cultural and educational expectations. Items such as 23 (necklace or earrings), 38 (shoe ‘feminine’), 48 (waist I), 49 (waist II), 50 (skirt ‘modeled’ to indicate pleats or draping), 52 (garb feminine), 54 (garb a definite ‘type’ house dress/apron), 67 (directed lines and form: breast), and 68 (directed lines and form: hip contour) suggest that all females (regardless of age as the DAW scoring guide is used to score young girls self-portraits as well) should be drawn with jewellery, high heels, skirts, dresses and aprons, and include breasts, waists and hips. Indeed, a young girl’s self-portrait might, quite rightfully, exclude a number of items as they are not relevant to her developmentally (i.e. breasts, hips, waist); however, the evaluation of this drawing using the DAW scoring guide is unlikely to produce a raw score, or resultant ability estimate, that truly reflects that child’s level of cognitive development as the scoring guide assumes that these items are ‘absent’ due to a lack of intellectual maturity rather than inappropriateness due to the drawer’s level of physical development. The researcher intentionally avoided stereotypical, sexist and age-inappropriate items in the production of the more general HFDC scoring guide (see Appendix E). In alignment with the philosophy of early childhood education, a key modification could be the freedom of choice for young children to select which type of human figure drawing they would like to complete for evaluative purposes. Given that the results of this research indicate that most children performed comparably on each of the sub-tests anyway, such a modification would involve minimal, if any, limitations on the results. Furthermore, the idea of facilitating children to choose which type of human figure drawing they would like to make (i.e., a man, woman, boy or girl), rather than dictating that children make three drawings of specific types, better aligns with the ideals of modern early childhood education which is flexible, play-based and child-centred (Creche and Kindergarten Association, 2008; Hendrick, 2003; Queensland Studies Authority, 2006).

Another important modification for this researcher, was the suggestion of a new name for the evaluation of children's drawings. This researcher proposed the 'Human Figure Drawing Continuum' (HFDC). The HFDC title indicates: that it is concerned with drawings of the human figure, and it implies that there is no right or wrong answer, as opposed to 'test' which is often concerned with right and wrong answers and often carries some negative connotations. The name suggests that children's drawing proficiency might move along the continuum, even in both directions. The idea of a drawing continuum better aligns with a Piagetian theoretical perspective, which acknowledges that children produce both progressive and regressive drawings throughout their development. Regarding the development of children's drawings, Piaget and Inhelder (1956) state that:

> ... one should not be surprised by the co-existence of two situations during development which might seem contradictory but for the general reorganization precipitated by the completion of elementary topological operations ... an operation can function implicitly before it has become completely formal and explicit ... [therefore] ... one should not be surprised to find that in the actual development of the individual the true order in which the operational mechanisms appear is sometimes confused and overlapped ... this can be brought about by the conflicts which beset growing consciousness and abstract thinking, assisting it in

> some contexts and hindering it in others. (Piaget & Inhelder, 1956, p. 485)

Furthermore, Rasch analysis of the drawings produced logits measurement scales – or continua – along which the person ability and item difficulty estimates were plotted. Therefore, the idea of a drawing 'continuum' is quite apt.

Question Seven

Question seven was: Is it possible to develop a Human Figure Drawing Continuum that includes some of the items of the current GHDT that could be useful for evaluating any human figure drawings made by young children?

The results presented in Chapter Four indicate that, yes, it was possible to develop a draft HFDC that includes items from the GHDT and is effective in evaluating human figure drawings made by the young Australian children in this sample.

The construction of a draft HFDC began with an examination of the fit statistics for each of the 50 GHDT common items. The 50 common DAM/DAW items were selected as a basis for the development of the HFDC as they yielded satisfactory Rasch fit statistics with only 50 items (see Chapter Four and Appendix D), and each of the items is applicable to drawings of both males and females. The fit statistics for each of these 50 items (in each of the sub-tests, that is) was carefully examined in order to determine which items could be removed in an effort to enhance the measurement properties of the proposed instrument. Various combinations of misfitting items were removed and repeated Rasch model analyses were conducted. Each iteration produced output that was inspected meticulously for adherence to the Rasch model's expectations. Whilst some items, such as DAM/DAW item 9 (nose present), were frequently detected as misfitting the Rasch model's expectations in the separate analyses of each sub-test, when removed from the instrument the person and item standard deviations, person separation indices, person and item reliability indices did not improve significantly; in fact, in some cases they became less satisfactory. Furthermore, the number of children with drawings detected as misfitting the Rasch model's expectations also increased with item 9 absent. This might have been linked to the fact that removal of item 9 affected other DAM/DAW items such as 10 (nose, two dimensions), 13/16 (both nose and lips in two dimensions), and 17/11 (bridge of nose); therefore, despite the frequent detection of item 9 as an overly erratic item, the Rasch output indicated that the HFDC instrument yielded more erratic results without it. Future research could include modification of item 9, to make a polytomous rather than a dichotomous item, together with application of the Rasch Partial Credit Model.

Finally, the particular combination of 45 items that best fit the expectations of the Rasch model was established (see Appendix E). The Rasch analysis results and common linking graphs presented in Chapter Four show that it was possible to propose a Human Figure Drawing Continuum (HFDC) that was verified by modern test theory as effective in the evaluation of young children's self-selected human figure drawings of men, women and children. The variable map in Figure 4.35 together with the summary and item statistics presented in Tables 4.41 and 4.42 respectively, reveal that the 45 item HFDC was almost exactly as effective in describing the children's drawings as any of the 70-plus item sub-tests of the GHDT. In fact, it could be argued that it was the most well-suited and reliable instrument. The HFDC person separation was 2.89, which was comparable with that produced by the SPM and SPW sub-tests (SPM = 2.33 to 2.98; SPW = 2.92 to 3.50), and was only slightly less than that of the DAM and DAW sub-tests (DAM = 3.15 to 3.37; DAW = 3.22 to 3.40). The HFDC yielded a mean person measure of -.59 which was the highest of all of the sub-tests (DAM: -1.30, -1.21 and -1.14; DAW: -1.37, -1.25 and -.94; SPM: -1.14, -1.00 and -1.08; SPW: -.83, -1.15, and -.79) indicating that it was the best-targeted test to this sample. The HFDC item reliability of .99 was also the highest of all the sub-test's item reliability indexes (DAM: .97, .96 and .94; DAW: .97, .96 and .95; SPM: .94, .90 and .84; SPW: .96, .94 and .92). Therefore, considering the four GHDT sub-tests and the proposed HFDC, it was the HFDC's item hierarchy which was most likely to be replicated when the test is administered to other suitable samples. And whilst the HFDC yielded a person reliability index of .89, this was only a slight decrease from those yielded by the DAM and DAW sub-tests (which ranged from low to mid .90s) and still above the lowest of .84 for the SPM sub-test.

Ideally, a scale featuring 50 to around 60 items would be more apt for the task of inferring young children's levels of intellectual development through the detail of their human figure drawings. Whilst more work is required in order to produce a series of around 50 to 60 items that adequately describe young children's human figure drawings, and fit the expectations of the Rasch model, the draft 45 item HFDC is a step closer toward that goal.

Implications and Recommendations

This research, which applied the Rasch model to the Goodenough-Harris Drawing Test (GHDT), has several implications for young children, teachers, psychologists and other interested parties, as well as for the body of knowledge related to young children's human figure drawings.

Young children's human figure drawings are a key developmental element of early childhood that should be nurtured. Whilst the statistics presented in Chapter Four can reveal much about the psychometric properties of the GHDT and can plot the development of the children's drawings, the statistics can not show the importance of the drawing process to young children. Young children do love to draw and their drawings are often both personal and significant to them. Drawing is not merely a fun activity, 'time-filler', or meaningless play; drawing can be considered an important developmental activity in itself, as well as a tool for future development.

Teachers, psychologists, parents, and other interested parties should offer multiple and varied opportunities for children to draw, as well as to share and discuss their drawings. Adults can infer much from young children's drawings, however, even more can be revealed by listening to the young child explaining the elements of their representation. For this researcher, asking children to talk about their human figure drawings was often worthwhile. What could otherwise be interpreted as a smudge or 'slip of the hand' was sometimes a belly button, pocket, neck tie, or necklace. A seemingly meaningless scribble beside a human figure was actually a bag, briefcase, pet, or plant. And incomprehensible additional markings on, or inside, some human figures were, in fact, self-portraits of children at their desks, babies inside their mothers, and hearts, lungs or stomachs. Indeed, some children's raw scores actually increased due to the researcher carefully questioning the children about their drawings. Therefore, for adults to be sensitive to children's drawings they must be prepared to delve beneath the surface features. Focus should be on the concepts and detail included in the drawings, rather than neatness, creativity, realism, or conformity to particular styles. Of course, with drawing being a key developmental feature of early childhood, adults should remain sensitive to the natural progression and regression in children's drawings. Similar to progress in other developmental domains, the development of drawing is not rigid, abrupt, nor sequentially invariant.

Young children's human figure drawings could be collected and examined profitably, approximately every six months in educational settings. This research has shown that some remarkable development occurs in young children's human figure drawings – even within such relatively short periods of time. Early childhood professionals could benefit from collecting human figure drawings from their students, approximately every six months. These drawings could be evaluated using the 45 item HFDC scoring criteria, and the Draft Number to Measure Conversion Table (in Appendix E) could be used to infer children's ability levels. These drawings, together with notes about what the child said or clarified in relation to the drawing and – perhaps – the interpolated drawing measure could be filed in individual portfolios together with other information which could then inform judgements about the child's development and achievement.

There are limitations to the GHDT which should be taken into consideration by users. The purpose of this research was to produce insights that might address some of the limitations of the GHDT in order to better align the instrument with today's societal, cultural and educational expectations. Indeed, the GHDT was, and continues to be, one of the most effective and widely-used non-verbal assessments of young children's intellectual development. However, users should be aware of its limitations; in particular, possible sub-test redundancy, the lack of sensitivity of the scoring guides to features of children's self-portraits, in particular, and the fact that the instrument yields ordinal-level counts rather than equal-interval measures.

The proposed HFDC mitigates many of the limitations of the GHDT, however, more research is needed. While the HFDC would offer children the freedom to self-select the type of human figure drawing they would like to make, it currently lacks the wealth of empirical support of the GHDT which has been used with thousands of children from diverse backgrounds. Although the HFDC scoring guide can be applied to this sample of young Australian children's human figure drawings of any type (men, women, boys and girls) and inferences about these children's abilities can be made by applying the Rasch model, the effectiveness of this method for other suitable samples is yet to be verified. For this researcher, however, perhaps the single most meaningful outcome of this research is the construction of an instrument which is better aligned with common values of early childhood education.

Limitations and Suggestions for Future Research

Like all research, this study has its own limitations which are discussed, and linked to recommendations for future research, below.

Foremost, this research should be replicated to investigate whether similar results can be achieved. Replication would provide further verification of the effectiveness, and limitations, of the both the GHDT and the draft HFDC.

This research involved only a very small sample size, and as the focus was on whether GHDT data were apt for Rasch analysis, the number of demographic and other variables related to the sample was quite limited. Since this research has indicated that the GHDT is suitable for Rasch analysis, future research should include not only a larger sample size, but also cover additional variables including children from diverse backgrounds (i.e., other cultures and countries) and children with special needs.

The longitudinal aspect of this study spanned only twelve months. Future research should track the development of young children's drawings over several years. Ideally, a project could follow the development of drawing from the very first marks produced at around 18 months of age, through to the product of an adolescent in order to thoroughly investigate the possible ceiling effect of human figure drawings.

This research investigated only drawings of the human figure. Future research could investigate what associations might exist between human figure drawings and other assessments such as Piagetian developmental tasks, IQ tests, and school achievement. Given that young children draw many different subjects, not only drawings of the human figure, future research might examine other subject matter of young children's drawings.

Whilst this research has produced a practical by-product in the draft of a proposed HFDC, this instrument is not a standardised, stand-alone assessment. The original ideas on which the HFDC is based are copyright by Goodenough (1926a), Harris (1963) and the Harcourt Brace Psychological Corporation. Given the resolution of any copyright issues, use of the draft HFDC must be part of a suite of information upon which judgements about young children are made; teachers, psychologists, and other interested parties must be aware of the risks of basing important decisions on the result of any single assessment. The researcher is adamant about not repeating the mistakes of the past whereby young children were underestimated or condemned as 'incompetent' because of a single less-than-favourable drawing. As Goodenough espoused in her book, *Mental Testing* (1949), "the use of ... [such] ... methods by many enthusiastic but poorly trained persons has not always worked out to the advantage of the tested individuals or of society" (p. vii).

Conclusion

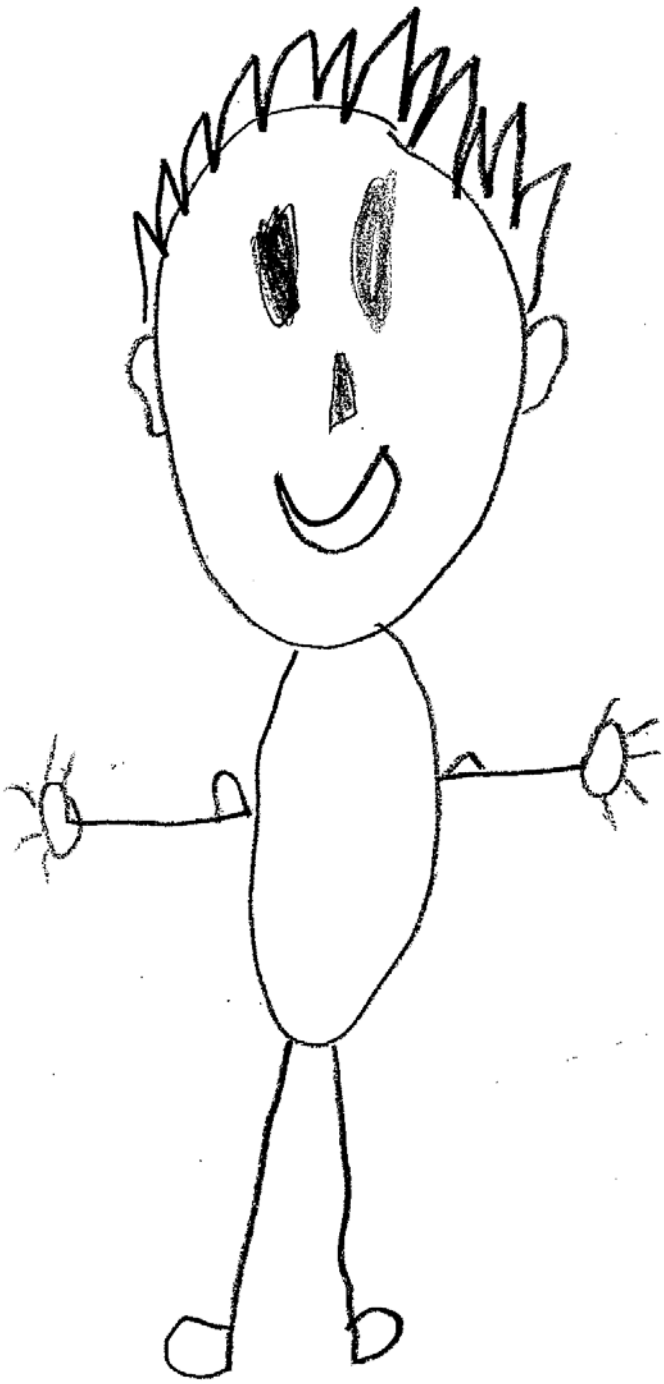
The results of this study have offered further verification of the premise that "drawings provide rich insights into young children's thinking" (Anning & Ring, 2004, p. x). Through the examination of the world's most famous and widely used drawing instrument from a latent trait theoretical perspective, it is intended that the significance of young children's human figure drawings might be re-acknowledged by teachers, psychologists, parents and others.

Whilst this research has involved an assessment instrument, the Rasch model, and statistics, it is fundamentally about young children, their human figure drawings, and the insights that these can reveal. After all, the researcher is an early childhood professional, not a psychometrician. The application of the Rasch model was to do with this researcher's interest in better understanding young children, not due to a fondness for mathematical formulae. Crucially: at a time when young children have not yet mastered speaking, reading or writing to a degree where they can utilise these forms of communication to represent their understandings, or levels of intellectual development, young children can – and do – draw.

> Children draw what they know rather than what they see. The truth of this oft-quoted statement has been recognised for decades. The little child does not care whether or not his pictures are beautiful, but he wants them to tell what he has in mind. (Goodenough & Tyler, 1959, p. 316)

Figure 5.1

A Self-Portrait Made by a Male aged 6 years



Note. After the boy had completed all three drawings, the researcher asked him to tell her about each of them. Looking at his self-portrait the boy explained, “yeah, I got them muscles ... you can’t really see ‘em here ... [points to his biceps] .... but I put ‘em in here ... [points to drawing]...”.

References

Andrich, D. A. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561-573.

Alland, A. (1983). *Playing with Form*. New York: Columbia University Press.

Anning, A. & Ring, K. (2004). *Making sense of children's drawings*. Berkshire, England: Open University Press.

Annual Report. (1939). State Colony for Epileptics and Feeble-Minded. Richmond.

Ashton, L. E. (1999). Drawing subjects: A poststructuralist study of generalist primary teacher and pre-service teacher discourse in art education. Unpublished Doctor of Philosophy Thesis. Townsville: James Cook University of North Queensland.

Ballard, P. B. (1913). What London children like to draw. *Journal of Exp. Ped*. 1, 185-197.

Baldwin, J. M. (1894). Mental development in the child and in the race. New York: Macmillan.

Baldwin, A. L. (1980). *Theories of child development* (2nd ed.). New York: John Wiley & Sons.

Barnes, E. (1891). Notes on children's drawings. *Pedagogical seminary, 1,* (445-447).

Barnes, E. (1893). A study of children's drawings. *Pedagogical Seminary, 2,* 451-463.

Barnes, E. (1894). The art of little children. *Pedagogical Seminary, 3*, pp. 302-147.

Barrett, M. D., Beaumont, A. V., & Jennett, M. S. (1985). Some children do sometimes do what they have been told to do: task demands and verbal instructions in children's drawing. In N. H. Freeman & M. V. Cox (Eds.), *Visual order: the nature and development of pictorial representation* (pp. 135-154). Cambridge: Cambridge University Press.

Barrett, M. D. & Bridson, A. (1983). The effect of instructions upon children's drawings. *British Journal of Developmental Psychology, 1*, 175-178.

Barrett, M. D. & Light, P. H. (1976). Symbolism and intellectual realism in children's drawings. *British Journal of Educational Psychology, 46*, 198-202.

Bentzen, W. R. (2000). Seeing young children: A guide to observing and recording behaviour (4^{th} ed.). New York: Thompson Learning.

Bentzen, W. R. (2009). Seeing young children: A guide to observing and recording behaviour (6th ed.). New York: Thompson Learning.

Berk, L. E. (2005). *Infants and children* (5th ed.). Boston: Allyn & Bacon.

Berrien, F. K. (1935). A study of the drawings of abnormal children. *Journal of Educational Psychology, XXVI*, 143-150.

Bidell, T. R. & Fisher, K. W. (1992). Beyond the stage debate: Action, structure, and variability in Piagetian theory and research. In R. J. Sternberg & C. A. Berg (Eds.) *Intellectual Development*. New York: Cambridge University Press.

Blain, G. J., Bergner, R. M., Lewis, M. L. & Goldstein, M. A. (1981). The use of objectively scorable House-Tree-Person indicators to establish child abuse. *Journal of Clinical Psychology*, *37* (3), 667-673.

Bond, T.G. (2001). Ready for School? Ready for learning? An empirical contribution to a perennial debate. *The Australian Educational and Developmental Psychologist, 18*(1), 77-80.

Bond, T. G. & Bunting, E. M. (1995). Piaget and measurement III: Reassessing the *méthode clinique*. *Archives de Psychologie, 63*, 231-255.

Bond, T. G. & Fox, C. M. (2001). *Applying the Rasch model: Fundmental measurement in the human sciences.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch model: Fundmental measurement in the human sciences* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Bond, T. G. & Parkinson, K. (1996). *Qualitative analysis of the méthode clinique II: The child's conception of area.* Poster presented at the Annual Symposium of the Jean Piaget Society, June.

Brill, M. (1935). The reliability of the Goodenough Draw A Man test and the validity and reliability of an abbreviated scoring method. *Journal of Educational Psychology, 26*(9), 701-708.

Brill, M. (1937). A study of instability using the Goodenough drawing scale. *Journal of Abnormal Psychology, 32*, 288-302.

Brooks, M. (2002). *Drawing to Learn*. Unpublished Ph.D Thesis. University of Alberta, Canada.

Brooks, M. (2003). Drawing to learn [Electronic Version]. *Beyond the Journal: Young Children on the Web.*

Brooks, M. (2004). Drawing: The social construction of knowledge. *Australian Journal of Early Childhood, 29* (2), 41-49.

Brooks, M. (2005). Drawing as a unique mental development tool for young children: interpersonal and intrapersonal dialogues. *Contemporary Issues in Early Childhood, 6*, 80-91.

Brown, D. D. (1897). *Notes on children's drawings*. California: University of California Publications.

Bruner, J. (1964). The course of cognitive growth. *American Psychologist, 19*, 1-15.

Buck, J. N. (1948a). The H-T-P test. *Journal of Clinical Psychology*, *4*, 151-159.

Buck J. N. (1948b). The H-T-P technique; a qualitative and quantitative scoring manual. *Journal of Clinical Psychology, 4*, 317-396.

Bunting, E. (1993). A *qualitative and quantitative analysis of Piaget's control of variables scheme.* Unpublished Honours Thesis. James Cook University, Australia.

Burt, C. (1921). *Mental and scholastic tests.* London: P. S. King & Son.

Carroll, J. B. (1993). Human cognitive abilities: A survey of factor-analytic studies. London: Cambridge University Press.

Chien, C. W. (2007). Using the Rasch model to validate the Peabody Developmental Motor Scales-Second Edition in infants and preschool children. Unpublished Masters Thesis. James Cook University, Australia.

Claparède, E. (1907). Plan d'experiences collectives sur le dessin des enfants. (Plan for collective experiments on the drawing of children.) *Arch. De Psychol. 6*, 276-278.

Clark, A. B. (1902). The child's attitude toward perspective problems. In E. Barnes, (ed.). *Studies in Education, 1*, (2nd ed.). Standford: University Press.

Collins English Dictionary. (1998). Glasgow: HarperCollins Publishers.

Cooke, E. (1885). Art teaching and child nature. *London Journal of Education*, 10.

Cox, M. (1992). *Children's Drawings*. London: Penguin Group.

Cox, M. V. (1993). *Children's drawings of the human figure.* Mahwah, New Jersey: Lawrence Earlbaum Associates.

Cox, M. (1997). *Drawings of people by the under-5s.* London: Falmer Press.

Cox, M. (2005). *The pictorial world of the child*. London: Cambridge University Press.

Cox, M. V. & Parkin, C. (1986). Young children's human figure drawing: cross-sectional and longitudinal studies. *Educational Psychology, 6*, 353-368.

Creche and Kindergarten Association. (2008). *Building Waterfalls: A living and learning curriculum framework.* Newmarket, QLD: Creche and Kindergarten Association.

Crook, C. (1985). Knowledge and appearance. In N. H. Freeman & M. V. Cox (Eds.), *Visual order: the nature and development of pictorial representation* (pp. 248-265). Cambridge: Cambridge University Press.

Darwin, C. (1877). A biological sketch of an infant. *Mind, 2,* 285-294.

Davis, A. M. (1985). The canonical bias: young children's drawings of familiar objects. In N. H. Freeman & M. V. Cox (Eds.), *Visual order: the nature and development of pictorial representation* (pp. 202-213). Cambridge: Cambridge University Press.

Denison, D. (1978). *Education under six*. London: Croom Helm.

Di Leo, J. H. (1970). *Young children and their drawings*. New York: Brunner/Mazel, Inc.

Di Leo, J. H. (1973). *Children's drawings as diagnostic aids.* New York: Brunner/Mazel, Inc.

Drake, C. (1998). Judgements versus justifications in Piaget's control of variables scheme. Unpublished Thesis. James Cook University, Australia.

Dunn, J. A. (1972). Goodenough-Harris drawing test. In O. K. Buros (Ed.). *The Seventh Mental Measurements Yearbook,* (pp. 671-673). Lincoln, Nebraska: Buros Institute of Mental Measurement.

Dunn, M. B. (1954). Global evaluation of children's drawings of "person" and "self". Unpublished Doctoral Dissertation. Colombia: Colombia University.

Dyson, H. A. (1982). The emergence of visible language: Interrelationships between drawing and early writing. *Visible Language 6,* 360-381.

Dyson, H. A. (1993). From prop to mediator: The changing role of written language in children's symbolic repertoires. In B. Spodek & O. N. Saracho, *Year Book in Early Childhood Education: Language and Literacy in Early Childhood Education*, Volume 4, pp21-41. New York: Teachers College Press.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93,* 179-360.

Embretson, S. E. (Ed.). (1985). *Test design: developments in psychology and psychometrics*. Orlando, Florida: Academic Press, Inc.

Flavell, J.H., Miller, P.H & Miller, S.A. (2002). *Cognitive development* (4^{th} ed.). Upper Saddle River, N.J.: Prentice Hall.

Ford, M. A., Stern, D. N. & Dillon, D. J. (1974). Performance of children ages 3 and 5 on the Draw-a-Person Task: Sex differences. *Perceptual and Motor Skills, 38*, 1188.

Fowlkes, S. A. (1980). The effect of colour on human figure drawings as related to level of social adaptability. Unpublished Masters Thesis. The Ohio State University, United States of America.

Freeman, N. H. (1972). Process and product in children's drawings. *Perception,* 1, 123-140.

Freeman, N. H. (1980). *Strategies of representation in young children.* London: Academic Press.

Freeman, N. H. & Cox, M. (1985). *Visual order.* Cambridge: Cambridge University Press.

Furth, G. M. & Kubler-Ross, E. (1988). *The secret world of drawings: A Jungian approach to healing through art.* Toronto: University of Toronto Press Inc.

Gablik, S. (1976). *Progress in Art*. London: Thames and Hudson.

Gardner, H. (1980). Artful scribbles: The significance of children's drawings. New York: Basic Books.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Gardner, H. (1999). Are there additional intelligences? The case for naturalist, spiritual, and existential intelligences, In J. Kane (Ed), *Education, Information, and Transformation*, Prentice Hall, pp. 111-131.

Gaultier, T. (1856). *L'art moderne.* Paris: Michel Lévy Frères.

Ginsburg, H. & Opper, S. (1988). *Piaget's theory of intellectual development* (3^{rd} ed.). New Jersey: Prentice Hall.

Golomb, C. (1981). Representation and reality: the origins and determinants of young children's drawings. *Review of Research in Visual Arts Education,* 14, 36-48.

Golomb, C. (1992). *The child's creation of a pictorial world.* Berkeley, CA: University of California Press.

Goodenough, F. L. (1926a). *The measurement of intelligence by drawing*. New York: World Book Co.

Goodenough, F. L. (1926b). A new approach to the measurement of the intelligence of young children. *Pedagogical Seminary and Journal of Genetic Psychology*, 33.

Goodenough, F. L. (1928a). The Kuhlman-Binet tests for children of preschool age: A critical study and evaluation. Minneapolis: University of Minnesota Press.

Goodenough, F. L. (1928b). Measuring behaviour traits by means of repeated short samples. *Journal of Juvenile Research, 12,* 230-235.

Goodenough, F. L. (1949). *Mental testing: its history, principles and applications*. New York: Rinehart and Company, Inc.

Goodenough, F. L., Foster, J. C. & Van Wagenen, M. J. (1932). *The Minnesota preschool tests: Manual of instructions. Forms A and B.* Minneapolis: Educational Testing Bureau.

Goodenough, F. L. & Harris, D. B. (1950). Studies in the psychology of children's drawings: II. *Psychological Bulletin, 47*, 369-433.

Goodenough, F. L. & Harris, D. B. (1963) *Goodenough-Harris Drawing Test.* San Antonio, Texas: Harcourt Assessment, Inc.

Goodenough, F. L. & Tyler, L. E. (1959). Developmental psychology (3rd ed.). New York: Appleton-Century-Crofts Inc.

Goodnow, J. (1977). *Children's drawing*. London: Open Books

Götze, K. (1898). *Das kind als Könstler.* Hamburg.

Gould, S. J. (1996). The mismeasure of man: Revised and expanded edition. London: Penguin Books.

Gould, S. J. (1977). *Ontogeny and phylogeny.* Cambridge, MA: Harvard University Press.

Groth-Marnat, G. & Roberts, L. (1998). Human figure drawings and House-Tree-Person drawings as indicators of self-esteem: A quantitative approach. *Journal of Clinical Psychology 54*(2), 219-222.

Harris, D. B. (1963). Children's drawings as measures of intellectual maturity: A revision and extension of the Goodenough draw-a-man test. New York: Harcourt, Brace & World, Inc.

Harris, D. B. (1977). Goodenough-Harris test estimates of intellectual maturity of youths 12-17 years. *Vital and Health Statistics, 11 (159).*

Harris, D. B. & Pinder, G. D. (1974). The Goodenough-Harris Drawing Test as a measure of intellectual maturity of youth aged 12 to 17 years: Demographic and socioeconomic factors. *Vital and Health Statistics,* 116(*11*).

Harris, D. B. & Roberts, J. (1972). Intellectual maturity of children: Demographic and socioeconomic factors. *Vital and Health Statistics, 105(11)*.

Hendrick, J. (2003). *Total Learning: Developmental curriculum for the young child* 6th Ed.). Upper Saddle River, New Jersey: Merrill Prentice Hall.

Herrick, M. A. (1893). Children's drawings. *Pedagogical Seminary, 3,* 338-339.

Hogan, L. (1898). *A study of a child.* New York: Harper & Brothers.

Houston, S. D. (Ed.). (2004). *The first writing: Script invention as history and process*. Cambridge: Cambridge University Press.

Ingram, N. (1985). Three into two won't go: Symbolic and spatial coding processes in young children's drawings. In N. H. Freeman & M. V. Cox (Eds.), *Visual order: the nature and development of pictorial representation* (pp. 135-154). Cambridge: Cambridge University Press.

Inhelder, B. & de Caprona, D. (1987). Introduction. In B. Inhelder, D. de Caprona, & A. Cornu-Wells (Eds.) *Piaget Today*. London: Lawrence Erlbaum.

James Cook University Research Committee. (n.d.). *James Cook University Statement and Guidelines on Research Practice* (draft). Townsville, Queensland: Author.

Jolley, R. P., Fenn, K. & Jones, L. (in press). The development of chidren's expressive drawing. *British Journal of Developmental Psychology.*

Katzaroff, M. D. (1910). Qu'est ce que les enfants dessinent? (What do children draw?). *Arch. de Psychol.,* 9, 125.

Kellogg, R. (1967). *The psychology of children's art.* San Diego: Random House.

Kellogg, R. (1970). *Analyzing children's art.* Pal Alto, California: Mayfield.

Kellogg, R. & O'Dell, S. (1967). *The psychology of children's art*. Random House.

Kendrick, M. & McKay, R. (2002). Uncovering literacy narratives through children's drawings. *Canadian Society for the Study of Education, 27(*1), 45-60.

Kerschensteiner, D. G. (1905). *Die entwickelung der zeichnerischen begabung.* (The development of drawing talent). Munich: Gerber.

Klinkenborg, V. (2000). The best clock in the world … and why we can't live without it. *Discover, 21*(6), 50-57. In T. G. Bond & C. M. Fox (2001), *Applying the Rasch model: Fundamental measurement in the human sciences.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Koppitz, E. M. (1968), Psychological evaluation of children's human figure drawings. New York: Grune & Stratton.

Lambert E. B. (2006). Can drawing facilitate problem-solving? An exploratory study. *Australian Journal of Early Childhood, 31*(2), 42-47.

Lamprecht, K. (1906). Les dessins d'enfants comme source historique (Children's drawings as an historical source.) *Bull. De l'Academie Royale de Belgique, Classe des Lettres,* 9-10, 457-469.

Lantz, B. (1955). *Easel age scale*. Los Angeles, California: Test Bur.

Levinstein, S. (1905). Kinderzeichnunger bis zum 14 Lebensjahre. Mit Parallelen aus der Urgeschichte, Kulturgeschichte, and Völkerunde. (Children's drawings to age fourteen. With parallels to the earliest history, the history of civilization, and folklore.) Liepzig: Voigtländer.

Lewis, D. & Greene, J. (1983). *Your child's drawings ... Their hidden meaning*. London: Hutchinson & Co.

Light, P. (1985). The development of view-specific representation considered from a soci-cognitive standpoint. In N. H. Freeman & M. V. Cox (Eds.), *Visual order: the nature and development of pictorial representation* (pp. 135-154). Cambridge: Cambridge University Press.

Light, P. H. & Simmons, B. (1983). The effects of a communication task upon the representation of depth relationships in young children's drawings. *Journal of Experimental Child Psychology*, 35, 81-92.

Linacre, J. M. (1994). Chi-square fit statistics. *Rasch Measurement Transactions, 8*(2), 350.

Linacre, J. M. (1995). Learning from qualitative data analysis. *Rasch Measurement Transactions,* 9(1), 405-407.

Linacre, J. M. (2002). What do infit and outfit, Mean-square and standardized mean? *Rasch Measurement Transactions, 16*(2), 878.

Linacre, J. M. (2009a). *WINSTEPS®* Version 3.68.1 [computer software]. Chicago, IL: http://www.winsteps.com.

Linacre, J. M. (2009b). A user's guide to WINSTEPS® Ministeps Rasch model computer program: Program manual 3.68.0. Chicago, IL:

Lobsien, M. (1905). Kinderzeichnung und Kunstkanon. *Zeitshrift für pädagogische Psychologie,* 7, 393-404.

Lowenfeld, V. (1939). *The nature of creative activity.* New York: Macmillan.

Lowenfeld, V. (1957). *Creative and mental growth,* (3rd Ed.). New York: Macmillan.

Lowenfeld, V. (1975). *Your child and his art.* New York: Macmillan.

Lukens, H. (1896). A study of children's drawings in the early years. *Pedagogical Seminary, 4*, 79-110.

Luquet, G. H. (1913). *Les dessins d́un enfant*. Paris: Librairie Felix Alcan.

Luquet, G. H. (1923). *Le dessin enfantin*. Paris: Delachaux et Niestle.

Luquet, G. H. (1930). *The art and religion of fossil man* (J. Townsend Russell, Trans.). New Haven, CT: Yale University Press.

Machover, K. A. (1949). Personality projection in the drawing of the human figure: A method of personality investigation. Springfield: C. C. Thomas.

MacPhail, A & Kinchin, G. (2004). The use of drawings as an evaluative tool: students' experiences of sport education. *Physical Education & Sport Pedagogy*, 9(*1),* 87-108.

Maitland, L. M. (1895). What children draw to please themselves. *The Inland Educator, I,* pp. 77-81.

Maley, C. R. (2005). Examining the role of cognitive development as an indicator of school readiness: An empirical contribution to a perennial debate. Unpublished Honours Thesis. James Cook University, Australia.

Märtin, H. (1939). Die motivwahl und ihr wandel in der freien zeichnung des grundschulkindes. *Zeitschrift für Paedagogische Psychologie*, 40, pp. 231-241.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.

McCarthy, D. (1944). A study of the reliability of the Goodenough drawing test of intelligence. *Journal of Psychology, 18*, 201-206.

McCarty, S. A. (1924). *Children's drawings*. Baltimore: William and Wilkins.

McCurdy, H. G, (1947). Group and individual variability on the Goodenough Draw-a-man test. Journal of Educational Psychology, 38, 428-436.

McElwee, E. W. (1932). The reliability of the Goodenough intelligence test used with sub-normal children fourteen years of age. *Journal of applied Psychology, 16*, pp. 217-218.

McInerney, D. M. & McInerney, V. (2006). *Educational psychology: constructing learning*, (4th ed.). Frenchs Forest: Pearson Education.

McHugh, G. (1945a). Changes in Goodenough IQ at the public school kindergarten level. *Journal of Educational Psychology, 36*, 17-30.

McHugh, G. (1945b). Relationship between the Goodenough Drawing a Man test and the 1937 revision of the Stanford-Binet test. *Journal of Educational Psychology, 36*, 119-124.

Medicus, G. (1992). The inapplicability of the biogenetic rule to behavioural development. *Human Development 35(1)*, 1-8.

Miller, S. A. (2004). Piagetian Research. In L.Nadelman, *Research manual in child development,* (pp. 111-123). Mahwah, NJ: Lawrence Erlbaum.

Nadelman, L. (2004). *Research manual in child development.* Mahwah, NJ: Lawrence Erlbaum.

National Health and Medical Research Council, Australian Research Council, & Universities Australia. (2007). *Australian Code for the Responsible Conduct of Research*. Canberra, ACT: Australian Government.

O'Shea, M. V. (1897). Some aspects of drawing. *Educational Review, 14,* 263-284

Perez, M. B. (1888). L'Art et la poesie chez l'enfant. Paris.

Phatak, P. (1959). A study of the revised Goodenough scale with reference to artistic and non-artistic drawings. Unpublished, mimeo. India: University of Baroda.

Piaget, J. (1953). *The origin of intelligence in the child.* . London: Routledge and Kegan Paul Ltd.

Piaget, J. (1961). *The child's conception of number* (2nd ed.). London: Routledge and Kegan Paul Ltd.

Piaget, J. (1971). *Biology and knowledge.* Chicago: University of Chicago Press.

Piaget, J. & Inhelder, B. (1956). *The child's conception of space.* London: Routledge and Kegan Paul.

Piaget, J. & Inhelder, B. (1971). *Mental Imagery in the Child*. London: Routledge & Kegan Paul Ltd.

Piaget, J. & Inhelder, B. (1969). *The psychology of the child.* London: Routledge and Kegan Paul.

Pikunas, J. & Carberry, H. (1961). Standardisation of the graphoscopic scale: the content of children's drawings. *Journal of Clinical Psychology,* 17, 297-301.

Poortinga, Y. H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology, 24,* 737-756.

Preyer, W. (1899). *The mind of the child* (W. H. Brown trans). New York: Appleton & Co.

Queensland Studies Authority. (2006). *Early Years Curriculum Guidelines.* Spring Hill, QLD: Queensland Studies Authority.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Paedagogiske Institut.

Rasch, G. (1980). Probabilistic models for some intelligence and attainment test (Expanded ed.). Chicago: University of California Press.

Read, H. (1943). *Education through art*. London: Faber and Faber.

Ricci, C. (1887). Ĺarte dei bambini (art of children). (Translation by Maitland in *Pedagogical Seminary* (1894), 3, 302-307).

Ring, K. (2001). *Young children drawing: the significance of the context.* Paper presented at the British Educational Research Association Annual Conference, University of Leeds, 13-15 September 2001.

Rorschach, H. (1942). Psychodiagnostics: A diagnostic test based on perception (P. Lemkau & B. Kronenberg, Trans.). Berne, Switzerland: Hans Huber.

Rouma, G. (1913). Le langage graphique de l'enfant. (The child's graphic language). Paris: Misch. Et Thron.

Rousseau, J. J. (1976). *Emile* (Allan Bloom trans.). New York: Basic Books.

Rowe, F. B., Crews, W. D., & Finger, F. W. (1993). John N. Buck (1906-1983): Did he practically establish clinical psychology in Virginia? *Journal of Clinical Psychology, 49*(3), 428-434.

Schuyten, M. C. (1901). Het oorspronkelijk teekenen als bijdrage tot kinderanalyse. *Paedologisch Jaarboek*, *2,* 112-126.

Schuyten, M. C. (1904). De oorspronkelijke 'Ventjes' der antwerpsche schoolkindern. *Paedologische Jaarboek, 5*, 1-87.

Schuyten, M. C. (1907). Note pedagogique sur le Dessins des Enfants. *Archives de Psychologie,* 6, 389-391.

Scott, L. H. (1981). Measuring intelligence with the Goodenough-Harris Drawing Test. *Psychological Bulletin, 89(3),* pp. 483-505.

Segall, M. H., Dasen, P. R., Berry, J. W., & Poortinga, Y. H. (1999). Human behaviour in global perspective: An introduction to cross-cultural psychology (2nd ed.). Boston, MA: Allyn and Bacon.

Selfe, L. (1983). Normal and anomalous representational drawing ability in children. London: Academic Press Inc.

Selfe, L. (1985). Anomalous drawing development: Some clinical studies. In N. H. Freeman & M. V. Cox (Eds.), *Visual order: the nature and development of pictorial representation* (pp. 135-154). Cambridge: Cambridge University Press.

Shinn, M. W. (1897). Notes on the development of a child. In D. D. Brown *Monograph*. Berkeley, California: University of California Studies.

Slack, J. D. (1996). The theory and method of articulation in cultural studies. In D. Morley and K. H. Chen (eds.), *Stuart Hall: Critical changes in cultural studies*. Pp 112-130. London: Routledge.

Smith, F. O. (1937). What the Goodenough intelligence test measures. *Psychological Bulletin, 34,* 760-761.

Spearman, C. (1904). 'General intelligence', objectively determined and measured. *American Journal of Psychology, 15*, pp. 201-293.

Stern, C. & Stern, W. (1910). Die zeichnerische entwickelung einer knaben von4 bis 7 jahre. *Zeitschrift für angewandte Psychologie, 3,* 1-31.

Sternberg, R. (1986). Intelligence applied: Understanding and increasing your own intellectual skills. New York: Harcourt Brace Jovanovich.

Sternberg, R. J. (1985). Beyond IQ: A triarchic theory of human intelligence. New York: Freeman.

Sternberg, R. J. (1998). Teaching triarchically improves school achievement. *Journal of Educational Psychology,* 33, 374-384.

Sternberg, R. J. (1999). Successful intelligence: finding a balance. *Trends in Cognitive Science, 3(11)*, pp.439-442.

Streiner, D. L. (1995). Learning how to differ: agreement and reliability statistics in psychiatry. *Canadian Journal of Psychiatry, 40(2),* p. 60.

Sully, J. (1907). *Children's ways*. New York: D. Appleton & Co.

Sully, J. (1908). *Studies of childhood*. New York: D. Appleton & Co.

The Psychological Corporation Harcourt Brace Jovanovich Inc. (1963). *The Goodenough-Harris Drawing Test*. Harcourt, Brace & World, Inc.

Thompson, D. 1990. Florence Laura Goodenough. In A. O'Connell and N. Russo, (Eds.). *Women in psychology: A bio-bibliographic sourcebook*. Westport, CT.: Greenwood Press.

Thorndike, E. L. (1913). The measurement of achievement in drawing. *Teach. Coll. Rec. 14(5).*

Töpffer, R. (1848). Réflexions et menus à propos d'un peintre génevois: ou essai sur le beau dans les arts, 2 vols. Paris: J. J. Dubochet, Lechevalier et cie.

Trochim, W. M. K. (2006). *Research methods knowledge base.* Retrieved January 17, 2008, from http://www.socialresearchmethods.net/kb/external.php

Van de Vijver, F. & Leung, K. (1997). *Methods and data analysis for corss-cultural research.* Thousand Oaks, California: Sage Publications, Inc.

Vernon, M. (1907). Kinderkunst and urgeschichte. *Korrespondénz der Deutscher Anthropologische Gesellschaft,* 327-337.

Vygotsky, L. S. (1978). *Mind in society.* Cambridge MA: Harvard University Press.

Wade, D. T. (1992). *Measurement in neurological rehabilitation.* Oxford medical publications. Oxford: Oxford University Press.

Wagner, M. E. & Schubert, H. J. (1955). *D. A. P. Quality Scale.* Kenmore, New York: Delaware Letter Shop.

Wallach, L. (1969). On the bases of conservation. In D. Elkind, & J. H. Flavell (Eds.). *Studies in cognitive development: Essays in honour of Jean Piaget*. New York: Oxford University Press.

Watson, J. B. (1926). Experimental studies of the growth of the emotions. In C. Murchison (Ed.), *Psychologies of 1925* (pp. 37-57). Worcester, Mass: Clark University Press.

Willats, J. (1977). How children learn to represent three-dimensional space in drawing. In G. E. Butterworth (Ed.). *The child's representation of the world*. New York: Plenum, pp. 189-202.

Willats, J. (1981). *Formal structures in drawing and painting.* Unpublished PhD thesis. Council for National Academic Awards.

Williams, J. H. (1935). Validity and reliability of the Goodenough intelligence test. *Sch. & Soc.,* 67, 533-538.

Williams, M. L . (1930). The growth of intelligence as measured by the Goodenough drawing test. *Journal of Applied Psychology, 14*, 239-256.

Wright, B. D. & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370.

Wright, B. D. & Masters, G. N. (1982). *Rating scale analysis.* Chicago: MESA Press.

Wright, B. D. & Mok, M. M. C. (2004). An overview of the family of Rasch measurement models. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch Measurement: Theory, models and application* (pp. 1-24). Maple Grove: JAM Press.

Wright, B. D. & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.

Wright, H. F. (1960). Observational child study. In P. Mussen (Ed.), *Handbook of Research Methods in Child Development* (pp. 71-139). New York: John Wiley.

Wolff, W. (1946). *The personality of the preschool child*. New York: Grune & Stratton.

Yepsen, L. N. (1929). The reliability of the Goodenough drawing test with feebleminded subjects. *Journal of Educational Psychology, 20,* 448-451.

Appendix A

Goodenough-Harris Drawing Test Administration Guide

Test Administration Procedure

The Goodenough-Harris Drawing Test may be administered to children individually or in groups using essentially the same directions. Preschool children, and children with special needs, should be examined individually. Although most children can generally print their names, where necessary, the examiner or classroom teacher is to provide assistance to complete the rest of the information on the front of the test booklet.

Test administration should be followed by some individual, informal questioning to clarify any ambiguous aspects of the drawings. The examiner should start by saying: “Tell me about our picture”. Throughout the questioning period the examiner should try to get at the child’s intentions in the drawings, if a child does not spontaneously identify an ambiguous part of his or her drawings, the examiner may ask (pointing): “What might that be?” The child’s responses should be recorded, and the identification of parts written directly on the drawings.

Each child should be provided with a pencil and a test booklet. Crayons should *not* be used. The number two or two-and-one-half pencil is preferred. See that pictures and books are put aside, to reduce the likelihood of copying.

Have children fill in the information requested on the cover sheet of the test booklet. With children of primary school age it is best to ask them as a group to complete the items one at a time, the examiner directing the task, as follows:

> **Where it says “Name”, *print* your name. Print your first name, and then your last name.**

Now draw a circle around one of the words "Boy" or "Girl", to show whether you are a boy or a girl.

Now print the name of this school.

Where is says "Date of Drawing" put today's date. This is _________.

Where is says "Grade", put your grade in school. (In groups, say: This is the ____ grade.)

Where it says "Age", write how old you are *now*.

Now listen carefully: When were you born? Where is says "Birth Date", first write the date of your birthday, and then write the month. Is it the fourteenth of November, or the second of January? Write whatever date it is. Then put the *year* you were born. Do you know that? If you do, put it down. If not, just leave it blank. (*Note*: Birth dates should always be checked with official records. Ages should be taken to the nearest month.)

When the children have finished supplying the face sheet data, have them fold it back so that the space for the first drawing, and *only* the first drawing, is exposed. Now say:

I am going to ask you to make three pictures for me today. We will make them one at a time. On this first page I want you to make a picture of a man. Make the very best picture that you can; take your time and work very carefully. I want to see whether the boys and girls in __________ School can do as well as those in other schools. Try very hard, and see what good pictures you can make. Be sure to make the whole man, not just his head and shoulders.

When the drawings have been completed, say a few words of praise and have the children turn over the sheets to the space for the second drawing. Then say:

> **This time I want you to make a picture of a woman. Make the very best picture that you can; take your time and work very carefully. Be sure to make the woman, not just her head and shoulders. (Note: With very young children it may be appropriate to say: … picture of a woman, a grown-up girls, or a muumy).**

When this drawing has been completed, praise a bit more lavishly than before as a means of keeping up interest. Then demonstrate how to refold the sheets so that the two completed drawings are inside and the space for the third drawing is now face up. Now say:

> **This picture is to be of someone you know very well, so it should be the best of all. I want each of you to make a picture of yourself – your whole self – not just your face. Perhaps you don't know it but many of the greatest artists like to make their own portraits, and these are often among their best and most famous pictures. So take care and make this last one the very best of the three.**

Children under age eight or nine should have short rest period between drawings two and three. Ask children to put down their pencils, stretch their arms and flex their fingers, to relax from the tension imposed by concentration and effort.

While the children are drawing, stroll about the room and encourage those who are slow or who seem to have difficulty by saying: "These drawings are very fine; you boys and girls are doing very well". *Do not make adverse comment or criticisms, and do* *not give suggestions*. If any child wishes to write about his picture, he may do so at the bottom of the sheet.

If children ask for further instructions, such as whether the man is to doing anything particular like working or running, say: “Do it whatever way you think is best”. Avoid answering “Yes” or “No” or giving any further specific instructions to the children.

The importance of avoiding every kind of suggestion cannot be overemphasized. The examiner must refrain from remarks that might influence the nature of the drawing. The examiner must also see to it that no suggestions come from the children. They should not hold up their drawings for admiration or comment. Young children sometimes accompany their work with a running commentary, such as: “I am giving my man a soldier hat”, or “Mine is a big, big man”. A firm but good-natured, “No one must tell about his / her picture now. Wait until everybody has finished”, will usually dispose of such cases without affecting the general interest or suppressing the child’s enthusiasm for his or her work.

There is no time limit for the test, but young children rarely take more than ten to fifteen minutes for all three drawings. If one or two children are slower than the rest, it is best to collect papers from those who have finished and allow them to go on with their regular work while the slower workers are finishing.

In older groups, above the fifth or sixth grade, it may be necessary to offer strong encouragement to some children, who will say they can’t do the task. In these groups it may also be desirable to say:

You are to make three drawings, one on each of the three pages of this folder. The instructions are at the top of each page. When you have finished one drawing go right on the next, until you have finished all three.

In this case, it is well to have two examiners who can walk about the room speaking to individuals who seem reluctant to attempt the task.

The following special circumstances should be noted: (1) A child may spoil his or her drawings and wish to start again. In such case he or she should be given a fresh test booklet and be allowed to try again. All such instances should be noted on the margin of the booklet after the child has finished his work. (2) Above the second grade (rarely below), a child may draw a bust picture only. When it is evident that this has been the intention, a fresh test booklet should be given, and the child told to, "Make a whole man". Both drawings should be preserved for comparison (Harris, 1963, pp. 239-242).

Goodenough-Harris Drawing Test Booklet

Goodenough-Harris Drawing Test

By Florence L. Goodenough and Dale B. Harris

Name__________ Boy_____ Girl_____

School__________ Date of Drawing__________

Grade__________ Age_____ Birth Date__________

Father's Occupation__________

Examiner's Notes

Summary

	Raw Score	Standard Score	Percentile Rank
Point Scale			
Man	_____	_____	_____
Woman	_____	_____	_____
Average		_____	
Self	_____	_____	_____
Quality Scale			
Man	_____	_____	_____
Woman	_____	_____	_____
Average		_____	

ISBN 015411855-9

9780154118554

THE PSYCHOLOGICAL CORPORATION®
HARCOURT BRACE JOVANOVICH, INC.

Printed in the United States of America.

09-118555

Note. From "The Goodenough-Harris Drawing Test", by The Psychological Corporation Harcourt Brace Jovanovich, Inc, 1963.

Make Your First Drawing Here

Draw a picture of a man. Make the very best picture you can. Be sure to make the whole man, not just his head and shoulders.

1.____	41.____
2.____	42.____
3.____	43.____
4.____	44.____
5.____	45.____
6.____	46.____
7.____	47.____
8.____	48.____
9.____	49.____
10.____	50.____
11.____	51.____
12.____	52.____
13.____	53.____
14.____	54.____
15.____	55.____
16.____	56.____
17.____	57.____
18.____	58.____
19.____	59.____
20.____	60.____
21.____	61.____
22.____	62.____
23.____	63.____
24.____	64.____
25.____	65.____
26.____	66.____
27.____	67.____
28.____	68.____
29.____	69.____
30.____	70.____
31.____	71.____
32.____	72.____
33.____	73.____
34.____	
35.____	
36.____	
37.____	
38.____	
39.____	
40.____	

Raw Score ____

Make Your Second Drawing Here

Draw a picture of a woman. Make the very best picture you can. Be sure to make the whole woman, not just her head and shoulders.

1.____	41.____
2.____	42.____
3.____	43.____
4.____	44.____
5.____	45.____
6.____	46.____
7.____	47.____
8.____	48.____
9.____	49.____
10.____	50.____
11.____	51.____
12.____	52.____
13.____	53.____
14.____	54.____
15.____	55.____
16.____	56.____
17.____	57.____
18.____	58.____
19.____	59.____
20.____	60.____
21.____	61.____
22.____	62.____
23.____	63.____
24.____	64.____
25.____	65.____
26.____	66.____
27.____	67.____
28.____	68.____
29.____	69.____
30.____	70.____
31.____	71.____
32.____	
33.____	
34.____	
35.____	
36.____	
37.____	
38.____	
39.____	
40.____	

Raw Score ____

Make Your Third Drawing Here

Draw a picture of yourself. Make the very best picture you can. Be sure to make your whole self, not just your head and shoulders.

1.____	41.____
2.____	42.____
3.____	43.____
4.____	44.____
5.____	45.____
6.____	46.____
7.____	47.____
8.____	48.____
9.____	49.____
10.____	50.____
11.____	51.____
12.____	52.____
13.____	53.____
14.____	54.____
15.____	55.____
16.____	56.____
17.____	57.____
18.____	58.____
19.____	59.____
20.____	60.____
21.____	61.____
22.____	62.____
23.____	63.____
24.____	64.____
25.____	65.____
26.____	66.____
27.____	67.____
28.____	68.____
29.____	69.____
30.____	70.____
31.____	71.____
32.____	72.____
33.____	73.____
34.____	
35.____	
36.____	
37.____	
38.____	
39.____	
40.____	

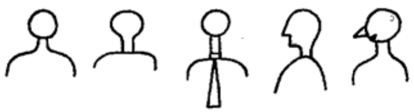
Raw Score ____

Goodenough-Harris Drawing Test Draw-a-Man Scoring Guide

Requirements for Scoring the Draw-a-Man Scale

ITEM	DESCRIPTION
1. **Head present**	Any clear method of representing the head. Features alone, without any outline for the head itself, are not credited for this point.
2. **Neck present**	Any clear indication of the neck as distinct from the head and the trunk. Mere juxtaposition of the head and the trunk is not credited.
3. **Neck, two dimensions**	Outline of neck continuous with that of the head, of the trunk, or of both. Line of neck must "flow" into head line or trunk line. Neck interposed as pillar between head and trunk does not get credit unless treated definitely to show continuity between neck and head or trunk or both, as by collar, or curving of lines.

Credit



No Credit

ITEM	DESCRIPTION
4. **Eyes present**	Either one or two eyes must be shown. Any method is satisfactory. A single indefinite feature, such as is occasionally found in the drawings of very young children, is credited.
5. **Eye detail: brow or lashes**	Brow, lashes or both shown.
6. **Eye detail: pupil**	Any clear indication of the pupil or iris as distinct from the outline of the eye. Both must appear if both eyes are shown.
7. **Eye detail: proportion**	The horizontal dimension of the eye must be greater than the vertical dimension. This requirement must be fulfilled in both eyes if both are shown; one eye is sufficient if only one is shown. Sometimes in profile drawings of a high grade the eye is shown in perspective. In such drawings any triangular form approximating the following examples is credited.

Credit

ITEM	DESCRIPTION
8. **Eye detail: glance**	**Full Face:** The eyes obviously glancing. There must be no convergence or divergence of the two pupils, either horizontally or vertically.

Credit

Profile: The eyes must either be shown as in the preceding point, or, if the ordinary almond form is retained, the pupil must be placed toward the front of the eye rather than in the center. The scoring should be strict.

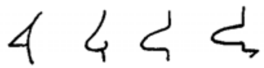
ITEM	DESCRIPTION
9. **Nose present**	Any clear method of representation. In "mixed profiles," the score is plus even though *two* noses are shown.
10. **Nose, two dimensions**	**Full Face:** Credit all attempts to portray the nose in two dimensions, when the bridge is longer than the width of the base or tip.

Credit

No Credit

Profile: Credit all crude attempts to show the nose in profile, provided tip or base is shown in some manner. Do not credit simple "button."

Credit



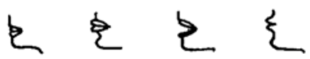
No Credit

ITEM	DESCRIPTION
11. **Mouth present**	Any clear representation.
12. **Lips, two dimensions**	**Full Face:** Two lips clearly shown.

Credit

Profile:

Credit



No Credit

13. Both nose and lips in two dimensions — Bonus point given when Items 10 and 12 are passed. See preceding items for accepted forms.

14. Both chin and forehead shown — **Full Face:** Both the eyes and mouth must be present, and sufficient space left above the eyes to represent the forehead; below the mouth to represent the chin. The scoring should be rather lenient. Where neck is continuous with face, placement of mouth with respect to narrowing of lower portion of head is important. The sketches below illustrate mouth placement.

Credit

No Credit

15. Projection of chin shown; chin clearly differentiated from lower lip — **Full Face:** Modeling of chin must be indicated in some way, as by a curved line below the mouth or lip, or point of chin indicated by appropriate facial modeling, or dot or line placed below mouth near lower limit of face. Beard obscuring chin does *not* score. *Note:* Distinguish carefully from Item 16. There must definitely be an attempt to show a "pointed" chin to credit this item. This point is credited most frequently in profiles.

Credit

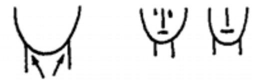
Items 15 and 16

Item 15 but not 16

Item 16 but not 15

16. Line of jaw indicated — **Full Face:** Line of jaw and chin drawn across neck but not squarely across. Neck must be sufficiently wide, and chin must be so shaped that the line of the jaw forms a well-defined acute angle with the line of the neck. Score *strictly* on the simple oval face.

Credit

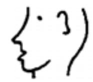


ACUTE ANGLES

No Credit

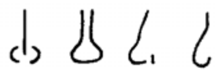
Profile: Line of jaw extends toward ear.

Credit

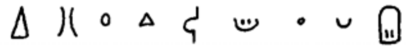


17. Bridge of nose — **Full Face:** Nose properly placed and shaped. The base of the nose must appear as well as the indication of a straight bridge. Placement of upper portion of bridge is important; must extend up to or between the eyes. Bridge must be narrower than the base.

Credit



No Credit



Profile: Nose at angle with face, approximately 35–45 degrees. Separation of nose from forehead clearly shown at eye.

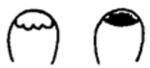
Credit

No Credit

18. Hair I — Any indication of hair, however crude.

19. Hair II — Hair shown on more than circumference of head and more than a scribble. Nontransparent, unless it is clear that a bald-headed man is portrayed. A simple hairline across the skull on which no attempt has been made to shade in hair does *not* score. If any attempt has been made, even in outline or with a little shading, to portray hair as having substance or texture, the item scores.

Credit



No Credit

20. Hair III — Any clear attempt to show cut or styling by use of side burns, a forelock, or conformity of base line to a "style." When a hat is drawn, credit the point if hair is indicated in front as well as behind the ear, or if hairline at back of neck or across forehead suggests styling.

21. Hair IV — Hair shaded to show part, or to suggest having been combed, or brushed, by means of *directed* lines. Item 21 is never credited unless Item 20 is; it is thus a "high-grade" point.

Credit

No Credit

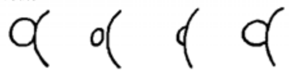
22. Ears present — Any indication of ears.

23. Ears present: proportion and position — The vertical measurement must be greater than the horizontal measurement. The ears must be placed somewhere within the middle two-thirds of the head.

Full Face: The top of the ear must be separated from the head line, and *both* ears must extend from the head.

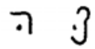
Credit

No Credit

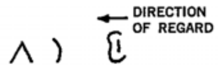


Profile: Some detail, such as a dot, to represent the aural canal must be shown. The shell-like portion of the ear must extend toward the back of the head. (Some children, especially retarded boys, tend to reverse this position, making the ear extend toward the face. In such drawings this item is never credited.)

Credit



No Credit



24. Fingers present — Any suggestion of fingers, separate from hand or arm. In drawings by older children, where there is a tendency to "sketch," credit this point if any suggestion of fingers occurs.

25. Correct number of fingers shown — Both hands necessary if both hands are shown. Credit this point in "sketchy" drawings by older children, even though five digits may not be definitely discerned.

26. Detail of fingers correct — "Grapes" or "sticks" do not score. Length of individual fingers must be distinctly greater than width. In well-executed drawings, where hand may appear in perspective, or where fingers are indicated by "sketching," credit this point. Credit also those cases in which, because the hand is obviously clenched, only the knuckles or part of the fingers appear. This last will occur only in high-quality drawings where there is considerable use of perspective.

27. Opposition of thumb shown — Fingers must be indicated, with a clear differentiation of the thumb from the fingers. Scoring should be very strict. The point is credited if one of the lateral digits is definitely shorter than any of the others (compare especially with the little finger), or if the angle between it and the index finger is not less than twice as great as that between any two of the other digits, or if its point of attachment to the hand is distinctly nearer to the wrist than that of the fingers. Conditions must be fulfilled on both hands if both are shown; one hand is sufficient if only one is shown. *Fingers must be present or indicated;* "mitt" hand does *not* score, unless figure is definitely in winter garb, wearing mittens.

Credit

No Credit

28. Hands present

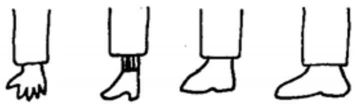
Any representation of the hand, apart from the fingers. When fingers are shown, a space must be left between base of fingers and edge of sleeve or cuff. Where no cuff exists, arm must broaden in some way to suggest palm or back of hand as distinct from wrist. Characteristic must appear on both hands if both are shown.

Marginal Credit

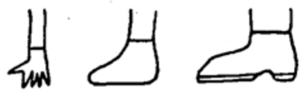
29. Wrist or ankle shown

Either wrist or ankle clearly indicated as separate from sleeve or trouser. A line across the limb to indicate the end of sleeve or trouser, although credited in Item 55, is *not* sufficient here.

Credit



No Credit



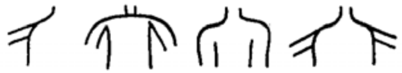
30. Arms present

Any method of representation clearly intended to indicate arms. Fingers alone are not sufficient, but the point is credited if any space is left between the base of the fingers and that part of the body to which they are attached. The number of arms must also be correct, except in profile drawings when only one arm may score.

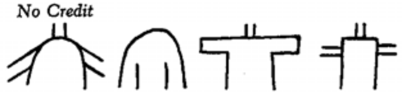
31. Shoulders I

Full Face: A change in the direction of the outline of the upper part of the trunk which gives an effect of concavity rather than convexity. The point is scored rather strictly. The ordinary elliptical form is never credited, and the score is always minus unless it is evident that there has been a recognition of the abrupt broadening out of the trunk below the neck which is produced by the shoulder blade and the collar bone. A perfectly square or rectangular trunk does not score, but if the corners have been rounded, the point is credited.

Credit



No Credit



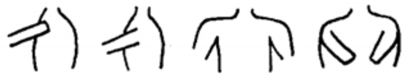
Profile: The scoring should be somewhat more lenient than in full-face drawings, since it is more difficult to represent the shoulders adequately in the profile position. A profile drawing, in this connection, should be understood to mean one in which the trunk, as well as the head, is shown in profile. If the lines forming the outline of the upper part of the trunk diverge from each other at the base of the neck in such a way as to show the expansion of the chest, the point is credited.

32. Shoulders II

Full Face: Score more strictly than previous item. Shoulders must be continuous with neck and arms, and "square," not drooping. If arm is held from the body, the armpit must be shown.

Profile: Shoulder joint in approximately correct position. Arm must be represented by double line.

Credit



No Credit

33. Arms at side or engaged in activity

Full Face: Young children generally draw the arms stiffly out from the body. Credit this point when at least one arm is down at the side, making an angle of no more than 10 degrees with the general vertical axis of the trunk, unless the arms are engaged in some definite activity, such as carrying an object. Credit when hands are in pockets, on hips, or behind back.

Credit

10° OR LESS

Profile: Credit if hands are engaged in definite activity, or if upper arm is suspended even though forearm is extended.

34. Elbow joint shown — There must be an abrupt bend (not a curve) at approximately the middle of the arm. One arm is sufficient. Modeling or creasing of the sleeve is credited.

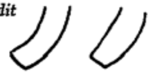
Full Face:

Credit

Profile:

Credit

No Credit



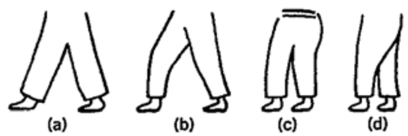
35. Legs present — Any method of representation clearly intended to indicate the legs. The number must be correct: two in full-face drawings; either one or two in profiles. Use common sense rather than a purely arbitrary scoring. If only one leg is present, but a rough sketch of a crotch is included, showing clearly what the child has in mind, score the item. On the other hand, three or more legs, or a single leg without logical explanation should be scored minus. A single leg to which two feet are attached is scored plus. Legs may be attached anywhere to the figure.

36. Hip I (crotch) — **Full Face:** Crotch indicated. This is most frequently shown by inner lines of the two legs meeting at point of junction with the body. (Young children usually place the legs as far apart from each other as possible, and this never scores.)

Credit

Profile: If only one leg shows, buttock must be shaped.

Credit

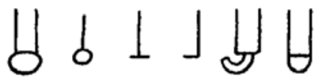


37. Hip II — Preceding item earned with credit to spare. Drawing gives a better idea of the hip than required for passing preceding item. Examples (b) and (d) on Item 36 are credited here also; (a) and (c) are not.

38. Knee joint shown — There must be, as in the case of the elbow, an abrupt bend (not curve) at about the middle of the leg, or, as is sometimes found in very high-quality drawings, a narrowing of the leg at this point. Knee-length trousers are not sufficient. Crease or shading to indicate knee is scored plus.

39. Feet I: any indication — Feet indicated by any means: two feet in full-face; one or two in primitive profile. Young children may indicate feet by attaching toes to the end of the leg. This is credited.

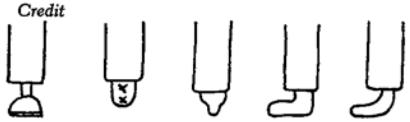
Credit



40. Feet II: proportion — The feet and legs must be shown in two dimensions. Feet must not be "clubbed"; that is, the length of the foot must be greater than its height from sole to instep. The length of the foot must be not more than one-third or less than one-tenth the total length of the leg. The item is also credited in full-face drawings in which the foot is shown in perspective, longer than wide, provided the foot is separated in some way from the rest of the leg, and not merely indicated by a line across the leg.

Full Face:

Credit

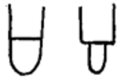


No Credit

41. Feet III: heel

Any clear method of indicating the heel. In full-face drawings, credit the item arbitrarily when the foot is shown as below, provided there is some demarcation between the foot and the leg. In the profile, the instep *must* be indicated.

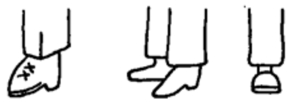
Credit



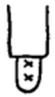
42. Feet IV: perspective

Foreshortening attempted in at least one foot.

Credit



No Credit



43. Feet V: detail

Any one item of detail such as lacing, tie, strap, or shoe sole indicated by a double line.

44. Attachment of arms and legs I

Both arms and both legs attached to the trunk at any point, or arms attached to the neck, or at the juncture of the head and the trunk when the neck is omitted. If the trunk is omitted, the score is always zero. If the legs are attached elsewhere than to the trunk, regardless of the attachment of the arms, the score is zero. If only one arm or leg is shown, either in full-face or in profile drawings, credit may be given on the basis of the limb that is shown. If both arms and legs are shown, the members of each pair must be attached approximately symmetrically. Arms attached to the legs score zero.

45. Attachment of arms and legs II

Legs attached to trunk, and arms attached to the trunk at the correct point. Do not credit if arm attachment occupies one-half or more of the chest area (neck to waist). When no neck is present, the arms must *definitely* be attached to the upper part of the trunk.

Full Face: When Item 31 is plus, the point of attachment must be exactly at the shoulders. If Item 31 is zero, the attachment must be exactly at the point which should have been indicated as the shoulders. Score very strictly, especially in those cases where Item 31 is zero.

Profile: Do not credit if both the lines delineating the arm extend from the outline of the back, or if the point of attachment either reaches the base of the neck, or falls below the greatest expansion of the chest line.

46. Trunk present

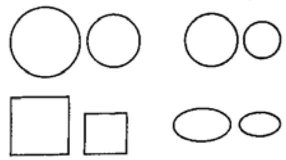
Any clear indication of the trunk, either one or two dimensional. Where there is no clear differentiation between the head and the trunk, but the features appear in the upper end of a single figure, the point is scored plus if the features do not occupy more than half the length of the figure; otherwise, the score is zero, unless a cross line has been drawn to indicate the termination of the head. A single figure placed between the head and the legs is always counted as a trunk, even though its size and shape may suggest a neck rather than a trunk. (This ruling is based on the fact that, when questioned, a number of children whose drawings showed this peculiarity, called the part a trunk.) A row of buttons extending down between the legs is scored zero for trunk but plus for clothing, unless a cross line has been drawn to show the termination of the trunk.

47. Trunk in proportion, two dimensions

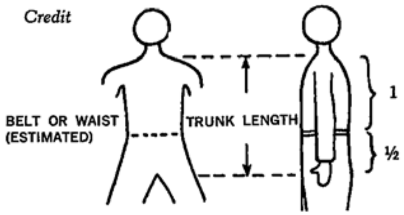
Length of the trunk must be greater than breadth. Measurement should be taken at the points of greatest length and of greatest breadth. If the two measurements are equal, or so nearly so that the difference is not readily determined, the score is zero. In most instances the difference will be great enough to be recognized at a glance, without actually measuring.

48. Proportion: head I

Area of the head not more than one-half or less than one-tenth that of the trunk. Score rather leniently. See below for a series of standard forms of which the first is double the area of the second in each pair.



49. Proportion: head II — Head approximately one-fourth trunk area. Score strictly; over one-third or under about one-fifth fails the item. Where crotch is not shown, as in some profiles, consider belt or waist at about two-thirds down total trunk length.



50. Proportion: face — Full Face: Length of head greater than its width. Should show a general oval shape.

Profile: Head definitely elongated. Face longer than "dome" of skull.

51. Proportion: arms I — Arms at least equal to the trunk in length. Tips of hands extend to middle of hip but not to knee. Hands need not necessarily extend to or below the crotch, especially if legs are unusually short. In full-face drawings, both hands must so extend. Score by relative *lengths*, not *position*, of arms.

52. Proportion: arms II — Arms taper; forearm narrower than upper arm. Any tendency to narrow the forearm except right at the wrist, is credited. If both arms show clearly, tapering must occur in both.

53. Proportion: legs — Length of the legs not less than the vertical measurement of the trunk nor greater than twice that measurement. Width of either leg less than that of the trunk.

54. Proportion: limbs in two dimensions — Both arms and legs shown in two dimensions. If the arms and legs are in two dimensions, the point is credited, even though the hands and feet are drawn in linear dimension.

55. Clothing I — Any clear representation of clothing. As a rule the earliest forms consist of a row of buttons running down the center of the trunk, or of a hat, or of both. Either alone scores. A single dot or small circle placed in the center of the trunk is practically always intended to represent the navel and should not be credited as clothing. A series of vertical or horizontal lines drawn across the trunk (and sometimes on the limbs as well) is a fairly common way of indicating clothing, and should be so credited. Marks to indicate pockets or sleeve-ends also get credit.

56. Clothing II — At least two articles of clothing (as hat and trousers) nontransparent; that is, concealing the part of the body which they are supposed to cover. In scoring this point it must be noted that a hat which is merely in contact with the top of the head but does not cover any part of it is *not* credited. Buttons alone, without any other indication of the coat, are *not* credited. Two of the following must be present to indicate coat: sleeves, collar or neckline, buttons, or pockets. Trousers must be clearly intended by belt, fly, pockets, cuff, or any separation of feet or leg from bottom of trouser leg. Foot as an extension of leg does not score, when a line drawn across the leg is the only way of indicating the separation of foot and leg.

57. Clothing III — Entire drawing free from transparencies of any sort. Both sleeves and trousers must be shown as distinct from wrists or hands and legs or feet.

58. Clothing IV — At least four articles of clothing definitely indicated. The articles should be among those in the following list: hat, shoes, coat, shirt, collar, necktie, belt, trousers, jacket, sport shirt, overalls, socks (pattern). *Note:* Shoes must show some detail, as laces, toe cap, or double line for the sole. Heel alone is not sufficient. Trousers must show some features, such as fly, pockets, cuffs. Coat or shirt must show either collar, sleeves, pockets, lapels, or distinctive shading, as spots or stripes. Buttons alone are not sufficient. Collar should not be confused with neck shown merely as insert. The necktie is often inconspicuous and care must be taken not to overlook it, but it is not likely to be mistaken for anything else.

59. Clothing V — Costume complete without incongruities. This may be a "type" costume (e.g., cowboy, soldier) or costume of everyday dress. If the latter, it should be clearly recognized as appropriate; e.g., sport shirt on man, cap appropriate to hunting outfit, overalls for farmer. This is a "bonus" point, and must show more than necessary for Item 58.

60. Profile I — The head, trunk, and feet must be shown in profile without error. The trunk may not be considered as drawn in profile unless the characteristic line of buttons has been moved from the center to the side of the figure, or some other indication, such as the position of the arms, pockets, or necktie shows clearly the effect of this position. The entire drawing may contain *one*, but not more than one, of the following three errors:

1. One body transparency, such as the outline of the trunk showing through the arm.
2. Legs not in profile. In a true profile at least the upper part of the leg which is in the background must be concealed by the one in the foreground.
3. Arms attached to the outline of the back and extending forward.

61. Profile II — The figure must be shown in true profile, without error or any body transparency.

62. Full face — (Include partial profile, where attempt is to show figure in perspective.) All major body parts in proper location and correctly joined unless hidden by perspective or clothing.

Essential items: Legs, arms; eyes, nose, mouth, ears; neck, trunk; hands and feet. Parts must be in two dimensions. Feet may be in perspective, but not in profile, unless they turn "out" in opposite directions.

63. Motor coordination: lines [4] — Look at the long lines in arms, legs, and trunk. Lines should be firm, well-controlled and free from accidental wavering. A few long lines may be retraced or erased. The drawing need not achieve very smoothly "flowing" lines to earn credit. Young children sometimes "color in" with their pencils; examine carefully the fundamental lines of their drawings. Older children frequently use a "sketching" technique readily distinguishable from the uncertain, wavering lines resulting from immature coordination. If the general effect is that of firm, sure lines showing that the pencil was under control, credit the item. The drawing may be quite immature and still score on this point.

64. Motor coordination: junctures — Look at the juncture points of lines. They must meet cleanly without a marked tendency to cross or overlap, or leave gaps between the ends. A drawing with few lines is scored more strictly than one with frequent changes in direction of line. A "sketchy" drawing is ordinarily credited even though the junctures of lines may seem uncertain, since this is a characteristic confined almost entirely to drawings of a mature type. Some erasures may be allowed.

65. Superior motor coordination — This is a "bonus" point for good pencil work on details as well as on major lines. Look at the small detail as well as at the character of the major lines. All lines should be firmly drawn, with correct joining. Pencil work in fine detail—facial features, small items of clothing, etc.—indicates a good control of the pencil. Scoring should be quite strict. Erasures and/or redrawing invalidate this item.

[4] Items 63, 64, and 65 concern the *quality* of the child's control of the pencil. These items evaluate the firmness and sureness of line, quality of line junctions, "corners," etc.

66. Directed lines and form: head outline [5] — Outline of head must be drawn without obviously unintentional irregularities. The point is credited only in drawings where the shape has developed beyond the first crude circle or ellipse. In profile drawings, a simple oval to which a nose has been added does *not* score. Scoring should be rather strict; the contour of the face must be developed as a unit, *not* by adding parts.

67. Directed lines and form: trunk outline — Same as for the preceding item, but here with reference to the trunk. Note that the primitive "stick," circle, or ellipse does *not* score. The body lines must show an attempt to follow an intentional deviation from the simple ovoid form.

68. Directed lines and form: arms and legs — Arms and legs must be drawn without irregularities, as in above item, and without tendency to narrowing at the points of junction with the body. Both arms and legs must be in two dimensions.

69. Directed lines and form: facial features — Facial features must be symmetrical in all respects. Eyes, nose, and mouth must all be shown in two dimensions.

Full Face: The features must be appropriately placed, regular and symmetrical, giving a clear appearance of the human form.

Profile: The eye must be regular in outline and located in the forward one-third of the head. The nose must form an obtuse angle with the forehead. The scoring should be strict; a "cartoon" nose is not credited.

70. "Sketching" technique — Lines formed by well-controlled short strokes. Repeated tracing of long line segments is not credited. "Sketching" technique appears in the work of some older children and almost never occurs under age eleven or twelve.

71. "Modeling" technique — "Lines" or shading must indicate one or more of the following: garment creases, wrinkles or folds, other than trouser press; fabric; hair; shoes; "coloring in"; or background features.

72. Arm movement — Figure must express freedom of movement in both shoulders and elbows. One arm suffices. Credit hands on hips or in pockets, if both shoulders and elbows are apparent. A definite activity need *not* be indicated.

73. Leg movement — Freedom of movement portrayed both in hips and knees of the figure.

[5] Items 66–69 concern the child's deliberate direction of the pencil to produce a good form. The child's work must show that he has exercised control, firmly and surely.

Goodenough-Harris Drawing Test Draw-a-Woman Scoring Guide

Requirements for Scoring the Draw-a-Woman Scale

ITEM — DESCRIPTION

1. Head present — Any clear method of representing the head. Features alone, without any outline for the head itself, are not credited for this point.

2. Neck present — Any clear indication of the neck as distinct from the head and the trunk. Mere juxtaposition of the head and the trunk is not credited.

3. Neck, two dimensions — Outline of neck continuous with that of the head, of the trunk or of both. Line of neck must "flow" into head line or trunk line. Neck interposed as pillar between head and trunk does not get credit unless treated definitely to show continuity between neck and head or trunk or both, as by collar, or curving of lines.

Credit

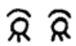
No Credit

4. Eyes present — Either one or two eyes must be shown. Any method is satisfactory. A single indefinite feature, such as is occasionally found in the drawings of very young children, is credited. Credit also, in mature drawings attempting perspective, any indication of the eye by contour of the profile, as:

5. Eye detail: brow or lashes — Brow, lashes or both shown.

Full Face:

Credit



Profile:

Credit

No Credit

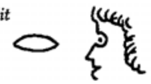
6. Eye detail: pupil — Pupil shown. Credit any clear indication of the pupil or iris as distinct from the outline of the eye. Both pupils must appear if both eyes are shown.

7. Eye detail: proportion — The horizontal measurement of the eye must be greater than the vertical dimension. This requirement must be fulfilled in both eyes if both are shown; one eye is sufficient if only one is shown. In profile drawings, any triangular forms which approximate the example below are credited.

Profile:

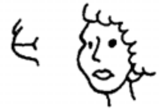
Credit

No Credit



8. Cheeks — Credit modeling or "shading" on cheeks or at mouth corners. Credit also "cosmetic cheeks"— circular spots on cheeks. In drawings which attempt perspective, credit any indication in *contour* of face.

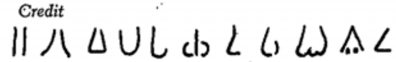
Credit



9. Nose present — Any clear method of representation. In "mixed profiles," the score is plus even though *two* noses are shown.

10. Nose, two dimensions — **Full Face:** Credit all attempts to portray the nose in two dimensions, when the bridge is longer than the width of the base or tip.

Credit



No Credit

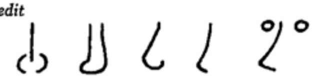
Profile: Credit all crude attempts to show the nose in profile, provided tip or base is shown in some manner. Do not credit simple "button."

No Credit

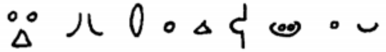
11. Bridge of nose

Full Face: Nose properly placed and shaped. The base of the nose must appear as well as the indication of a straight bridge. Placement of upper portion of bridge is important; must extend *up to or between the eyes. Bridge must be narrower than the base.*

Credit



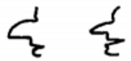
No Credit



Profile: Nose at angle with face, approximately 45 degrees. Separation of nose from forehead clearly shown at eye.

Credit

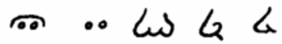
No Credit



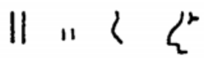
12. Nostrils shown

Any attempt to portray nostrils as holes, dots, or to show "wings."

Credit



No Credit



13. Mouth present

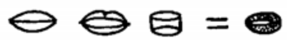
Any clear representation.

14. Lips, two dimensions

Two lips clearly shown.

Full Face:

Credit



Profile:

Credit

No Credit

15. "Cosmetic lips"

Any clear attempt to show "Cupid's bow." Score based on the outer shape. Two lips need *not* be shown.

Credit

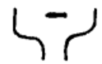
16. Both nose and lips in two dimensions

Bonus point given when both Items 10 and 14 are passed.

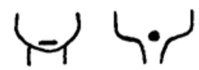
17. Both chin and forehead shown

Full Face: Sufficient space must be left above the eyes to represent the forehead, and below the mouth to represent the chin. The scoring should be rather lenient. Where neck is continuous with face, placement of mouth with respect to narrowing of lower portion of head is important.

Credit



No Credit



Profile: The point may be credited when the eyes and mouth are omitted, if the outline of the face shows clearly the limits of the chin and forehead. Score leniently if forehead is covered by hat brim; more strictly if covered by hair.

18. Line of jaw indicated

Full Face: Line of jaw and chin drawn across neck but not squarely across. Neck must be sufficiently wide, and chin must be so shaped that the line of the jaw forms a well defined acute angle with the line of the neck. Score *strictly* on the simple oval face.

Credit

ACUTE ANGLES

Profile: Line of jaw extends toward (but not all the way to) the ear or across the neck.

Credit

No Credit

19. Hair I — Any indication of hair, however crude.

20. Hair II — Scribble closely conforming to head, *or*

Full Face: Shaped masses suggesting braids or locks each side of face.

Credit

Profile: Mass dependant in back.

Credit

21. Hair III — Style suggested by indentation at temple, or bangs, or shaped at lower ends, or both. General "style" achieved. Distinctly better design than Item 20.

22. Hair IV. — Use of directed lines to indicate a part, texture, or combing. Superior style achieved.

Caution: Score strictly; superior style may be achieved with *outline* sketching, but this does not score. Directed lines to indicate hair texture must appear, and be better than "coloring in."

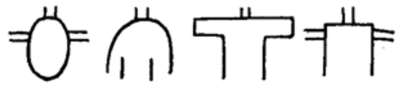
23. Necklace or earrings — Any clear indication. Distinguish necklace from neckline or collar of dress. Earrings without ears (which may be concealed by hair) should be credited.

24. Arms present — Any method of representation clearly intended to indicate arms. Fingers alone are not sufficient, but the point is credited if any space is left between the base of the fingers and that part of the body to which they are attached. The number of arms must be correct, except in profile drawings when only one arm may score.

25. Shoulders — **Full Face:** A distinct change in the direction of the upper part of the trunk, which gives the effect of a "*rounded corner.*" The ordinary elliptical form is never credited. There must be an abrupt broadening of the trunk below the neck, which then turns downward into the arms or sides of the trunk. Square corners fail.

Credit

No Credit



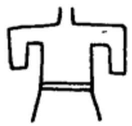
Profile: Somewhat more lenient where the trunk as well as the head is shown in profile. If the lines that form the upper part of the trunk diverge from each other at the base of the neck so as to show the expansion of the chest, credit the point.

26. Arms at side (or engaged in activity or behind back) — **Full Face:** Young children generally draw the arms held stiffly out from the body. Credit this point when at least one arm is down at the side, making an angle of no more than 10 degrees with the general vertical axis of the trunk, unless the arms are engaged in some definite activity, such as carrying an object. Credit when hands are placed on hips or behind the back.

Credit

10° OR LESS

No Credit



Profile: Credit if *hands are engaged in definite activity,* or if upper arm is suspended, even though forearm is extended.

Credit

No Credit

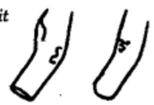
27. Elbow joint shown There must be an abrupt bend (not a curve) at approximately the middle of the arm. One arm is sufficient. Modeling or creasing of the sleeve is credited.

Full Face:

Credit

Profile:

Credit



No Credit

28. Fingers present Any indication of fingers. Mitt hand does not score even if thumb is shown.

29. Correct number of fingers shown If both hands are shown, the correct number on each is necessary, unless there is a clear attempt to portray hand activity which would conceal the correct number. Credit drawings produced by older children who try a "sketching" technique, even though five digits may not be definitely discerned.

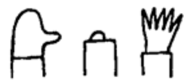
Credit

30. Detail of fingers correct "Grapes" or "sticks" do not score. Length of individual fingers must be distinctly greater than width. In well-executed drawings, where hand may appear in perspective, or where fingers are indicated by "sketching," credit this point. Credit also those cases in which, because the hand is obviously clenched, only the knuckles or part of the fingers appear. This last will occur only in high-quality drawings where there is considerable use of perspective.

31. Opposition of thumb shown A clear differentiation of the thumb from the fingers. Scoring should be very strict. The point is credited if one of the lateral digits is definitely shorter than any of the others (compare especially with little finger), or if the angle between it and the index finger is not less than twice as great as that between any two of the other digits, or if its point of attachment to the hand is distinctly nearer to the wrist than that of the fingers. Conditions must be fulfilled on both hands if both are shown, unless hand is grasping something; one hand is sufficient if only one is shown. Five digits are necessary for thumb to score. *Fingers must be present or indicated;* "mitt" hand does *not* score unless subject is definitely shown in winter garb, wearing mittens.

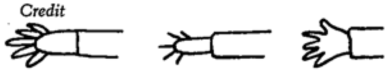
Credit

No Credit

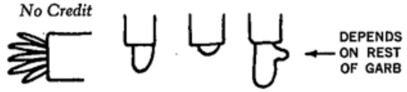


32. Hands present Any representation of the hand, apart from the fingers. When fingers are shown a space must be left between base of fingers and edge of sleeve or cuff. Where no cuff exists, arm must broaden in some way to suggest palm or back of hand as distinct from wrist. Characteristic must appear on both hands, if both are shown. "Mitt" hand with thumb does not score unless figure obviously is wearing mittens.

Credit



No Credit

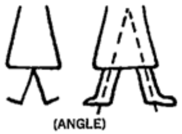


Marginal Credit

33. Legs present — Any method of representation clearly intended to indicate the legs. There must be two legs in full-face drawings, and either one or two, in profiles. Credit where long skirt hides legs or feet.

34. Hip — **Full Face:** The principal axes of the legs must form a distinct angle. The distance between the ankles must be greater than the distance between the inner surfaces of the legs at the skirt line, and the difference must be more than can be accounted for by contours of the calf and ankle. Do not credit in the case of a long gown.

Credit



No Credit

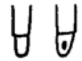
Profile: Credit when legs form angle, as in walking. Credit in standing figure, when one leg is shown, or when two appear in true profile.

Credit

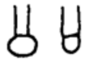
35. Feet I: any indication — Feet indicated by any means: two feet in full-face; one or two in profile. In the case of a long gown, credit this item.

36. Feet II: proportion — **Full Face:** Feet must be longer than wide, or drawn in perspective.

Credit

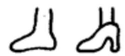


No Credit

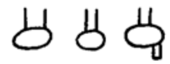


Profile: Horizontal dimension of fore-part of foot must be greater than vertical dimension. In the case of a long gown, credit only when foot is indicated in some way, as by the tip appearing beneath the edge of the gown, etc.

Credit



No Credit



37. Feet III: detail — Foot or shoe must show some ornamentation, such as a buckle, tie, strap, or sole. In the case of a long gown, do *not* credit unless foot is shown.

38. Shoe I: "feminine" — Credit any clear attempt to depict a feminine shoe as opposed to "brogan" or other thick, solid shoe. Note especially attempts to depict slender toe or arch, high heel, open toe, or straps. If heel is crucial point, it should be at least one-third of total height of shoe at that point. Shoe must be marked off from leg, either by a line or by profile shaping. In the case of a long gown, credit only when shoe is shown.

Credit

39. Shoe II: style — Shoe must be clearly feminine and "styled," i.e., clearly a pump, tie, open toe, wedgie, saddle-shoe, etc. In the case of a long gown, credit only when clearly shown.

40. Placement of feet appropriate to figure

Full Face: Feet turned "in" or "out," or in perspective. Do not credit primitive feet.

No Credit

Profile: Credit both feet turned in direction of head. Do not credit when feet are absent, except where long gown hides feet.

41. Attachment of arms and legs I

Both arms and legs attached to the trunk at any point, or arms attached to the neck, or at juncture of head and trunk when neck is omitted. Do not credit if either arms or legs are missing. Credit where dress hides legs and/or feet. If the trunk is omitted, the score is always zero. If the legs are attached elsewhere than to the trunk, regardless of the attachment of the arms, the score is zero. If only one arm or leg is shown, either in full-face or profile drawings, credit may be given on the basis of the limb that is shown. If both arms and legs are shown, the members of each pair must be attached approximately symmetrically. Credit where long dress hides legs and/or feet. Be careful to distinguish this item from Item 25.

Credit

42. Attachment of arms and legs II

Arms attached to the trunk at the correct position. Legs attached to the bottom of the trunk or skirt and not continuous with vertical line or drape of the skirt. Credit this point if both feet and legs are hidden by long gown.

Legs:

Credit

No Credit

Arms: **Full Face:** Where Item 25 is failed, attachment must be exactly at the point where the shoulders should have been indicated. Score very strictly, especially when Item 25 is zero. Do not credit if arms at their place of attachment occupy as much as one-half or more of the distance from the neck to the waist. The following sketch illustrates when Item 41 but *not* Item 42 scores:

(See also Item 25, a, e, h, for examples which credit Item 41 but *not* 42.)

Arms: **Profile:** The attachment of the arms must be indicated at a point approximately on the median line of the trunk, at a short distance below the neck, this point coinciding with the broadening of the trunk which indicates the chest and shoulders. If the arms extend from the line which outlines the back, or if the point of attachment reaches the base of the neck, or falls below the greatest expansion of the chest, the point is not credited. Credit Item 41 but *not* Item 42.

43. Clothing indicated

Clothing indicated by buttons or pockets on the simple ellipse, triangle, or trapezoid figure. Credit if there is definitely a skirt, even if no buttons or pockets are shown.

44. Sleeve I

Indicated by any means.

45. Sleeve II

Indicated by more than a simple cross line. Must show button, cuff, double line, puffed sleeve (long or short), or sleeve definitely wider than the arm which protrudes from it. Where a strap or strapless gown is clearly indicated, credit both Items 44 and 45. When hands are so placed that possible cuff is hidden, do *not* credit unless short sleeve is clearly indicated. *Note:* Be careful not to confuse bracelet or wristwatch with sleeve.

46. Neckline I

Any dress line at neck other than that produced by chin or jaw. Any crude single line, straight or semicircular. Distinguish carefully from necklace.

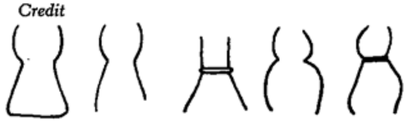
47. Neckline II: collar

Collar indicated. Neckline must be "V'd" or definitely shaped in some other manner.

48. Waist I

Whether or not a belt is shown, the direction of the body contour must change perceptibly at and/or below the waist. *If no belt or waist is drawn*, a gentle, continuous curve does not score; there must be an abrupt change in body line.

Credit

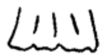


No Credit

49. Waist II
A distinct belt (two lines), sash, sweater, or blouse hem must be indicated by means ***better than a single horizontal line.***

50. Skirt "modeled" to indicate pleats or draping
Irregular hemline not sufficient; lines, shading, or sketching must appear.

Credit

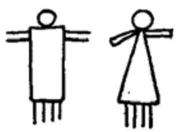


51. No transparencies in the figure
There must be a garment on the figure that is clear and complete. Clothing must show neckline, sleeves, skirt hem, or slacks. No body lines may show through clothes that would ordinarily conceal them.

52. Garb feminine
Young Children (under 8): Skirt must be a distinct feature, and the body must appear in *two* distinct segments.

Credit

No Credit



Older Children (8 and over): Credit any dress or skirt. Where slacks, breeches, or overalls are shown, credit only if the style of blouse or pants is distinctly feminine, apart from hair, face, or breast indication. Slacks may be judged by absence of fly and by placement of pockets.

53. Garb complete, without incongruities
Garb must contain all these elements: shoes, sleeves (hands must protrude), dress and neckline or sleeves, or skirt and blouse (or jacket). *Exceptions:* Slacks, blue jeans, sports garb, formal dress which may obscure shoes. These are credited.

54. Garb a definite "type"
Types may include: formal gown, sports garb (shorts, slacks), "school garb," "dress up," house dress (should include apron), or "suit" (jacket and skirt).

55. Trunk present
Any clear indication of the trunk, either one or two dimensional.

56. Trunk in proportion, two dimensions
Length of trunk greater than breadth. In drawings by younger children, where the trunk may not be clearly differentiated from the skirt, judge body area as including skirt.

57. Head-trunk proportion
Young Children (under 8): Score in relation to *body area,* excluding head when no differentiation between waist and terminus of trunk or no indication of skirt is shown.

Older Children (8 and over): Credit drawings that indicate a garment but do not suggest a waistline, if the head is no larger than one-fourth or smaller than one-eighth of the body (including garment) area.

Profile: Score more leniently. Judge more on the length of head in relation to the length of chest area. If two lengths are about equal, or if head is the shorter length but not less than one-fourth the chest length, credit the item.

58. Head: proportion
Full Face: Length of head greater than its width. Should show a general oval shape.

Profile: Same requirement as full-face drawing, but exclude hair in estimating width.

59. Limbs: proportion
Length of arms and legs greater than width. When arms score, credit the item even if feet are concealed by long dress.

60. Arms in proportion to trunk
Both arms *longer* than length of trunk from shoulder (or base of neck) to waist, but not more than twice this length.

Young Children (under 8): Arms must be equal to body length.

Older Children (8 and over): Credit drawings that portray dress or skirt if arm length is at least half of dress length (shoulder to hem of skirt) but not as long as hem.

61. Location of waist

This item evaluates child's ability to *locate* the waist. Waist located below one-third of total length of figure, crown to toe, but not below one-half of total length. (Crown is considered the top of the head, including hair but not hat.) Waistline must be indicated by belt, or by some distinct change in body contour. Do not credit when trunk and dress are indicated by uninterrupted curve, with no indication of waistline.

62. Dress area

Dress area below waist must be as large or larger than trunk area above waist but not more than twice as large (three times as large in profile). Credit if formal gown is clearly represented. For slacks, include the area occupied by the legs but not the feet. Define as waist a waist *line* however indicated, or estimate location from an obvious narrowing of body, or widening of hips. Do not credit in drawings by young children showing no trunk or body contours.

63. Motor coordination: junctures

All lines meet cleanly, without overlap or intervening space. Emphasis is on the juncture of lines, regardless of the character of lines.

64. Motor coordination: lines

Lines are firm, cleanly made, continuous and "controlled." If "sketchy" judge the basic character of the body lines created by the shorter pencil strokes. Both curved and straight lines must be handled with assurance. Do not credit in a drawing with extensive redrawing and erasures.

65. Superior motor coordination

Credit this point in all cases where Item 64 is achieved without redrawing or erasures, and where the total effect of lines is neat, clean, and "sure."

66. Directed lines and form: head outline

The drawing must show the contours of the head and/or face. Simple circle or ellipse to which projecting features have been added does not score.

No Credit

67. Directed lines and form: breast

Any attempt, by modeling or by contour, to indicate the feminine breast. In full-face drawings, credit strapless gown if top is curved.

Credit

68. Directed lines and form: hip contour

Full Face: Hips indicated by distinct convexity below waistline. This must occur on both sides. Note that wide, uniformly curved bell-shaped flaring skirt does *not* score.

Profile: Convexity must be indicated over hips and buttocks.

Credit

No Credit

69. Directed lines and form: arms taper

Wrist and/or forearm distinctly narrower than upper arm. Credit the point whether achieved by narrowing of sleeve or by shaping the bare arm. Where long, full sleeves are clearly indicated, credit this item.

70. Directed lines and form: calf of leg

Leg shaped better than a taper. Definite calf must be shown. Score strictly.

71. Directed lines and form: facial features

Facial features must be symmetrical in all respects. Eyes and mouth must be shown in two dimensions; nose may be indicated by dots.

Full Face: Features must be appropriately placed, regular and symmetrical, giving a clear appearance of the human form.

Profile: The eye must be regular in outline and located in the forward one-third of the head. The bridge of the nose must form an obtuse angle with the forehead. The scoring should be strict; a "cartoon" nose does not get credit.

Appendix B

Information Sheet and Informed Consent Form

JAMES COOK UNIVERSITY

TOWNSVILLE Queensland 4811 Australia Telephone: (07) 4781 4111

INFORMATION SHEET

Contact Details

Researcher:	Ms. Claire Maley	Claire.Maley@jcu.edu.au	Phone: 0402 459 446
Supervisor:	Prof. Trevor Bond	tbond@ied.edu.hk	Phone: (852) 2948 8473
Co-supervisor:	Dr. David Lake	David.Lake@jcu.edu.au	Phone: 47814918
Ethics Administrator:	Ms. Tina Langford	Tina.Langford@jcu.edu.au	Phone: 47814342; Fax: 47815521

Dear Parents / Guardians,

I am a PhD student at James Cook University. My PhD Research Project looks at children's human figure drawings and how these relate to their thinking. This research will be conducted under the supervision of Professor Trevor Bond, Head of Department of Educational Psychology, Counselling and Learning Needs at the Hong Kong Institute of Education, and Dr. David Lake, Lecturer at the School of Education, James Cook University.

I will ask children who take part in this research to do three individual drawings: one each of a man, a woman and themselves. This will take approximately 10-15 minutes. We will use a quiet, comfortable and air-conditioned area within the school. The aim is to collect three drawings from each child which are drawn to the best of his/her ability. As long as they draw a whole person each time (i.e. head to feet), there is no right or wrong type of drawing. I am interested in what the children *can* draw and understand; not what they cannot draw or understand. I will not be comparing results of different children. Once drawings are completed and collected, participating children will be identified by a unique code number, for confidentiality.

All drawings will be scored and analysed to indicate the children's levels of understanding. Also, children will be randomly selected to provide longitudinal data, to track the development of their drawings over time. This will require these participating children to complete the same three drawings every 3 to 6 months, for the duration of the Researcher's Doctoral studies (approximately 1.5 years).

Children may benefit from the focused engagement in their drawings which not only indicate levels of understanding but can act as stimulus for further development. Also, the drawings are not unrelated to the learning activities presented in primary school.

I will use this research as the data for my Doctoral Thesis, and may also use it in articles published in professional and academic literature. No information that might identify the school or the children will be included in any of these. To ensure that data remains confidential, all data collected (such as signed consent forms and drawings) will be stored securely in a locked cabinet at the School of Education at JCU.

If you have any questions about the nature of this project, or about how the information from it will be used, please contact either myself or my supervisors (details above). If you allow your child to participate in the project, and during the process have any questions about the way I have conducted the research, please contact the Ethics Administrator at JCU, Ms Tina Langford (details above). If you are willing for your child to participate, please complete, sign and return the attached Consent Form to the school by … 2007. Please retain this Information Sheet for your future reference.

Thank you,

Claire Maley

JAMES COOK UNIVERSITY

TOWNSVILLE Queensland 4811 Australia Telephone: (07) 4781 4111

INFORMED CONSENT FORM

Principal Investigator:	Ms. Claire Maley
Project Title:	A Cross-Cultural Examination of the Goodenough-Harris Drawing Test
School:	School of Education
Contact Details:	Claire.Maley@jcu.edu.au *or* 0402 459 446

I seek permission from you to allow me to (please tick if you are willing for me to do so):	**Yes**	**No**
1. Ask your child to draw three individual drawings: one each of a man, woman, and themselves. The drawings will take approximately 10 to 15 minutes in total, and will be conducted in a quiet, comfortable and air-conditioned area within the school. The task will require your child to manipulate a Number 2 pencil and draw, to the best of his/her ability, the three above-mentioned pictures on separate pages in a booklet.	☐	☐
2. Ask your child some informal questions to clarify any ambiguous aspects of his/her drawings, if necessary.	☐	☐
3. Ask your child to complete the same three drawing tasks detailed above, in intervals of approximately 3 to 6 months, for the duration of the Researchers Doctoral studies (approximately 1.5 years).	☐	☐

All drawings will be collected, de-identified, scored and analysed to determine the children's levels of understanding.

The results of the research will be used as the data for the researcher's Doctoral Thesis, and may also be used in articles published in professional and academic journals. No personal identifying information will be included in any of these. All data collected (such as signed Consent Forms and the children's drawings) will be stored separately in locked cabinets at the School of Education at JCU.

Participation in this Research Project will involve minimal interruption of your child's school experience, and he/she may benefit from the close attention to his/her thinking and drawings.

The aims of this study have been clearly explained to me and I understand what is wanted of me / my child. I know that taking part in this study is voluntary and I am aware that I can stop my child taking part in it at any time and he / she may refuse to complete the drawings or answer any questions. I understand that any information I / my child gives will be kept strictly confidential and that no names will be used to identify me / my child with this study without my approval.	
Child's Name:	
Parents / Guardians Name: *(printed)*	
Signature:	**Date:**

Appendix C

Ethics

ETHICS REVIEW COMMITTEE
Human Ethics Committee
APPROVAL FOR RESEARCH OR TEACHING INVOLVING HUMAN SUBJECTS

PRINCIPAL INVESTIGATOR	**Ms Claire Maley**				
SUPERVISORS	Prof Trevor Bond & Dr David Lake (Education)				
SCHOOL	Education				
PROJECT TITLE	**A cross-cultural examination of the Goodenough-Harris Drawing test**				
APPROVAL DATE	10 Oct 2006	**EXPIRY DATE**	30 Nov 2009	CATEGORY	**1**

This project has been allocated Ethics Approval Number with the **following conditions**:	H	**2450**

All subsequent records and correspondence relating to this project must refer to this number.
That there is **NO** departure from the approved protocols unless prior approval has been sought from the Human Ethics Committee.

The Principal Investigator **must advise** the responsible Ethics Monitor appointed by the Ethics Review Committee:

periodically of the progress of the project;
when the project is completed, suspended or prematurely terminated for any reason;
notify within 48 hours of any adverse effects on participants occur; and if any unforeseen events occur that might affect continued ethical acceptability of the project.

4. In compliance with the National Health and Medical Research Council (NHMRC) *"National Statement on Ethical Conduct in Research Involving Humans, 1999*, it is **MANDATORY** that you provide an annual report on the progress and conduct of your project. This report must detail compliance with approvals granted and any unexpected events or serious adverse effects that may have occurred during the study.

Human Ethics Advisor:	Vick, Dr Malcolm
Email:	malcolm.vick@jcu.edu.au
ASSESSED AT MEETING **APPROVED**	Date: ***27 Sep 2006*** Date: ***10 Oct 2006***

Professor Peter Leggat Chair, Human Ethics Committee	
Tina Langford, Ethics Officer, **Research Office** Tina.Langford@jcu.edu.au	Date: 13 October 2006

Appendix D

50 Common Items Draw-a-Man & Draw-a-Woman

Rasch Analysis Output

Variable Maps

Draw-a-Man

```
Persons - MAP - Items
              <more>|<rare>
    6               +  45
                    |T
                    |
                    |  19    50
    5               +
                    |
                    |  44
                    |  12
    4               +  49
                    |  11
                    |
                    |  28
    3            .  +S 43
                 .  |  22
                 .  |  18
                .# T|  14    23
    2           .#  +  48
                 .  |
               .##  |  15    33    39    7
             #####  |  5
    1          .## S+
             .####  |  21    3     34
              ####  |  26
          ########  |  32    9
    0          .##  +M 27
              ####  |  17    30    47
              .### M|  24    41    6
              .###  |  36
   -1    .#########  +  46
            .#####  |
             .####  |  2     20    8
             .####  |  13
   -2      .####### S+  31    40    42
                .#  |  38
                .#  |
                    |  16
   -3           .#  +S
                 .  |
                   T|
                 .  |
   -4            .  +
                    |  10
                 #  |
                    |
   -5               +  35
                    |  25    37
                    |
                 .  |T
   -6               +
                    |
                    |  29
                    |
   -7               +
                    |  4
                    |
                    |
   -8               +  1
              <less>|<frequ>
 EACH '#' IS 3.
```

```
 Persons - MAP - Items
              <more>|<rare>
   6                +
                    |
                    |T
                    |  49
   5                +
                    |  50
                    |  19
                    |
   4                +  28
                    |  45
                 .  |
                    |
   3                +  22
                 .  |S
                 .  |  12    23    44    48
                 . T|  14
   2           ###  +  11    18
                 #  |  15
                .#  |  7
           .######  |
   1          .### S+  21    33    43    5
                .#  |  26    3     30
             .####  |  39
                .#  |  34    9
   0      .#######  +M 27    32    41
            .#####  |  24    36    47    6
              .### M|
         .########  |  2
  -1         #####  +  46
             .####  |  17    20    40    8
        .#########  |  13
              .### S|
  -2           .##  +  31
              .###  |  42
                .#  |
                    |S 38
  -3             .  +
                .# T|
                    |
                    |  16    35
  -4             .  +
                    |
                 .  |
                    |  10
  -5                +  37
                    |  25
                    |T 29
                    |
  -6                +
                    |
                    |
                    |
  -7                +  1     4
              <less>|<frequ>
EACH '#' IS 3.
```

Summary Statistics

Draw-a-Man

SUMMARY OF 246 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	21.5	50.0	-.53	.45	1.01	.0	.97	.1
S.D.	7.3	.0	1.43	.05	.26	.9	1.19	1.0
MAX.	40.0	50.0	3.07	.82	3.07	3.1	9.90	9.1
MIN.	4.0	50.0	-5.75	.41	.51	-2.5	.08	-1.2

REAL RMSE	.48	ADJ.SD	1.34	SEPARATION	2.83	Person RELIABILITY .89
MODEL RMSE	.45	ADJ.SD	1.35	SEPARATION	3.01	Person RELIABILITY .90
S.E. OF Person MEAN = .09						

LACKING RESPONSES: 75 Persons
VALID RESPONSES: 99.9%
Person RAW SCORE-TO-MEASURE CORRELATION = .99 (approximate due to missing data)
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .89 (approximate due to missing data)

SUMMARY OF 48 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	105.1	246.0	.00	.27	.99	-.1	1.34	.1
S.D.	78.7	.0	2.92	.19	.17	1.9	1.88	2.3
MAX.	245.0	246.0	5.18	1.03	1.57	5.0	9.90	7.5
MIN.	2.0	246.0	-7.21	.15	.71	-4.7	.06	-3.0

REAL RMSE	.34	ADJ.SD	2.90	SEPARATION	8.59	Item RELIABILITY .99
MODEL RMSE	.33	ADJ.SD	2.90	SEPARATION	8.68	Item RELIABILITY .99
S.E. OF Item MEAN = .43						

MAXIMUM EXTREME SCORE: 1 Items
MINIMUM EXTREME SCORE: 1 Items

SUMMARY OF 246 MEASURED Persons

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	21.9	50.0	-.47	.43	1.01	.0	.97	.1
S.D.	7.8	.0	1.39	.05	.25	1.0	1.04	.7
MAX.	43.0	50.0	3.40	.69	2.03	3.5	9.90	5.0
MIN.	6.0	50.0	-4.47	.39	.56	-2.7	.15	-.9

```
| REAL RMSE    .46  ADJ.SD    1.31  SEPARATION  2.85  Person RELIABILITY  .89 |
|MODEL RMSE    .44  ADJ.SD    1.32  SEPARATION  3.03  Person RELIABILITY  .90 |
| S.E. OF Person MEAN = .09                                                   |
```

LACKING RESPONSES: 75 Persons
VALID RESPONSES: 99.9%
Person RAW SCORE-TO-MEASURE CORRELATION = .99 (approximate due to missing data)
CRONBACH ALPHA (KR-20) Person RAW SCORE RELIABILITY = .90 (approximate due to missing data)

SUMMARY OF 49 MEASURED Items

	RAW SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	105.1	246.0	.00	.26	.99	-.1	.98	.0
S.D.	76.6	.0	2.74	.17	.20	2.3	.54	2.2
MAX.	245.0	246.0	5.24	1.01	1.61	7.0	2.59	7.7
MIN.	2.0	246.0	-6.93	.15	.72	-4.4	.16	-3.3

```
| REAL RMSE    .31  ADJ.SD    2.72  SEPARATION  8.73  Item   RELIABILITY  .99 |
|MODEL RMSE    .31  ADJ.SD    2.72  SEPARATION  8.84  Item   RELIABILITY  .99 |
| S.E. OF Item MEAN = .40                                                     |
```

MINIMUM EXTREME SCORE: 1 Items
UMEAN=.000 USCALE=1.000

Item Statistics

Draw-a-Man

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
45	0	246	7.10	1.83	Maximum Estimated Measure			
19	2	246	5.18	0.72	0.97	0.20	9.90	5.00
50	2	246	5.18	0.72	0.91	0.10	0.14	-1.70
44	4	246	4.45	0.52	0.99	0.10	0.35	-1.00
12	5	246	4.21	0.46	0.92	-0.10	0.28	-1.20
49	6	246	4.02	0.43	1.02	0.20	2.22	1.60
11	7	246	3.85	0.40	0.95	0.00	2.95	2.20
28	11	246	3.34	0.32	1.12	0.60	9.90	7.50
43	15	246	2.97	0.28	0.87	-0.60	0.36	-1.70
22	19	246	2.68	0.26	0.85	-0.80	0.49	-1.40
18	24	246	2.39	0.23	0.98	-0.10	0.69	-0.80
14	25	246	2.33	0.23	1.26	1.60	1.60	1.50
23	27	246	2.23	0.22	0.95	-0.30	0.75	-0.70
48	33	246	1.95	0.21	0.90	-0.80	1.07	0.30
33	42	246	1.60	0.19	0.92	-0.70	0.62	-1.50
7	43	246	1.56	0.19	0.79	-2.10	0.49	-2.30
39	46	246	1.46	0.18	1.57	4.80	2.30	3.80
15	48	246	1.39	0.18	1.08	0.80	1.03	0.20
5	52	246	1.26	0.18	1.06	0.70	0.99	0.00
34	71	246	0.71	0.16	0.83	-2.20	0.66	-2.00
3	73	246	0.66	0.16	0.86	-1.80	0.77	-1.30
21	74	246	0.64	0.16	1.32	3.70	1.46	2.30
26	77	246	0.56	0.16	0.78	-3.10	0.65	-2.20
32	89	246	0.26	0.16	0.74	-3.80	0.60	-3.00
9	94	246	0.14	0.15	0.95	-0.70	0.88	-0.80
27	102	246	-0.05	0.15	0.71	-4.70	0.63	-2.90
30	111	246	-0.25	0.15	0.77	-3.60	0.70	-2.40
17	112	246	-0.28	0.15	0.96	-0.60	0.89	-0.80
47	116	246	-0.37	0.15	1.12	1.80	1.16	1.20
6	123	246	-0.52	0.15	1.08	1.10	1.29	2.10
41	123	246	-0.52	0.15	1.12	1.70	1.11	0.90
24	126	246	-0.59	0.15	1.00	0.10	0.92	-0.60
36	133	246	-0.75	0.15	0.91	-1.40	0.86	-1.10
46	147	246	-1.07	0.15	1.20	2.90	1.28	2.00
8	162	246	-1.42	0.16	1.11	1.60	1.25	1.60
2	163	246	-1.45	0.16	1.19	2.60	1.81	4.40
20	168	246	-1.57	0.16	1.06	0.90	1.15	1.00
13	174	246	-1.72	0.16	1.10	1.30	1.43	2.30
40	181	246	-1.91	0.17	1.45	5.00	2.54	5.90
42	187	246	-2.08	0.17	0.85	-1.80	0.60	-2.20
31	188	246	-2.11	0.17	0.84	-1.90	0.63	-2.00
38	194	246	-2.29	0.18	0.85	-1.60	0.63	-1.80
16	207	246	-2.74	0.20	0.93	-0.50	0.69	-1.10
10	233	246	-4.25	0.31	0.83	-0.70	1.09	0.40

35	239	246	-5.02	0.42	0.79	-0.50	0.39	-1.00
25	240	246	-5.21	0.45	0.85	-0.30	1.51	0.90
37	240	246	-5.21	0.45	0.96	0.00	0.37	-1.00
29	244	246	-6.47	0.74	1.04	0.30	0.35	-1.00
4	245	246	-7.21	1.03	0.98	0.30	0.06	-2.30
1	246	246	-8.46	1.84	Minimum Estimated Measure			
Mean	105.9	246	-0.03	0.34	0.99	-0.10	1.34	0.10
SD	81.1	0	3.26	0.36	0.17	1.90	1.88	2.30
r	0.99							

Draw-a-Woman

Item #	Raw Score	Count	Difficulty Estimate	Model S.E.	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
49	2	246	5.24	0.72	1.01	0.20	0.38	-0.90
50	3	246	4.82	0.59	0.88	0.00	0.16	-1.60
19	4	246	4.52	0.52	0.99	0.10	1.41	0.70
28	6	246	4.08	0.43	1.02	0.20	0.46	-0.80
45	9	246	3.63	0.35	1.03	0.20	2.26	1.80
22	17	246	2.88	0.27	1.01	0.10	0.52	-1.10
44	22	246	2.56	0.24	0.89	-0.70	0.43	-1.70
12	23	246	2.50	0.24	0.81	-1.20	0.36	-2.10
48	23	246	2.50	0.24	1.00	0.10	0.96	0.00
23	25	246	2.39	0.23	1.07	0.50	1.15	0.50
14	29	246	2.19	0.22	1.18	1.30	1.55	1.50
18	33	246	2.01	0.21	0.81	-1.50	0.54	-1.60
11	34	246	1.97	0.21	0.83	-1.40	0.53	-1.70
15	41	246	1.69	0.19	0.98	-0.10	1.00	0.10
7	50	246	1.38	0.18	0.84	-1.70	0.63	-1.70
5	62	246	1.01	0.17	0.90	-1.10	1.42	1.90
33	63	246	0.98	0.17	0.91	-1.00	0.69	-1.70
43	64	246	0.95	0.17	0.75	-3.10	0.57	-2.50
21	66	246	0.90	0.17	1.28	3.10	1.43	2.00
30	68	246	0.84	0.17	1.61	6.10	2.25	5.00
26	69	246	0.82	0.16	0.73	-3.50	0.54	-3.00
3	72	246	0.74	0.16	0.75	-3.30	0.66	-2.20
39	82	246	0.48	0.16	1.57	6.30	1.84	4.50
9	87	246	0.35	0.16	0.99	-0.10	0.89	-0.80
34	90	246	0.28	0.16	0.75	-3.70	0.62	-3.10
41	99	246	0.07	0.15	1.02	0.40	1.36	2.60
27	101	246	0.02	0.15	0.72	-4.40	0.64	-3.30
32	107	246	-0.12	0.15	0.80	-3.20	0.72	-2.50
36	112	246	-0.23	0.15	0.92	-1.10	0.97	-0.20
24	113	246	-0.25	0.15	1.09	1.30	1.09	0.80
6	116	246	-0.32	0.15	1.01	0.10	1.09	0.80
47	117	246	-0.34	0.15	0.92	-1.30	0.87	-1.20
2	140	246	-0.86	0.15	1.08	1.20	1.01	0.20
46	142	246	-0.90	0.15	1.13	2.00	1.30	2.30
17	154	246	-1.18	0.15	0.91	-1.40	0.78	-1.70
8	161	246	-1.34	0.15	1.15	2.10	1.24	1.60
20	162	246	-1.37	0.16	1.15	2.10	1.43	2.70
40	162	246	-1.37	0.16	1.56	7.00	2.59	7.70
13	166	246	-1.47	0.16	1.18	2.40	1.45	2.60
31	187	246	-2.02	0.17	0.82	-2.30	0.70	-1.60
42	193	246	-2.20	0.18	0.81	-2.20	0.59	-2.00
38	210	246	-2.79	0.20	0.90	-0.80	1.08	0.40
16	228	246	-3.72	0.27	0.90	-0.40	0.54	-1.00
35	229	246	-3.80	0.27	0.89	-0.50	1.05	0.30
10	239	246	-4.85	0.40	1.00	0.10	0.63	-0.50
37	240	246	-5.03	0.43	0.94	-0.10	1.95	1.40

25	241	246	-5.23	0.47	0.94	0.00	0.92	0.10
29	242	246	-5.48	0.52	0.94	0.00	0.33	-1.10
4	245	246	-6.93	1.01	1.01	0.30	0.23	-1.30
1	246	246	-8.15	1.83	Minimum Estimated Measure			
Mean	107.9	246	-0.16	0.29	0.99	-0.10	0.98	0.00
SD	78.3	0	2.94	0.28	0.20	2.30	0.54	2.20
r	0.99							

Appendix E

Human Figure Drawing Continuum

Human Figure Drawing Continuum Person Statistics

Case	Raw		Ability	Model	Infit		Outfit	
#	Score	Count	Estimate	S.E.	MNSQ	ZSTD	MNSQ	ZSTD
100 F 3	36	45	3.34	0.54	1.61	1.80	4.05	2.00
177 F 1	35	45	3.06	0.53	0.99	0.10	0.59	-0.20
174 F 1	34	45	2.79	0.51	0.59	-1.60	0.26	-0.90
177 F 3	34	45	2.79	0.51	1.69	2.10	2.66	1.60
174 F 3	33	45	2.53	0.50	0.76	-0.80	0.78	0.00
109 M 1	32	45	2.28	0.49	1.04	0.20	1.57	0.90
174 F 2	32	45	2.28	0.49	0.69	-1.20	0.35	-0.90
177 F 2	32	45	2.28	0.49	0.81	-0.60	0.41	-0.70
150 M 3	32	45	2.28	0.49	0.93	-0.20	0.46	-0.60
100 F 1	31	45	2.04	0.49	0.98	0.00	0.68	-0.20
196 F 1	31	45	2.04	0.49	1.20	0.80	1.46	0.80
205 F 2	31	45	2.04	0.49	1.29	1.10	1.61	0.90
179 F 3	31	45	2.04	0.49	1.60	2.00	1.94	1.20
102 F 1	30	45	1.81	0.48	1.20	0.80	0.90	0.10
108 M 1	30	45	1.81	0.48	1.08	0.40	0.86	0.10
125 F 2	30	45	1.81	0.48	0.82	-0.60	0.44	-0.70
196 F 2	30	45	1.81	0.48	0.68	-1.30	0.46	-0.70
204 F 2	30	45	1.81	0.48	1.17	0.70	1.02	0.30
112 F 3	30	45	1.81	0.48	0.72	-1.10	0.36	-0.90
196 F 3	30	45	1.81	0.48	0.94	-0.20	0.87	0.10
107 F 1	29	45	1.58	0.47	0.67	-1.40	0.36	-0.90
113 F 1	29	45	1.58	0.47	0.55	-2.00	0.28	-1.10
116 F 1	29	45	1.58	0.47	0.77	-0.80	0.45	-0.70
122 F 1	29	45	1.58	0.47	0.92	-0.20	0.59	-0.40
112 F 2	29	45	1.58	0.47	0.97	0.00	0.85	0.10
193 F 3	29	45	1.58	0.47	0.93	-0.20	1.12	0.40
112 F 1	28	45	1.35	0.47	0.78	-0.80	0.74	-0.10
125 F 1	28	45	1.35	0.47	1.23	0.90	1.50	0.80
201 F 1	28	45	1.35	0.47	0.86	-0.50	0.52	-0.50
206 M 1	28	45	1.35	0.47	1.19	0.80	0.75	-0.10
101 F 2	28	45	1.35	0.47	1.49	1.70	1.64	0.90
166 F 2	28	45	1.35	0.47	1.06	0.30	0.71	-0.10
172 M 3	28	45	1.35	0.47	0.88	-0.40	0.66	-0.20
178 M 3	28	45	1.35	0.47	1.15	0.60	0.74	-0.10
146 F 1	27	45	1.14	0.47	0.82	-0.60	0.77	0.00
150 M 1	27	45	1.14	0.47	1.61	2.10	1.70	1.00
195 F 1	27	45	1.14	0.47	0.85	-0.50	0.45	-0.60
204 F 1	27	45	1.14	0.47	0.84	-0.60	0.55	-0.40
167 F 2	27	45	1.14	0.47	1.12	0.50	0.73	-0.10
166 F 3	27	45	1.14	0.47	1.13	0.60	0.86	0.10
175 F 3	27	45	1.14	0.47	0.57	-1.90	0.36	-0.80
195 F 3	27	45	1.14	0.47	1.32	1.20	0.97	0.20
166 F 1	26	45	0.92	0.46	0.90	-0.30	0.63	-0.20
121 F 2	26	45	0.92	0.46	0.91	-0.30	0.54	-0.40
123 F 2	26	45	0.92	0.46	0.92	-0.20	0.78	0.00
175 F 2	26	45	0.92	0.46	0.78	-0.90	0.48	-0.50

121 F 3	26	45	0.92	0.46	1.03	0.20	0.66	-0.20
167 F 3	26	45	0.92	0.46	1.35	1.30	1.39	0.70
192 M 3	26	45	0.92	0.46	1.20	0.80	1.05	0.30
153 F 1	25	45	0.71	0.46	0.96	-0.10	0.57	-0.30
155 F 1	25	45	0.71	0.46	1.47	1.80	5.73	3.00
175 F 1	25	45	0.71	0.46	0.84	-0.60	0.60	-0.20
113 F 2	25	45	0.71	0.46	0.99	0.10	0.74	0.00
120 F 2	25	45	0.71	0.46	0.96	-0.10	0.99	0.30
122 F 2	25	45	0.71	0.46	0.74	-1.10	0.41	-0.60
150 M 2	25	45	0.71	0.46	1.26	1.00	0.95	0.20
172 M 2	25	45	0.71	0.46	0.90	-0.30	0.52	-0.40
125 F 3	25	45	0.71	0.46	0.73	-1.10	0.47	-0.50
120 F 1	24	45	0.50	0.46	0.98	0.00	0.66	-0.10
121 F 1	24	45	0.50	0.46	0.86	-0.50	0.48	-0.40
124 F 1	24	45	0.50	0.46	1.28	1.10	1.25	0.60
167 F 1	24	45	0.50	0.46	1.26	1.10	1.48	0.80
193 F 1	24	45	0.50	0.46	1.63	2.30	1.57	0.80
116 F 2	24	45	0.50	0.46	0.97	0.00	0.66	-0.10
126 F 2	24	45	0.50	0.46	1.43	1.60	1.04	0.40
155 F 2	24	45	0.50	0.46	1.15	0.70	0.89	0.20
156 F 2	24	45	0.50	0.46	1.21	0.90	0.79	0.10
170 M 2	24	45	0.50	0.46	0.99	0.00	0.61	-0.20
179 F 2	24	45	0.50	0.46	0.88	-0.40	0.51	-0.30
180 M 3	24	45	0.50	0.46	0.77	-0.90	0.49	-0.30
101 F 1	23	45	0.29	0.46	1.06	0.30	7.56	3.30
179 F 1	23	45	0.29	0.46	0.81	-0.70	0.45	-0.40
192 M 1	23	45	0.29	0.46	0.85	-0.60	0.52	-0.30
124 F 2	23	45	0.29	0.46	1.22	0.90	0.87	0.20
146 F 2	23	45	0.29	0.46	1.01	0.10	0.66	-0.10
173 M 2	23	45	0.29	0.46	0.84	-0.60	0.76	0.10
192 M 2	23	45	0.29	0.46	0.79	-0.90	0.62	-0.10
193 F 2	23	45	0.29	0.46	1.46	1.80	1.43	0.70
101 F 3	23	45	0.29	0.46	0.94	-0.20	0.67	-0.10
116 F 3	23	45	0.29	0.46	1.08	0.40	0.63	-0.10
122 F 3	23	45	0.29	0.46	0.98	0.00	0.66	-0.10
123 F 3	23	45	0.29	0.46	1.06	0.30	0.73	0.00
105 F 1	22	45	0.08	0.45	1.32	1.30	1.83	1.00
119 M 1	22	45	0.08	0.45	0.88	-0.40	0.51	-0.30
154 F 1	22	45	0.08	0.45	0.90	-0.30	0.73	0.00
202 M 1	22	45	0.08	0.45	1.01	0.10	0.70	0.00
205 F 1	22	45	0.08	0.45	1.02	0.20	0.66	-0.10
151 F 2	22	45	0.08	0.45	1.25	1.10	1.45	0.70
178 M 2	22	45	0.08	0.45	0.95	-0.20	0.58	-0.20
202 M 2	22	45	0.08	0.45	0.96	-0.10	0.67	-0.10
139 F 3	22	45	0.08	0.45	0.79	-0.90	0.58	-0.20
126 F 1	21	45	-0.12	0.46	0.80	-0.80	0.46	-0.40
156 F 1	21	45	-0.12	0.46	0.95	-0.10	0.53	-0.30
172 M 1	21	45	-0.12	0.46	0.88	-0.50	0.48	-0.30
180 M 2	21	45	-0.12	0.46	0.77	-1.00	0.44	-0.40
199 F 2	21	45	-0.12	0.46	1.16	0.80	0.73	0.00

110 M 3	21	45	-0.12	0.46	0.91	-0.30	0.48	-0.30
164 F 3	21	45	-0.12	0.46	1.45	1.80	1.18	0.50
199 F 3	21	45	-0.12	0.46	1.03	0.20	0.57	-0.20
202 M 3	21	45	-0.12	0.46	1.25	1.10	0.90	0.20
151 F 1	20	45	-0.33	0.46	0.99	0.00	0.80	0.10
157 F 1	20	45	-0.33	0.46	1.12	0.60	0.92	0.30
178 M 1	20	45	-0.33	0.46	1.17	0.80	0.67	0.00
114 M 2	20	45	-0.33	0.46	1.13	0.60	0.91	0.20
117 M 2	20	45	-0.33	0.46	1.41	1.70	1.17	0.50
118 M 2	20	45	-0.33	0.46	1.07	0.40	0.64	-0.10
147 F 2	20	45	-0.33	0.46	0.96	-0.10	0.71	0.00
181 M 2	20	45	-0.33	0.46	0.96	-0.10	0.54	-0.20
115 M 3	20	45	-0.33	0.46	1.03	0.20	0.59	-0.10
135 F 3	20	45	-0.33	0.46	0.86	-0.60	0.86	0.20
141 F 3	20	45	-0.33	0.46	0.76	-1.10	0.41	-0.40
111 M 1	19	45	-0.54	0.46	0.94	-0.20	0.56	-0.10
145 F 1	19	45	-0.54	0.46	1.17	0.80	0.78	0.10
173 M 1	19	45	-0.54	0.46	1.20	0.90	0.86	0.20
180 M 1	19	45	-0.54	0.46	1.08	0.40	1.44	0.70
118 M 3	19	45	-0.54	0.46	0.95	-0.20	0.78	0.10
104 F 1	18	45	-0.75	0.46	0.92	-0.30	0.60	-0.10
127 F 1	18	45	-0.75	0.46	1.00	0.10	0.63	-0.10
141 F 1	18	45	-0.75	0.46	1.05	0.30	0.68	0.00
181 M 1	18	45	-0.75	0.46	0.72	-1.30	0.37	-0.50
135 F 2	18	45	-0.75	0.46	0.86	-0.60	0.60	-0.10
154 F 2	18	45	-0.75	0.46	0.79	-0.90	0.51	-0.20
182 M 2	18	45	-0.75	0.46	1.22	1.00	0.84	0.20
194 F 2	18	45	-0.75	0.46	0.82	-0.80	0.44	-0.30
140 M 3	18	45	-0.75	0.46	0.90	-0.40	0.49	-0.30
194 F 3	18	45	-0.75	0.46	0.78	-1.00	0.41	-0.40
117 M 1	17	45	-0.97	0.47	1.01	0.10	0.88	0.30
118 M 1	17	45	-0.97	0.47	0.98	0.00	0.55	-0.10
138 F 1	17	45	-0.97	0.47	1.49	1.90	1.47	0.70
165 F 1	17	45	-0.97	0.47	0.61	-1.90	0.30	-0.50
194 F 1	17	45	-0.97	0.47	0.60	-1.90	0.29	-0.60
199 F 1	17	45	-0.97	0.47	1.41	1.60	1.14	0.50
110 M 2	17	45	-0.97	0.47	1.21	0.90	0.90	0.30
115 M 2	17	45	-0.97	0.47	1.08	0.40	0.63	0.00
136 F 2	17	45	-0.97	0.47	0.73	-1.20	0.49	-0.20
148 M 2	17	45	-0.97	0.47	0.85	-0.60	0.48	-0.20
152 F 2	17	45	-0.97	0.47	1.01	0.10	1.19	0.50
163 M 2	17	45	-0.97	0.47	0.95	-0.10	0.64	0.00
120 F 3	17	45	-0.97	0.47	1.05	0.30	0.58	-0.10
147 F 3	17	45	-0.97	0.47	0.55	-2.20	0.28	-0.60
106 M 1	16	45	-1.20	0.48	1.04	0.30	0.57	-0.10
110 M 1	16	45	-1.20	0.48	0.88	-0.40	0.54	-0.10
135 F 1	16	45	-1.20	0.48	1.76	2.70	5.53	2.40
147 F 1	16	45	-1.20	0.48	0.64	-1.60	0.31	-0.50
182 M 1	16	45	-1.20	0.48	1.16	0.70	0.88	0.30
197 M 1	16	45	-1.20	0.48	1.33	1.30	1.09	0.50

103 M 2	16	45	-1.20	0.48	0.96	-0.10	0.53	-0.10
157 F 2	16	45	-1.20	0.48	0.71	-1.30	0.35	-0.40
158 F 2	16	45	-1.20	0.48	0.53	-2.30	0.26	-0.60
165 F 2	16	45	-1.20	0.48	0.56	-2.10	0.27	-0.60
129 M 3	16	45	-1.20	0.48	1.06	0.30	1.27	0.60
152 F 3	16	45	-1.20	0.48	0.86	-0.50	1.17	0.50
163 M 3	16	45	-1.20	0.48	1.07	0.40	0.76	0.10
165 F 3	16	45	-1.20	0.48	0.66	-1.50	0.32	-0.50
173 M 3	16	45	-1.20	0.48	1.14	0.60	1.08	0.40
176 M 3	16	45	-1.20	0.48	0.94	-0.10	0.65	0.00
114 M 1	15	45	-1.43	0.49	0.81	-0.70	0.48	-0.20
128 M 1	15	45	-1.43	0.49	0.92	-0.20	0.43	-0.30
132 M 1	15	45	-1.43	0.49	0.74	-1.10	0.35	-0.40
133 F 1	15	45	-1.43	0.49	1.27	1.10	0.70	0.10
161 F 1	15	45	-1.43	0.49	0.72	-1.10	0.34	-0.40
163 M 1	15	45	-1.43	0.49	1.02	0.20	0.51	-0.10
170 M 1	15	45	-1.43	0.49	1.24	1.00	0.91	0.30
127 F 2	15	45	-1.43	0.49	1.24	1.00	1.46	0.70
132 M 2	15	45	-1.43	0.49	0.74	-1.10	0.36	-0.40
171 F 2	15	45	-1.43	0.49	0.68	-1.30	0.32	-0.50
159 F 3	15	45	-1.43	0.49	0.69	-1.30	0.33	-0.40
168 F 3	15	45	-1.43	0.49	0.88	-0.40	0.45	-0.20
171 F 3	15	45	-1.43	0.49	0.92	-0.20	0.44	-0.20
182 M 3	15	45	-1.43	0.49	0.75	-1.00	0.38	-0.30
184 M 3	15	45	-1.43	0.49	1.03	0.20	0.53	-0.10
185 F 3	15	45	-1.43	0.49	0.57	-2.00	0.27	-0.60
103 M 1	14	45	-1.67	0.50	0.57	-1.90	0.26	-0.50
115 M 1	14	45	-1.67	0.50	1.11	0.50	1.52	0.80
123 F 1	14	45	-1.67	0.50	0.92	-0.20	0.42	-0.30
129 M 1	14	45	-1.67	0.50	1.34	1.30	2.69	1.40
136 F 1	14	45	-1.67	0.50	1.32	1.20	2.46	1.30
183 F 1	14	45	-1.67	0.50	1.01	0.10	0.46	-0.20
184 M 1	14	45	-1.67	0.50	0.80	-0.70	0.36	-0.40
198 M 1	14	45	-1.67	0.50	0.90	-0.30	0.43	-0.20
203 M 1	14	45	-1.67	0.50	0.91	-0.30	0.43	-0.20
133 F 2	14	45	-1.67	0.50	0.82	-0.70	0.37	-0.30
137 M 2	14	45	-1.67	0.50	1.05	0.30	0.49	-0.20
139 F 2	14	45	-1.67	0.50	1.37	1.40	1.09	0.50
141 F 2	14	45	-1.67	0.50	0.75	-0.90	0.46	-0.20
149 M 2	14	45	-1.67	0.50	1.14	0.60	1.80	1.00
153 F 2	14	45	-1.67	0.50	1.19	0.80	1.43	0.70
161 F 2	14	45	-1.67	0.50	0.56	-1.90	0.26	-0.60
176 M 2	14	45	-1.67	0.50	1.35	1.30	1.08	0.50
191 F 2	14	45	-1.67	0.50	0.80	-0.70	0.36	-0.40
131 M 3	14	45	-1.67	0.50	1.12	0.50	0.56	-0.10
160 F 3	14	45	-1.67	0.50	0.59	-1.70	0.29	-0.50
188 M 3	14	45	-1.67	0.50	1.37	1.40	1.46	0.70
139 F 1	13	45	-1.93	0.52	1.10	0.40	0.51	-0.10
140 M 1	13	45	-1.93	0.52	1.34	1.20	1.70	0.90
142 M 1	13	45	-1.93	0.52	0.95	-0.10	0.62	0.00
148 M 1	13	45	-1.93	0.52	0.89	-0.30	0.41	-0.30
158 F 1	13	45	-1.93	0.52	0.79	-0.70	0.35	-0.40
129 M 2	13	45	-1.93	0.52	1.01	0.10	2.10	1.10
169 F 2	13	45	-1.93	0.52	0.67	-1.30	0.28	-0.50
186 M 2	13	45	-1.93	0.52	0.97	0.00	0.48	-0.20

187 F 2	13	45	-1.93	0.52	0.64	-1.40	0.27	-0.60
189 M 2	13	45	-1.93	0.52	1.06	0.30	0.93	0.30
190 F 2	13	45	-1.93	0.52	1.20	0.80	0.58	-0.10
169 F 3	13	45	-1.93	0.52	0.99	0.00	0.47	-0.20
189 M 3	13	45	-1.93	0.52	1.04	0.20	2.17	1.20
191 F 3	13	45	-1.93	0.52	1.06	0.30	0.94	0.30
149 M 1	12	45	-2.20	0.54	0.82	-0.50	0.44	-0.30
152 F 1	12	45	-2.20	0.54	0.50	-1.90	0.20	-0.70
159 F 1	12	45	-2.20	0.54	1.10	0.40	0.76	0.10
185 F 1	12	45	-2.20	0.54	0.86	-0.40	0.36	-0.40
186 M 1	12	45	-2.20	0.54	1.35	1.20	0.75	0.10
119 M 2	12	45	-2.20	0.54	0.89	-0.30	0.42	-0.30
128 M 2	12	45	-2.20	0.54	0.71	-1.00	0.27	-0.60
142 M 2	12	45	-2.20	0.54	0.90	-0.20	0.42	-0.30
159 F 2	12	45	-2.20	0.54	0.87	-0.30	0.38	-0.40
160 F 2	12	45	-2.20	0.54	0.88	-0.30	0.36	-0.40
168 F 2	12	45	-2.20	0.54	0.80	-0.60	0.35	-0.40
185 F 2	12	45	-2.20	0.54	0.99	0.10	0.42	-0.30
188 M 2	12	45	-2.20	0.54	1.34	1.10	7.16	2.90
143 M 3	12	45	-2.20	0.54	1.08	0.40	0.54	-0.10
157 F 3	12	45	-2.20	0.54	0.92	-0.20	0.46	-0.20
134 M 1	11	45	-2.50	0.56	1.02	0.20	0.81	0.20
137 M 1	11	45	-2.50	0.56	0.98	0.00	0.41	-0.30
164 F 1	11	45	-2.50	0.56	1.74	2.00	9.90	5.20
171 F 1	11	45	-2.50	0.56	0.62	-1.20	0.24	-0.60
176 M 1	11	45	-2.50	0.56	1.01	0.20	0.46	-0.20
187 F 1	11	45	-2.50	0.56	1.59	1.60	1.27	0.60
190 F 1	11	45	-2.50	0.56	0.95	0.00	0.46	-0.20
184 M 2	11	45	-2.50	0.56	1.04	0.20	0.50	-0.20
162 F 1	10	45	-2.84	0.59	1.89	2.00	9.90	9.90
168 F 1	10	45	-2.84	0.59	0.82	-0.40	0.41	-0.30
169 F 1	10	45	-2.84	0.59	0.91	-0.10	0.47	-0.20
131 M 1	9	45	-3.21	0.63	1.26	0.70	0.58	0.00
143 M 1	9	45	-3.21	0.63	0.59	-1.00	0.19	-0.70
144 M 1	9	45	-3.21	0.63	0.80	-0.40	0.33	-0.40
188 M 1	9	45	-3.21	0.63	0.71	-0.60	0.31	-0.40
200 M 1	9	45	-3.21	0.63	1.86	1.80	1.01	0.40
143 M 2	9	45	-3.21	0.63	0.67	-0.70	0.48	-0.10
131 M 2	8	45	-3.64	0.68	1.09	0.30	0.45	-0.20
186 M 3	8	45	-3.64	0.68	0.63	-0.70	0.55	0.00
189 M 1	7	45	-4.13	0.72	1.71	1.40	0.91	0.30
160 F 1	6	45	-4.68	0.76	2.35	2.20	1.92	1.00
130 M 2	6	45	-4.68	0.76	1.37	0.90	0.96	0.40
191 F 1	5	45	-5.27	0.79	2.77	2.90	9.90	9.90
130 M 1	4	45	-5.93	0.84	0.57	-1.00	0.09	-0.80
Mean	19.10	45.00	-0.60	0.49	1.01	0.00	0.95	0.10
SD	6.70	0.00	1.58	0.06	0.30	1.00	1.33	1.10
r	0.89							

HFDC Item Number Conversion Table

DAM Item Number	DAW Item Number	HFDC Item Number	Item Description
1	1	1	Head Present
2	2	2	Neck Present
3	3	3	Neck, 2 Dimensions
4	4	4	Eyes Present
5	5	5	Eye Detail: brow or lashes
6	6	6	Eye Detail: pupil
7	7	7	Eye Detail: proportion
9	9	8	Nose Present
10	10	9	Nose, 2 Dimensions
11	13	10	Mouth present
12	14	11	Lips, 2 Dimensions
13	16	12	Both nose & lips 2 Dimensions
16	18	13	Line of jaw indicated
17	11	14	Bridge of nose
18	19	15	Hair I
19	20	16	Hair II
20	21	17	Hair III
21	22	18	Hair IV
24	28	19	Fingers present
26	30	20	Detail of fingers correct
27	31	21	Opposition of thumb
28	32	22	Hands present
30	24	23	Arms present
31	25	24	Shoulders I
33	26	25	Arms at side or engaged in activity
34	27	26	Elbow joint shown
35	33	27	Legs present
39	35	28	Feet I: any indication
40	36	29	Feet II: proportion
41	38	30	Feet: Heel
43	37	31	Feet V: detail
44	41	32	Attachment of arms and legs I
45	42	33	Attachment of arms and legs II
46	55	34	Trunk present
47	56	35	Trunk in proportion, 2 dimensions
51	60	36	Proportion: arms I
55	43	37	Clothing I
57	51	38	Clothing III
58	53	39	Clothing IV
59	54	40	Clothing V
63	64	41	Motor coordination: lines
64	63	42	Motor coordination: junctures
65	65	43	Superior motor coordination
66	66	44	Directed lines and form head outline
69	71	45	Directed lines and form: facial features

Human Figure Drawing Continuum Booklet and Scoring Guide

Human Figure Drawing Continuum

Name: ____

Date of Birth: ____________________ Today's Date: ____________________

Grade: ________________________ Sex: M / F

Notes:

____

____

____

____

____

Instructions

When completing the task with older children in small groups, facilitate the children in completing the above sections one at a time. When completing the task with younger children or children with special needs, have an adult complete the above sections.

Ask the child/ren to complete a human figure drawing:

> "I would like you to make a drawing of a person for me today. You could draw a picture of yourself, or a friend, or your mummy, daddy, grandma or grandpa (nanny/poppy). It is up to you who you draw, and there is no right or wrong type of drawing. It just needs to be a drawing of a whole person – from top to toe. Take your time and work very carefully. Make the very best drawing that you can."

It is important not to comment, criticize or offer suggestions whilst children are drawing. If a child asks a question or for further instruction, simply say, "Do it whichever way you think is best". There is no time limit. However, children rarely take longer than about five minutes to complete a drawing.

The HFDC is, of course, designed to be used as part of a suite of information when making judgements about young children.

Note. This HFDC is the by-product of research completed by Claire Maley. This is not a standardised standalone test. The original ideas upon which this HFDC is based are copyright to Goodenough, Harris and The Harcourt Brace Psychological Corporation.

1	❑	26	❑
2	❑	27	❑
3	❑	28	❑
4	❑	29	❑
5	❑	30	❑
6	❑	31	❑
7	❑	32	❑
8	❑	33	❑
9	❑	34	❑
10	❑	35	❑
11	❑	36	❑
12	❑	37	❑
13	❑	38	❑
14	❑	39	❑
15	❑	40	❑
16	❑	41	❑
17	❑	42	❑
18	❑	43	❑
19	❑	44	❑
20	❑	45	❑
21	❑		
22	❑		
23	❑		
24	❑		
25	❑		

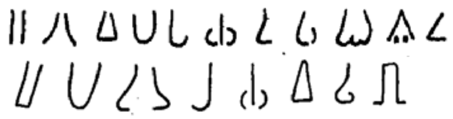
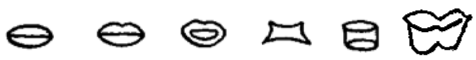
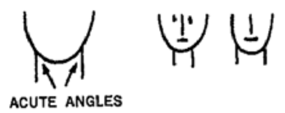
Note. This HFDC is the by-product of research completed by Claire Maley. This is not a standardised standalone test. The original ideas upon which this HFDC is based are copyright to Goodenough, Harris and The Harcourt Brace Psychological Corporation.

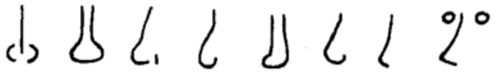
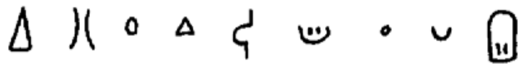
Human Figure Drawing Continuum

Scoring Guide

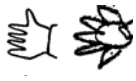
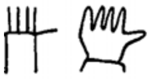
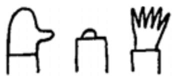
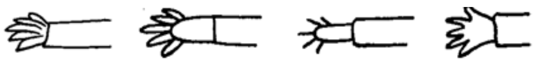
Items	Criteria
1. Head present	Any clear method of representing the head. Features alone, without any outline for the head itself, are not credited.
2. Neck present	Any clear method of representing the neck as distinct from the head and the trunk. Mere juxtaposition of the head and the trunk is not credited.
3. Neck present, two dimensions	Out of neck continuous with that of the head, of the trunk, or both. Line of neck must 'flow' into head line or trunk line. Neck as pillar between head and trunk is not credited. Credit:  No Credit: 
4. Eyes present	Any clear method of representing the eyes (either one or two eyes). A single indefinite feature, such as is occasionally found in the drawings of very young children, is credited.
5. Eye detail: brows or lashes	Brow, lashes, or both shown. Credit: 
6. Eye detail: pupil	Any clear method of representing the pupil or iris as distinct from the outline of the eye. Both pupils must appear if both eyes are shown.
7. Eye detail: proportion	The horizontal measurement of the eye must be greater than the vertical measurement. If two eyes are shown both must meet this requirement. One eye is sufficient if only one is shown.
8. Nose present	Any clear method of representing the nose.

Note. This HFDC is the by-product of research completed by Claire Maley. This is not a standardised standalone test. The original ideas upon which this HFDC is based are copyright to Goodenough, Harris and The Harcourt Brace Psychological Corporation.

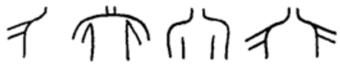
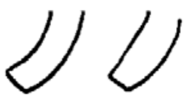
9. Nose present, two dimensions	All attempts to portray the nose in two dimensions is credited. The bridge must be longer than the width of the base or tip. Credit:  No credit: 
10. Mouth present	Any clear method of representing a mouth.
11. Lips, two dimensions	Two lips clearly shown. Credit: 
12. Both nose & lips two dimensions	Bonus credit given when both items 9 and 12 are credited.
13. Line of jaw indicated	The line of the jaw and chin are drawn across the neck, but not squarely. Neck must be sufficiently wide, and jaw/chin must be so shaped that the line of the jaw forms a well defined acute angle with the line of the neck. Score strictly on the simple oval-shaped face. Credit:  No Credit: 

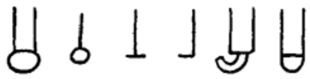
14. Bridge of nose	Nose must be accurately placed and shaped. The base of the nose must appear as well as the indication of a bridge. Placement of upper portion of bridge is important; must extend up to or between the eyes. Bridge must be narrower than the base. Credit:  No Credit: 
15. Hair I	Any clear representation of hair. A single mark above the head is credited. Credit: 
16. Hair II	Credit any marks that conform closely to the head. Credit: 
17. Hair III	Any clear method of representing hair that is distinctly more conceptually advanced than item 16. Credit: 
18. Hair IV	Representation of hair is distinctly more advanced than item 17. Score strictly. Credit: 

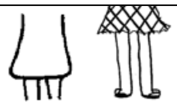
Note. This HFDC is the by-product of research completed by Claire Maley. This is not a standardised standalone test. The original ideas upon which this HFDC is based are copyright to Goodenough, Harris and The Harcourt Brace Psychological Corporation.

19. Fingers present	Any clear method of representing fingers. ‘Mitt’ hand does not score even if thumb is indicated. No Credit: 
20. Fingers present: Proportion	Length of individual fingers must be distinctly greater than width. “Grapes” or “sticks” are not credited. Credit:  No Credit: 
21. Opposition of thumb	A clear differentiation of the thumb from the fingers is represented. The item is credited where there is a distinct difference in the length of the fingers and the thumb, or if the angle between the fingers and thumb is greater than the distance between any two fingers. Score strictly. Credit:  No Credit: 
22. Hands present	Any clear method of representing hands, apart from the fingers. When fingers are shown a space must be left between the base of the fingers and the wrist/sleeve/cuff. Where no wrist/sleeve/cuff exists, the arms must broaden to suggest a palm or back of the hand as distinct from the wrist. Characteristics must appear in both hands is both are shown. Credit:  No Credit: 

Note. This HFDC is the by-product of research completed by Claire Maley. This is not a standardised standalone test. The original ideas upon which this HFDC is based are copyright to Goodenough, Harris and The Harcourt Brace Psychological Corporation.

23. Arms present	Any clear method of representing the arms. Must show two arms. Fingers alone are not sufficient, but the item is credited if any space is left between the base of the fingers and that part of the trunk to which they would be attached.
24. Shoulders	A distinct change in the direction of the upper part of the trunk which gives the effect of shoulders. The elliptical form is not credited. Credit:  No Credit: 
25. Arms at side of engaged in activity	Credit when one, or both, arms are shown as making an angle of no more than 10 degrees with the general axis of the trunk. Also credit when arms are represented as engaged in an activity, arms are on hips or in pockets, or behind back. Credit:  10° OR LESS
26. Elbow joint shown	There must be an abrupt bend (not a curve) at approximately the middle of the arm. Once is sufficient. Modeling or creasing of the sleeve is also credited. Credit:  No Credit: 
27. Legs present	Any clear method of representing the legs. There must be two legs. Also credit where long clothing hides the legs and/or feet.

28. Feet I: any indication	Any clear method of representing two feet. Also credit where long clothing hides the legs and/or feet. Credit: 
29. Feet II: proportion	Horizontal dimension of the fore-part of the foot must be greater than the vertical dimension. Also credit where long clothing hides the legs and/or feet. Credit:  No Credit: 
30. Feet III: Heel	Any clear method of representing the heel. Also credit the item arbitrarily when the foot is shown 'front on' (see first two examples below). Credit: 
31. Feet V: detail	Any clear method of representing detail such as toes, ties, straps, laces, shoe sole, heel, etc. Also credit where long clothing hides the legs and/or feet.
32. Attachment of arms & legs I	Any clear method of representing both the arms and legs attached to the trunk at any point.
33. Attachment of arms & legs II	The arms and legs are attached to the trunk at accurate points. Do not credit if arm attachment occupies more than one-half of the chest area (neck to waist). When no neck is indicated, the arms must definitely be attached to the upper part of the trunk where the shoulders would generally be. Legs attached to the bottom of clothing must be distinct from the clothing. Arms and legs that are drawn as continuous with clothing are not credit. Also credit where long clothing hides the arms, legs and/or feet. Credit:

	 No Credit: 
34. Trunk present	Any clear method of representing the trunk, either one or two dimensional. Where there is no clear differentiation between the head and the trunk, but the features appear in the upper section of a single figure (head/trunk), the item is credited. A figure drawn between the head and legs is credited as a trunk, even if the shape/size suggests a neck.
35. Trunk: proportion, two dimensions	Vertical dimension of trunk must be greater than the horizontal dimension. Measurement must be taken at the points of greatest length and width. If the two dimensions are equal, or nearly so that the difference is not readily determinable, the item is not credited.
36. Limbs: proportion	Length of arms and legs must be greater than width. Where legs are concealed by long clothing, credit the item if the arms are longer than wide.
37. Clothing I	Any clear method of representing clothing.
38. Clothing II	At least two articles of clothing that are nontransparent. That is, the clothing conceals the part of the body which it is intended to cover. Trousers/pants are not credited when the foot is a continuous extension of the leg and the only separation (or indication of trousers/shorts/pants) is a line drawn across the leg.
39. Clothing III	At least two articles of clothing are indicated and the entire drawing is free from transparencies of any sort. Both sleeves and trousers/shorts/pants must be shown as distinct from wrists/hand and legs/feet.
40. Clothing IV	At least four articles of clothing definitely indicated. Clothing must show some detail such as button, fly, pocket, collar, zip, cuff, lapels, spots or stripes, etc. Shoes must show some detail such as laces, heel, sole, strap, or tie. Clothing must be distinct from body and there must be no transparencies of any sort.

Note. This HFDC is the by-product of research completed by Claire Maley. This is not a standardised standalone test. The original ideas upon which this HFDC is based are copyright to Goodenough, Harris and The Harcourt Brace Psychological Corporation.

41. Motor coordination: lines	Main lines of drawing should be well-controlled and free from accidental wavering. The drawing need not contain much detail or many concepts to be credited with this item.
42. Motor coordination: junctures	Points of juncture should meet cleanly without a marked tendency to overlap, cross, or leave gaps between the ends. A drawing with fewer lines is scored more strictly than a drawing with multiple lines. The drawing need not contain much detail or many concepts to be credited with this item.
43. Superior motor coordination	Credit this item where both items 41 and 42 have been credit. Excessive erasures or redrawing invalidate this item.
44. Directed lines & form: head outline	Outline of the head must be drawn without obvious unintentional irregularities. The item is credited only where the shape of the head has developed beyond the first oval or ellipse shape. The face should be developed as a 'unit' rather than by adding 'parts'.
45. Directed lines & form: facial features	Facial features must be symmetrical in all respects and appropriately placed. Eyes, nose and mouth must be shown in two dimensions.

Human Figure Drawing Continuum Variable Map

```
            Persons - MAP - Items
               <more>|<rare>
    6                +T
                     |
                     |
                     |  Directed lines & form: facial features
    5                +  Hair IV
                     |
                     |  Clothing V
                     |
    4                +  Directed lines & form: head outline
                  .  |  Elbow joint shown
                     |
                  .  |  Both nose & lips two dimensions     Clothing IV
    3             #  +S Detail of fingers correct
                  .  |  Lips two dimensions
                .## T|  Opposition of thumb
                 ##  |  Line of jaw indicated   Hair III  Superior motor coordination
    2           .##  +
          .########  |  Clothing III
             ######  |  Bridge of nose     Eye detail: Proportion
             .#####  |  Feet: heel     Eye detail: brow or lashes
    1          #### S+
         .#########  |  Shoulders     Neck two dimensions
            .######  |  Feet V: detail
             .#####  |  Nose two dimensions
    0         #####  +M Arms at side or engaged in activity     Feet II: proportion
       ############  |  Limbs: proportion
              .#### M|  Hands present        Attachment of arms & legs
          #########  |  Motor coordination: junctures   Eye detail: pupil     Hair II
   -1    ##########  +
         .#########  |  Neck present      Motor coordination: lines
       .###########  |  Nose present
       .###########  |  Fingers present
   -2    ##########  +
          .######## S|  Feet I: any indication     Clothing I
             .#####  |
                ###  |  Trunk in proportion two dimensions
   -3                +S
              .####  |
                     |  Hair I
                 .# T|
   -4             .  +
                     |
                     |  Attachment of arms & legs
                 .#  |  Mouth present
   -5                +
                  .  |
                     |  Arms present     Trunk present
                  .  |
   -6                +T
                     |  Legs present
                     |
                     |
   -7                +
                     |  Eyes present
                     |
                     |
   -8                +  Head present
               <less>|<frequ>
 EACH '#' IS 3.
```

Note. This HFDC is the by-product of research completed by Claire Maley. This is not a standardised standalone test. The original ideas upon which this HFDC is based are copyright to Goodenough, Harris and The Harcourt Brace Psychological Corporation.

Draft Number to Measure Conversion Table

HFDC Raw Score	Rasch Measure	S.E.
0	-9.22E	1.92
1	-7.76	1.17
2	-6.70	0.93
3	-5.93	0.84
4	-5.27	0.79
5	-4.67	0.76
6	-4.13	0.72
7	-3.64	0.68
8	-3.21	0.63
9	-2.84	0.59
10	-2.50	0.56
11	-2.20	0.54
12	-1.93	0.52
13	-1.67	0.50
14	-1.43	0.49
15	-1.19	0.48
16	-0.97	0.47
17	-0.75	0.46
18	-0.54	0.46
19	-0.33	0.46
20	-0.12	0.46
21	0.08	0.45
22	0.29	0.46
23	0.50	0.46
24	0.71	0.46
25	0.92	0.46
26	1.14	0.47
27	1.36	0.47
28	1.58	0.47
29	1.81	0.48
30	2.04	0.49
31	2.28	0.49
32	2.53	0.50
33	2.78	0.51
34	3.05	0.53
35	3.34	0.54
36	3.64	0.56
37	3.96	0.58
38	4.31	0.61
39	4.70	0.65
40	5.15	0.71
41	5.72	0.81
42	6.58	1.08
43	7.89E	1.87

Note. This HFDC is the by-product of research completed by Claire Maley. This is not a standardised standalone test. The original ideas upon which this HFDC is based are copyright to Goodenough, Harris and The Harcourt Brace Psychological Corporation.