Abstract

TAYROSE, MICHELLE PARKER. The Developmental Organization of Cognitive Abilities of Deaf and Hard of Hearing (D/HOH) People as Compared to those who are Normal Hearing. (Under the direction of Jeffery P. Braden.)

The major purpose of the present study was to better understand whether the cognitive abilities of D/HOH and normal-hearing people are organized similarly throughout the lifespan, from preschool through adulthood. Previous research separating children into homogeneous age groups suggests differences in the cognitive ability structures of D/HOH and normal-hearing children at younger but not at older ages. However, no studies have examined this relationship using analytical techniques that are consistent with present-day intelligence theory and statistical practice. Thus, the current study used confirmatory factor analysis with CHC-driven hierarchical models to determine whether cognitive abilities were similar for both groups, (i.e., factor models adequately fit the data from normal-hearing and D/HOH people) and subsequent multisample confirmatory factor analysis to assess for invariance across groups (i.e., to determine whether intelligence tests are measuring the same construct in both groups). A bibliographic search yielded eight samples. Results suggest that factor structure differences do not exist between older (i.e., age 11 through adulthood) D/HOH children/adults and their normal-hearing peers, but were inconclusive in showing differences at younger ages. Implications regarding the use of cognitive test batteries and the application of CHC theory with these populations, as well as suggestions for future research are discussed.

The Developmental Organization of Cognitive Abilities of Deaf and Hard of Hearing (D/HOH) People as Compared to those who are Normal Hearing

by
Michelle Parker Tayrose

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Psychology

Raleigh, North Carolina

2011

APPROVED BY:

Jeffery P. Braden, Ph.D. Chair of the Advisory Committee	William P. Erchul, Ph.D.
Adam W. Meade, Ph.D.	John L. Nietfeld, Ph.D.

Biography

Born and raised in Chapel Hill, North Carolina, Michelle Parker Tayrose is the daughter of loving parents, Edythe and Barnett Parker, and older sister to wonderful siblings, Allison and Jeremy. Michelle graduated from East Chapel Hill High School in 1998 and then attended the University of North Carolina at Chapel Hill where she majored in Psychology and minored in Advertising. Upon receiving her Bachelor of Arts degree in May 2002, Michelle worked for two years at the Duke University ADHD Program on the Multimodal Treatment Study of Children With and Without Attention-Deficit/Hyperactivity Disorder (MTA) before entering the North Carolina State University School Psychology program. She received her Master of Science degree in May 2008 and completed an internship with the Scarsdale County School District during the 2009-2010 school year. She will receive her Doctor of Philosophy degree in May 2011. She currently resides in Manhattan with her amazing husband, Greg, and their baby girl, Hallie, and dog, Bumble.

Acknowledgements

First and foremost, I would like to thank my advisor, Jeff Braden, who has pushed me to think critically and independently over the last six years and has always helped me to "find the horse." He has invested an extraordinary amount of time and effort, and has provided much guidance in both my professional and personal development. The mentorship he has offered has been invaluable, inspiring and motivating me throughout my graduate career. I would also like to extend my gratitude to my committee members, Bill Erchul, Adam Meade, and John Nietfeld, for their participation in my preliminary examinations, as well as for their time, insightful feedback, and collaboration on this project. In addition to his support on this project, I would like to thank Bill Erchul for his encouragement since I entered this program. I would also like to thank Ann Schulte for all of her academic and emotional support.

This dissertation would not have been possible without the help of Tom Fagan, Erica Miller, Irene Mulkerin, and Steven Sligar (and his graduate students, Kyle Slough and Danielle Nilges), all of whom provided me with the test kits and reference materials necessary to conduct my analyses. Additionally, I would like to thank Mimi Riggs and Jim Clark, who worked tirelessly to help me access dissertations as part of my bibliographic search. I know they are extremely busy and thus deeply appreciate the time and effort they dedicated to this project. Special thanks to Victoria Briones Chiongbian for all of her assistance with my statistical procedures.

I would also like to thank Megan Bennett and Lauren Joyce for their coding assistance, and more importantly, for the support they have provided throughout my graduate career. They, as well as those in my cohort and the cohort below me, have served as a source of constant encouragement and motivation, and I am so happy to have met them through this experience. They made it bearable! I would also like to thank my longtime friend, Anna Sheahan. I'm not sure I would have made it through my dissertation if she had not been writing one also. Whether we were working, venting, or distracting one another, going through this process together made it all the more enjoyable.

It is virtually impossible to show due appreciation to my parents, who have always been there for me, encouraging and believing in me. They have shown me constant love and support, and just knowing I will make them proud is deeply motivating. I also wish to thank my brother and sister, whose happiness in my achievements means so much, and whose love and support helped me through this endeavor.

Thanks go, as well, to my extended family. They are all such admirable people who have provided me with much encouragement and confidence throughout my life. I take this opportunity to offer deepest thanks to my grandparents who had such a lasting, positive impact on my life. I know they are here in spirit and would be so proud of me, and for these reasons I will be eternally grateful.

And last but certainly not least, I would like to thank my husband, Greg, whose unwavering love and support made this work possible. Whether it was a much-needed hug, words of encouragement, a reminder that I am "the woman for the job," or just some fun and relaxation, he has been there for me in more ways than I can express. I feel so lucky to have him in my life, and I can't imagine a better teammate. Together we have started our own family, and for that I feel so blessed. I could not be more excited about the arrival of our little "tootsie pop," Hallie.

Table of Contents

List of Tables

List of Figures

Chapter 1

Review of the Literature and Statement of the Problem

The focus of the present study is the structure of cognitive abilities of people who are deaf and hard of hearing (D/HOH). More specifically, the proposed study seeks to compare the organization of human cognitive abilities for D/HOH and hearing people across the lifespan. Findings yielding similar structures would serve to validate the assumption that cognitive abilities, as conceptualized and measured by modern-day test batteries, are similarly organized across groups. However, it is necessary to review some concepts relevant to the D/HOH population, including characteristics of hearing loss and what is meant by the term "deafness." Following this review, information on the construct of intelligence/organization of cognitive abilities, methods, and validity with D/HOH samples is offered. Finally, a summary of the extant factor-analytic literature surrounding the intellectual structure within the D/HOH population is provided.

Defining Deafness

There exist three main structures of the ear, namely: (a) the outer ear, which contains the pinna/auricle (or the external portion of the ear) and the ear canal; (b) the middle ear, comprised of the eardrum (or a thin membrane separating the outer and middle ears) and several small bones connecting the eardrum to the inner ear; and (c) the inner ear, consisting of the cochlea and vestibular labyrinth, chambers housing the organs of hearing and balance, respectively. Essentially, sound reaches the outer ear and causes vibrations to be funneled into the ear canal and then into the eardrum. These vibrations create small movements in the middle ear bones, move into the inner ear, and finally disseminate into the cochlea. Here, the sound is converted into nerve impulses that are transferred to the auditory centers of the brain via the eighth cranial (vestibulochlear/auditory) nerve. Damage or blockage to some aspect of this sequence from the outer ear to the brain results in hearing loss, which varies in severity depending on where it occurs (e.g., middle ear problems tend to be less severe, cochlea or auditory nerve problems tend to be more severe; Marschark, Lang, & Albertini, 2002; Moore & Dalley, 1999).

The term “deafness” does not apply solely to those who experience a complete lack of hearing. Rather, it is heterogeneous in nature and can be used to describe a wide range of conditions (e.g., normal auditory sensitivity but an inability to differentiate foreground from background noise). Hearing loss (i.e., deafness) is classified according to the following three dimensions, each of which is described more thoroughly below: (a) etiology, (b) severity, and (c) onset.

Etiology. Interference in the transmission of sound to auditory brain centers is what causes hearing impairment; however, the specific etiology or organic description of the hearing loss depends on where these problems arise and dictates the type of hearing loss experienced. Difficulties in the middle ear (e.g., rupture of the eardrum, calcification of the small bones) affect the conduction of sound waves to the oval window (in the inner ear), thereby causing decreased sensitivity to sound, called a “conductive” hearing loss. Typically conduction problems can be alleviated through medical intervention, amplification of sound impulses, or a combination of the two.

Problems with the inner ear or auditory nerve are termed “sensorineural” impairments. Here, problems occur in the transformation of motion to nerve impulses making it difficult to detect and identify sounds, as well as understand speech. Sensorineural impairments tend to be resistant to medical intervention and are instead treated with hearing aids (which work to amplify sound) or cochlear implants (electronic devices that send electrical signals directly to cochlea nerve fibers). In a “mixed” hearing loss, elements of both conductive and sensorineural impairments are present (i.e., there are problems in the outer or middle ear and the inner ear).

Finally, a “central” hearing loss is a dysfunction with the auditory brain centers, resulting in difficulties localizing sounds, following directions, and attending. Interventions for central hearing losses target skills training in the areas of auditory/phonological awareness, language processing, and functional organization. The scope of impairment, that is whether the hearing loss is unilateral (i.e., present in one ear) or bilateral (i.e., present in both ears), is also part of the etiology/organic description (Marschark, Lang, & Albertini, 2002; Moore & Dalley, 1999; Northern & Downs, 2002).

Severity. The second aspect of hearing loss is severity, which is evaluated based on one’s ability to hear sounds at varying frequencies (i.e., pitches, as measured in hertz [Hz] or cycles per second) and intensities (i.e., levels of loudness, as measured in decibels [dB]). Hearing loss ranges from slight (a loss of 16-25 dB) to profound (a loss of 91 dB or greater), depending on the intensity needed to perceive sound at a specific frequency. Because most speech sounds occur between 500 and 2000 Hz, hearing impairments are classified on the basis of what intensity is needed to perceive sounds across these frequencies. It is important to note that there is not one agreed-upon classification scheme, but Table 1 outlines those

Table 1

Degrees of Hearing Loss

Classification of Hearing Loss	Criteria	Associated Problems
Normal	Loss of 0 – 15 dB	No hearing impairment
Slight	Loss of 16 – 25 dB	Hear vowel sounds clearly but often miss unvoiced consonant sounds
Mild	Loss of 25 – 30 dB	Often miss soft or whispered speech, tend to have mild speech problems
Moderate	Loss of 31 – 50 dB	Have difficulty hearing speech sounds at normal conversational levels when there is background noise, tend to have moderate speech problems

Table 1 (continued)

Severe	Loss of 51 – 70 dB	Hear only the loudest speech sounds, cannot detect speech sounds at normal conversational levels, oral speech either does not exist or differs in quality as compared to normal-hearing people
Profound	Loss of 71 dB or greater	Tend to have complete lack of hearing, often have inarticulate speech

characteristics commonly used by audiologists for evaluating the levels of hearing loss (Northern & Downs, 2002).

Onset. Onset represents the third characteristic of hearing loss and describes when, in an individual's lifetime, hearing loss begins. Whereas medical definitions tend to differentiate between hearing impairments that are congenital (i.e., present at birth) and those that are adventitious (i.e., acquired after birth), psychological definitions differentiate between pre and postlingual hearing loss. That is, when the hearing loss emerges in relation to when oral language is acquired. However, there exists some debate as to what constitutes "prelingual" as opposed to "postlingual." Some argue that the separation occurs around two years of age (which is when children begin speaking one-word utterances), whereas others suggest five years of age (which is when children acquire such language skills as basic grammar and syntax; Braden, 1994). Children with prelingual hearing impairments experience much greater difficulty acquiring speech as compared to those who do not suffer hearing loss until after language has developed (Sattler & Hoge, 2006).

Distinguishing People who are Deaf from those who are Hard of Hearing

Another important distinction is that between deafness and hard of hearing. Generally, individuals are considered to be hard of hearing if they experience mild to moderate hearing loss (Marschark, Lang, & Albertini, 2002). However, Braden (1994) provides a more specific definition in which hard of hearing describes those whose hearing impairments are mild to moderate, bilateral, and are of prelingual onset. He considers the most important factor differentiating deafness from hard of hearing to be the “degree to which auditory input is available and useful for the acquisition of speech” (p. 22). Consistent with this notion, Marschark, Lang, and Albertini (2002) say many authors define hard of hearing people as those who, regardless of the extent of their hearing loss, use spoken language.

Deaf versus deaf

It is also important to distinguish between the terms “deaf” and “Deaf.” When left uncapitalized, deafness refers to the physical loss of hearing and generally relates to those who experience hearing impairment as a result of older age. Regardless of the extent of their hearing loss, these late-deafened individuals continue to identify with hearing culture. When capitalized, Deaf describes those whose hearing loss occurs at a young age and who therefore grow up as members of the Deaf community. This community is separate from the hearing world and retains its own primary language (American Sign Language [ASL]), social structures, art, organizations, values, and cultural history. "Deaf people, the Deaf community, and Deaf culture all indicate a group of individuals who constitute linguistic and cultural minorities in many countries and who often see themselves as an international, cross-cultural community" (Marschark, Lang, & Albertini, 2002, p. ix; Marschark, 1997).

Language Exposure

Because severe to profound hearing loss is associated with an inability to acquire language, and most D/HOH children are raised in households in which both parents have normal-hearing (83.4%) and signing is not regularly used (71.0%), D/HOH children often experience extremely limited language exposure (Braden, 1994; Gallaudet Research Institute, 2006). D/HOH children's exposure to language is said to vary as a function of: (a) degree of hearing loss (including duration and onset), (b) the modalities in which language is presented (e.g., oral/auditory, gestural/visual, combined), (c) the particular language presented (i.e., sign language, gestural representations of spoken languages, or a mixture of the two, referred to as gestural pidgins), and (d) neurolingusitic aspects of nonstandard language exposure. Although it does not appear that reduced language exposure results in less complex brain development, "deaf and hard of hearing children are exposed to nonstandard language models less frequently, less intensively, for shorter durations, and later in life than their normal-hearing peers" (Braden, 1994, p. 41). As a result, it can be said that D/HOH children experience language deprivation, which is fundamental to the discussion of how cognitive abilities are organized within this population.

The Construct of Intelligence and the Organization of Cognitive Abilities

Referred to as the "father of cognitive and intellectual assessment," Binet, along with his collaborator, Simon, developed the first practical intelligence scale in the early 1900s (Wasserman & Tulsky, 2005, p. 6). This Binet-Simon Scale was designed to identify those children who would benefit from a strongly academic education and, despite the fact that its formation marked the beginning of modern intelligence testing, Binet never formally defined the term "intelligence." The lack of clarity that existed at that time with regard to the nature of the construct has continued to the present day; however, there does seem to be consensus amongst theorists that adaptability to one's environment, judgment, rational thinking, creativity, problem solving, and the ability to use information for learning are central to its conceptualization (Akamatsu, Mayer, & Hardy-Braz, 2008; Wasserman & Tulsky, 2005).

CHC theory. It was not until the latter part of the twentieth century that scholars began to converge on a cohesive theory describing the nature and structure of cognitive abilities. The Cattell-Horn-Carroll (CHC) model of intellectual abilities was developed following Carroll's (1993) systematic survey of hundreds of existing factor-analytic studies of human cognitive abilities. Considered to be an expansion and extension (Carroll, 2005, p. 74) of previous theories (e.g., those of Spearman, Thurstone, Horn and Cattell), advocates of CHC theory postulate that cognitive abilities are hierarchically organized into three strata or levels. At the highest stratum, III, there exists only one superordinate characteristic, labeled *g*, or general intelligence. Below stratum III is stratum II, which is comprised of various "broad" abilities (e.g., fluid intelligence, crystallized intelligence, general memory and learning, processing speed). The first stratum, I, represents the narrowest abilities within the model (e.g., lexical knowledge, numeric fluency, simple reaction time).

The three-stratum model posited by CHC theory highlights the multifactorial nature of intelligence, emphasizing that intelligence should not be thought of as a single entity, nor as a collection of independent abilities, but rather as a hierarchical organization of related abilities (Carroll, 2005; Maller & Immekus, 2005; Wasserman & Tulsky, 2005). Ultimately, intelligence is operationally measured by intelligence tests (Maller & Immekus, 2005), and cannot be directly measured, as is true of any latent variable (i.e., it is characterized by its measurement). Thus, there exists a complementary relationship between test and theory, a notion which acts as the basis for Ong and van Dulmen's (2006) "proof of concept" model. Here, Ong and van Dulmen argue that validity encompasses both the validation of the construct as well as of the measuring instrument. They argue that, to validate a construct, "evidence is sought in support of the hypothesized relations among aspects of the same construct and among different constructs" (p. 19). Ong and van Dulmen posit the more classical definition of construct validity (i.e., the degree to which a test measures the traits which it purports to measure) serves as evidence of the validity of the measuring instrument. Thus, it is necessary to garner validity evidence at both the theoretical and measurement levels. Importantly, however, these types of validity evidence are not mutually independent (e.g., a test found to have a factor structure consistent with underlying theory validates both the test and the theory).

As previously noted, Carroll (1993) factor analyzed hundreds of datasets from the extant literature resulting in a theory that provides a "map" of all human cognitive abilities.

Carroll's work thus serves as factor-analytic (i.e., structural) evidence in support of CHC theory. However, CHC theory is also supported by other non-factor-analytic evidence, including: (a) outcome-criterion research, which shows differential relations between CHC abilities and educational/occupational outcomes; (b) neurocognitive evidence, which show links between CHC abilities and aspects of physiological and neurological functioning; (c) heritability studies, which show differential heritability rates for CHC abilities; and (d) developmental research, which shows CHC abilities to have different patterns of growth and decline across the lifespan (Horn & Noll, 1997). Importantly, studies have also shown the factor structure underlying CHC theory to be invariant across the lifespan (Bickley, Keith, & Wolfe, 1995; Carroll, 1993) and across gender, cultural, and racial groups (Carroll, 1993; Gustafsson & Balke, 1993; Keith, 1999). However, Carroll does note that different environmental experiences may influence differences in cognitive factor structure.

Because CHC theory boasts an extensive network of empirical support, its framework is one that has influenced the design and the interpretation of many contemporary intelligence tests. That is, "today…nearly every intelligence test developer acknowledges the importance of CHC theory in defining and interpreting cognitive ability constructs, and most have used this theory to guide directly the development of their intelligence tests" (Alfonso, Flanagan, & Radwan, 2005, p. 188). In particular, CHC theory has served as the blueprint for the development of the Stanford-Binet Intelligence Scales, Fifth Edition (SB5; Roid, 2003), the Kaufman Assessment Battery for Children, Second Edition (KABC-II; Kaufman & Kaufman, 2004), the Woodcock Johnson III Tests of Cognitive Abilities (WJ III COG;

Woodcock, McGrew, & Mather, 2001), and the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV; Wechsler, 2003).

Verbal and nonverbal intelligence tests. There exist two main types of intelligence tests, those that are verbal in nature and those that are nonverbal in nature. Verbal intelligence test batteries require examinees to use language in most or all aspects of the testing process, including task comprehension, task solution, and task response. As such, these types of tests are language-loaded, in that they require an examinee to meet a threshold of exposure to the standard form of the dominant language. As previously noted, D/HOH children experience severe language deprivation and atypical language exposure, and as a result, results from verbal tests are often invalid as indicators of cognitive abilities.

In contrast, nonverbal tests reduce or eliminate language use in task comprehension, task solution, and task response. Nonverbal (or language-reduced) tests can be further classified as either performance-based or motor-free, depending on demands for manual dexterity and speed in task responses. Although many tests claim to be nonverbal in nature, nonverbal tests are those that do not require examinees to use language in understanding, doing, or responding to test items (Braden & Athanasiou, 2005). For example, performance-based tasks that require an examiner to read directions aloud would not be nonverbal (McCallum, 2003). Thus, in contrast to verbal test batteries, language-reduced or nonverbal tests are considered more appropriate for use with examinees who have not had customary exposure to the test language, such as those who are D/HOH, speak a different language, or who have suffered a traumatic brain injury (Braden & Athanasiou, 2005; McCallum, 2003). The question remains as to whether nonverbal tests measure the same processes as do verbal tests, but measure them nonverbally (i.e., they measure intelligence nonverbally), or whether they measure distinct nonverbal cognitive processes (i.e., they measure nonverbal intelligence). Some argue the former (e.g., McCallum, 2003), but others note scholars continue to debate this question (e.g., Braden & Athanasiou, 2005). Findings from factor-analytic studies may illuminate this issue. That is, evidence of parallel factor structures between verbal and nonverbal tests would suggest the tests are measuring similar cognitive processes, whereas evidence of different factor structures would indicate the tests are measuring distinct processes.

Validity and Test Bias

In specifying what is meant by the term "validity," Messick (1995) argued that it should be viewed as a unified concept in which all aspects of content, criteria, and consequences are integrated into a comprehensive framework to understand the meaning of a test result. More specifically, Messick argued that the characterization of three different types of validity (i.e., content, criterion, and construct) was incorrect; rather, criterion-related evidence (both predictive and concurrent) and content evidence should be subsumed under the superordinate umbrella of construct validity, or the degree to which a test measures what it intends to measure (Carmines & Zeller, 1979; Maller & Immekus, 2005). Messick also identified two new domains of validity evidence (i.e., response processes and test consequences), and noted that this evidence, as well as evidence internal to the test (e.g., scale consistency, factor structure), were needed to provide for a complete understanding of a test result (i.e., to establish the test's construct validity).

Test bias is said to occur when specific groups respond to items differently than the normative group due to such factors as their language or culture (Maller, 2003a). Thus, bias is defined as differential reliability or validity across subgroups and, if found to occur, jeopardizes the construct validity of the instrument. Messick (1995) explicated two major sources of bias, and therefore potential threats to construct validity: (a) construct-irrelevant variance and (b) construct underrepresentation. In the former, the assessment presumes no variance in one or more access skills (skills needed to perform well on a test); however, when test-takers do not have one or more of those skills, factors extraneous to the intended construct influence their score. In the latter, changes to the assessment intended to reduce/remove the influence of one or more access skills may inadvertently reduce the degree to which the test measures the intended construct, resulting in inadequate sampling of the construct of interest. These validity threats are relevant to intelligence testing with D/HOH populations. More specifically, it is important to be aware of these threats to ensure sources of assessment bias are controlled so that test results are not misleading.

Construct-irrelevant variance and construct underrepresentation. As mentioned earlier, D/HOH people are unlikely to have been exposed to language in the same ways as their normal-hearing peers. That is, D/HOH people are often denied early and consistent access to language because they tend to live in households in which the native language is communicated orally and sign language is not regularly used. Thus, D/HOH children often experience delayed language exposure. Even for the 24.3% of children whose family members do use sign language regularly (and this is likely an overestimate given that most adult language models are not fluent), because sign language must be seen, D/HOH children do not acquire as much incidental language as normal-hearing children because they are unable to overhear conversations (Akamatsu, Mayer, & Hardy-Braz, 2008; Braden, 2005; Gallaudet Research Institute, 2006). Furthermore, the quality of the language to which they have access is different both in the medium in which it is transmitted and received (visual), and in its grammatical content (which is typically an ASL/English Pidgin; Braden, 1994).

As a result, many researchers argue that verbal (language-loaded) tests of intelligence are inappropriate for use in estimating D/HOH people's intelligence because they presume D/HOH children have received adequate opportunities to acquire spoken (English) language (Braden & Athanasiou, 2005; Maller, 2003a). Use of verbal subtests with D/HOH examinees is likely to introduce construct-irrelevant variance because they measure factors extraneous to underlying cognitive ability, such as hearing loss, language exposure, and knowledge of English. Those opposing the use of verbal tests for assessing the intelligence of D/HOH people also cite evidence showing D/HOH people score roughly one standard deviation lower on language-loaded tests than they do on language-reduced tests (Braden, 1992, 1994), which they view as evidence of bias.

Although D/HOH people score approximately one standard deviation below the means of standardization samples on verbal scales of intelligence, their nonverbal scale scores tend to fall within the average range and with a distribution nearly identical to that of normal-hearing people (Braden, 1994). These findings, along with findings that nonverbal subtest/composite scores are not correlated with degree of hearing loss, have led researchers to conclude that performance-based intelligence scores can be used to differentially diagnose the effects of deafness from those of cognitive deficits (Braden, 1992, 1994; Vernon, 1967/2005).

Although measuring D/HOH people's intelligence with nonverbal instruments remains best practice, some researchers still advocate for the inclusion of verbal scales. They cite many reasons; however, perhaps the most important is to avoid construct underrepresentation. The structure of human cognitive abilities is currently best described within a CHC three-tier model. Three stratum II factors in particular (i.e., crystallized ability, literary or reading/writing ability, and auditory processing) are strongly related to language, and are therefore often excluded from nonverbal tests. However, eliminating these abilities from the assessment leads to a reduction in the representation of abilities, and therefore a narrower estimate of general intelligence or *g* (Braden & Athanasiou, 2005). This issue of construct underrepresentation is often ignored, which Braden (2001) postulates is due to: (a) a lack of consensus with regard to which are the essential cognitive abilities contributing to general intelligence, (b) the belief that all tasks measure general intelligence and therefore *g* can be sampled accurately whether using verbal or nonverbal scales, and (c) the general gap between research and practice with D/HOH people.

Factor Analysis

The primary form of evidence scholars use to develop and support CHC theory is structural (i.e., factor analytic). Ong and van Dulmen (2006) specify that in validating a multidimensional phenomenon, such as intelligence, it is necessary to provide evidence that the "dimensions of the phenomenon exhibit a factor structure that is consistent with the underlying theory" (p. 19). It is also imperative that this factor structure be invariant across groups, as this is a necessary structural condition for construct validity. That is, the same construct must be measured and present with different samples. Factor invariance studies employ factor-analytic techniques to evaluate the consistency of internal test structure with specified models, and across samples, as a means of testing theory consistency. Findings that are consistent with the expected factor structure validate both the test and the guiding theory.

Factor analysis is the most widely used method for garnering evidence of validity of intelligence test theory with D/HOH people. Essentially, factor-analytic methods determine whether D/HOH and normal-hearing people obtain similar factor structures because differences in structure suggest the construct (i.e., intelligence) is not the same for both groups. Thus, differences in factor structure imply that cognitive structures differ between groups, and therefore test scores would have different meanings for each group (Braden, 1994; Maller, 2003a). Historically, deafness research has used factor analysis for two distinct but related purposes: (a) to validate testing instruments (i.e., to show that test results have the same meaning for D/HOH examinees as for examinees with normal hearing), and (b) to investigate whether D/HOH people organize their cognitive abilities in a manner similar to their normal-hearing peers. That is, researchers ask two related questions—"Does the test work with D/HOH people?", and "Is intelligence the same thing in D/HOH people?"

Factor analyses are delineated into two types: (a) exploratory and (b) confirmatory. Researchers use exploratory factor analysis (EFA) when they do not have an *a priori* theory with regard to the underlying structure of an instrument, whereas they use confirmatory factor analysis (CFA) to test a hypothesized model. As such, EFA helps develop theory, whereas CFA tests theory. Because researchers rarely undertake projects without some guiding theory, EFA is a less appropriate approach than is CFA, and should be used infrequently (Keith, 2005; Maller, 2003a). However, because CFA is a relatively new research method, it is only recently that researchers have drawn a sharp distinction between theoretical development and confirmation (and thus between EFA and CFA). As such, most of the investigations surrounding the organization of cognitive abilities within the D/HOH population have used the EFA method. Concurrent research examining CHC theory employs both EFA and CFA techniques. That is, CHC theory was originally developed through the use of EFA (e.g., Carroll's systematic survey of factor analyses) which provided a framework from which to make predictions that have subsequently been tested using CFA (e.g., Keith, 2005; Keith, Fine, Taub, Reynolds, & Kranzler, 2006; Phelps, McGrew, Knopik, & Ford, 2005).

With CFA, researchers are able to specify the nature of the model (factor structure) underlying the data (e.g., number of factors, variables that load on each factor). The fit of the model to the data (i.e., that matches the covariances among the variables) is then evaluated using a variety of goodness-of-fit statistics, such as the comparative fit index (CFI), the non-normed fit index (NNFI), the root mean squared error of approximation (RMSEA), and the standardized root mean square residual (SRMR; Hu & Bentler, 1999; Vandenberg & Lance, 2000). The CFI and the NNFI test the difference in fit of the null and target models as compared to the fit of the null model; however, the CFI is designed to estimate the fit in the population, whereas the NNFI appears to be relatively unaffected by sample size. The RMSEA calculates the average difference between the empirical and modeled covariance matrices, and the SRMR represents the standardized difference between the observed and predicted covariance matrices. Here, a model found to fit a given dataset can be applied to a second dataset. If an adequate fit is not found between the model and the second sample, but is present between the model and the first (usually normative) sample, it implies that test constructs are different for each of the groups.

However, even if a common model fits both groups (which would serve as validity evidence at the theoretical level), progressively more restrictive models may be tested for measurement invariance (to gain validity evidence at the measurement level) using a type of CFA called multisample CFA (MCFA). Measurement invariance refers to a consistency in measurement (i.e., the yielding of similar attributes) under different conditions (i.e., across two or more groups) and offers one way to investigate test bias (i.e., differential validity of test scores). Ultimately, the goal of measurement invariance is to ensure that responses, and therefore scores, are not related to group membership. Invariance of factor structure suggests common constructs and behavior across groups. In contrast, variability of factor structure suggests the presence of differential validity and thus jeopardizes the construct validity of the instrument. Variability implies either the groups do not share the same underlying construct(s), or the construct(s) are not being measured the same for each group. Therefore, test scores will have different meanings for members of each group. As a result, a lack of invariance precludes any cross-group comparisons because the "basis for drawing scientific inference is severely lacking" (Horn & McArdle, 1992, p. 117).

Invariance testing via MCFA involves a series of tests, each of which requires the application of more constraints, thereby yielding increasingly more restrictive models. These models are nested within one another and are compared via the likelihood ratio test (i.e.,

$\Delta\chi^2$). A significant change in the likelihood ratio test indicates a significant decline in fit between models and suggests differences between groups in the constrained matrix (i.e., covariances among the variables; Kline, 2005). However, a nonsignificant change suggests equivalence across groups and thus the presence of measurement invariance (French & Finch, 2006). Although researchers rely on $\Delta\chi^2$ to compare competing nested models, goodness-of-fit statistics may be used to describe the fit of single models (e.g., Keith et al., 1995; Reynolds & Keith, 2007). The most common of these fit statistics is chi-square (χ^2), but it is directly related to sample size such that large samples are more likely to lead to rejection of a model (Keith, 2005). Thus, it is recommended that other fit indices, such as the CFI, NNFI, and RMSEA be examined in conjunction with χ^2 (Hu & Bentler, 1999; Vandenberg & Lance, 2000).

Factor-analytic studies in the extant literature that examine the intellectual structure within the D/HOH population have, collectively, yielded mixed findings and thus inconclusive results. Early studies (Hine, 1970; Farrant, 1964) suggested that D/HOH and normal-hearing people have different factor structures. Specifically, Farrant's results indicated the factor structures were similar in many ways, but also showed the "intellective abilities of the deaf and hard of hearing children were less correlated with each other than was the case with the normal hearing children" (pp. 322-323). Hine obtained similar results, finding three factors to account for the majority of the variance in intelligence and achievement in D/HOH children. Hine's three-factor model contrasts with the two-factor model (consisting of verbal and performance abilities) on which the WISC is based, and led him to conclude that "the three abilities identified here tend to develop relatively more independently in partially hearing children than in unhandicapped [sic] children" (p. 177).

However, Braden (1985) argued that these conclusions of factorial independence are likely the result of methodological variability (i.e., Varimax rotation creates orthogonal factors, whereas oblique rotation allows for factors to be correlated). Also, both Farrant and Hine used samples comprised of D/HOH children with a wide range of hearing loss and also included language-loaded achievement tests in their assessment batteries. Because hearing loss is known to correlate highly with "language-related achievements," hearing loss serves as a source of extraneous variance in the deaf samples (Braden, 1985, p. 496).

More contemporary studies that incorporate both verbal and nonverbal subtests in the analyses also suggest differences in factor structure between D/HOH and normal-hearing children (Sullivan & Montoya, 1997; Sullivan & Schulte, 1992). Sullivan and Montoya and Sullivan and Schulte employed principal axis factoring and rotated the solution to a Varimax criterion, resulting in two interpretable factors, labeled Language Comprehension (*l*) and Visual-Spatial Organization (*v-s*). These two-factor solutions depart from the three- and four-factor solutions found for hearing children on the WISC-R and WISC-III, respectively. That is, Freedom from Distractibility and Processing Speed factors were not obtained for D/HOH children. Sullivan and Montoya note that this two-factor model (*l* and *v-s*) is "consistent with Wechsler's…original belief that two factors, verbal and nonverbal, underlie intelligence" (p. 319) but does not suggest identical factor structures for D/HOH and hearing children.

Factor analyses that use only nonverbal tests yield different conclusions depending on whether they use relatively homogeneous age groups (e.g., children grouped into age strata of 2 – 6 years) or heterogeneous age groups (e.g., children aged 6 – 17 are combined). That is, when the cognitive ability structures of D/HOH children are compared with those of their normal-hearing same age peers in homogeneous age groups, there exist differences in the structures at younger ages (roughly 10 and under) but not older ages (Bolton, 1978; Braden, 1985; Zweibel & Mertens, 1985). However, research using heterogeneous age groups suggests D/HOH and normal-hearing children have similar factor structures on performance intelligence tests (Braden, 1984). Each of these studies is described in more detail below.

Bolton (1978) sought to examine the differential ability structure of D/HOH and hearing children using the Hiskey-Nebraska Test of Learning Ability (HNTLA), a test designed specifically for use with D/HOH children. He extracted factors using maximum likelihood estimation and retained two for both D/HOH and hearing children aged 3 through 10 years. Following oblique rotation of the factors, results indicated that two of subtests (Memory for Color and Block Patterns) measured different cognitive abilities in each of the two groups. Bolton took this finding as evidence that D/HOH and hearing children have different "thinking processes," thereby supporting Mykleburst's organismic shift hypothesis (p. 148).

Braden (1984, 1985) used principal factor analysis, thereby allowing *g* variance to be "captured and identified," and statistically compared the factorial similarity between D/HOH and hearing children via the Burt-Tucker Coefficient of Congruence (Braden, 1985, p. 497). Specifically, he extracted and compared factors using the Performance scales/nonverbal subtests of the WISC, WISC-R, and HNTLA. When the intercorrelation matrices of the WISC-R Performance subtests were examined for the WISC-R standardization sample and a large heterogeneous sample of D/HOH children, only one factor was extracted. Based on previous research that found Performance subtests to load highly on a general factor (*g*) and a performance factor (*P*), Braden (1984) subsequently named this factor *g* + *P*. A high degree of factorial similarity was found between groups, suggesting the structure of intelligence is the same for both D/HOH and hearing children. One *g* + *P* factor was also found when intercorrelation matrices of Performance subtests for D/HOH and hearing children from the normative samples of the WISC and WISC-R were compared (Braden, 1985).

Unlike those of the Wechsler Scales, which included ages 6-17 in a single sample, the standardization sample for the HNTLA is divided into younger (ages 3 to 10) and older (ages 11 to 17) age groups. For the older cohort, one factor, labeled *g* + *P* emerged from the HNTLA data; however, for the younger group, two principal factors were extracted. The first of these factors had a similar character for both D/HOH and hearing children and accounted for similar amounts of variance in both groups' scores; thus, it was referred to as *g* + *P*. The second factor accounted for similar (though small) amounts of variance in each group, but the Burt-Tucker Coefficient of Congruence showed this factor exhibited a different pattern of relationships within each of the two groups. Results yielded from both these studies led Braden to conclude that "deaf and hearing people do not differ in flexible, intelligent, nonverbal problem-solving ability" (Braden, 1985, p. 499). Differences were found in the character of secondary factors for younger children on the HNTLA. Braden interprets this to mean that nonverbal intelligence may not be completely identical for both groups, but “differences are small and ancillary to general nonverbal reasoning structures” (p. 500). However, younger D/HOH children appear to vary more from their normal-hearing peers of similar age than do older D/HOH children.

In an effort to better understand the similarities and differences in the developmental progression of cognitive structure of D/HOH and hearing children, Zweibel and Mertens (1985) conducted separate factor analyses for younger (6 – 9), middle (10 – 12), and older (13 – 15) groups. This procedure was intended to improve upon other studies for which data were aggregated across age groups. One psychologist administered the Snijders-Oomen Nonverbal Intelligence Test (SON) to all children and Zweibel and Mertens analyzed the results via principal axis factoring and Varimax rotation. Analyzing D/HOH and hearing groups as a whole yielded one factor for D/HOH children (overall general intelligence that is heavily perception-oriented) and two for hearing children (overall general intelligence and abstract thinking).

When analyzed by age, Zweibel and Mertens’ results differed between both D/HOH and hearing children of the same age and D/HOH children of different ages. For D/HOH children aged 6 to 9, a three-factor solution was found consisting of: (a) General Comprehension and Perceptual Thinking, (b) Perceptual and Concrete Thinking, and (c) Picture Memory. For the D/HOH children in the middle age group (10 to 12), two factors emerged, General Intelligence and Perceptual Skills. Figure Analogy loaded heavily for the hearing group but not for the D/HOH group suggesting abstract thinking skills to be absent or weak for D/HOH children (or problems with the way in which these abilities were measured by the test). For the older D/HOH children (13 to 15), an abstract thinking factor emerged similar to the one found for hearing children, along with a general intelligence factor. Differences within this age group included perception orientation to constitute a greater portion of the general intelligence factor for D/HOH than for hearing children and a greater emphasis on memory skills in the abstract thinking factor for D/HOH than for hearing children. Importantly, memory was found to contribute largely to the intellectual structure of D/HOH children across age groups.

In sum, these factor-analytic results suggest that disaggregating data by age level for D/HOH children may yield different results, leading Zweibel and Martens to suggest the "trend in the development of cognitive structure for the deaf subjects appears to be from a less organized to a more organized state of general intelligence and from a perceptual and visual orientation to a perceptual and abstract thinking orientation" (p. 30). Similar factor structures exist for older D/HOH and hearing children, but not for younger children, suggesting these two groups demonstrate different patterns of cognitive development but ultimately approach similar cognitive structures (i.e., equifinality). Braden (1994) suggests that D/HOH children's nonverbal intellective structures begin to look similar to those of their normal-hearing peers at about 11 years of age.

Overall, results from these factor analytic investigations are mixed, making it difficult to synthesize the findings. First, it is difficult to compare results yielded from studies that employ different factor analytic techniques (e.g., Oblimin versus Varimax rotations). Second, neither statistical procedure (i.e., oblique [Oblimin] or orthogonal [Varimax] rotation of extracted factors) is hierarchical in nature; thus, both are out of line with contemporary intelligence theories which posit all cognitive abilities are hierarchically organized. Third, the methods used in these studies (all variants of EFA) do not test the fit of the model to the data. That is, EFA techniques do not account for the fact that these intelligence tests were developed according to a theoretical model. Fourth, there is some evidence to suggest that results may vary across the developmental span from early childhood through adolescence, with younger D/HOH children showing greater differences from their normal-hearing peers than older D/HOH children and adolescents.

As a result of these issues, it is preferable to use CFA, a newer technique that allows for a more direct test of the degree to which factor structures are consistent with theory. The extant literature includes two studies that have employed MCFA to determine similarity of factor structure in relatively heterogeneous groups of D/HOH and hearing children (Maller & Ferron, 1997; Maller & French, 2004). Although results suggest some differences in the factor structures of D/HOH and normal-hearing children, intelligence is similarly organized for both groups.

In the first study, the factor structures of the WISC-III were compared for a sample of severely and profoundly deaf children (using a version of the test translated into American Sign Language) and the standardization sample (Maller & Ferron, 1997). Covariance matrices of the scaled scores from each of the samples were used to conduct MCFA. Maximum likelihood estimation was used to test the fit of the four-factor WISC-III model for both groups. Because an adequate fit was found for deaf and hearing children, parameter estimates were constrained across groups to test their invariance. Although constraint resulted in some statistically reliable differences for some minor parameters, the essential organization of the four-factor model (which was consistent with the theoretical grounding of the test, and fit the data from the normative sample) held for the D/HOH sample.

The second study examined the factor structure of the UNIT (Maller & French, 2004). Covariance matrices of the scaled scores for a sample of D/HOH participants and the standardization sample were used to test maximum likelihood estimation of a common factor structure for a D/HOH sample and a normal-hearing sample. The fit of each of two theoretical models was tested separately for each sample using the following fit indices: GFI, CFI, and RMSEA. The primary model is comprised of Memory and Reasoning and the second, Symbolic and Nonsymbolic factors. Both the primary and secondary models adequately fit for both samples, supporting both theoretical models. As a result of the invariance, further analyses were conducted. Important findings from these analyses include: (a) invariance of all pattern coefficients except Mazes for the primary model, (b) D/HOH children to score lower than hearing children on Analogic Reasoning, (c) D/HOH children to score lower than hearing children on the memory subtests, and (d) invariance of three pattern coefficients for the secondary model. That is, normal-hearing and D/HOH children scored significantly differently on the Cube Design, Spatial Memory, and Mazes subtests, which comprise the Nonsymbolic factor. Thus, it can be said that this factor does not fit for D/HOH children, which led Maller and French to conclude that the primary model is preferable over the secondary. They describe lack of invariance with regard to pattern coefficients as the "most critical concern regarding construct validity…because pattern coefficients indicate the relationship between variable and factor" (p. 657). This finding of partial measurement invariance of the primary model suggests that intelligence, as it is conceptualized and measured here, is organized similarly for both groups, but that there are some reliable differences that exist as well.

The Present Study

The purpose of the current study is to build upon this existing research base by exploring whether the cognitive abilities of D/HOH people are organized similarly to those of normal-hearing people across the developmental span, from preschool through early adulthood. More specifically, the present study seeks to employ CFA (using hierarchical models) and MCFA methods to better understand whether age interacts with hearing status with respect to factor structures on cognitive test batteries. Research conducted by Bolton (1978), Braden (1985), and Zweibel and Mertens (1985) suggests differences in the developmental progression of cognitive ability structure between D/HOH and normal-hearing children (with samples of younger D/HOH children showing greater differences relative to same-age normal-hearing peers than samples of older D/HOH examinees); however, no study has disaggregated samples by age groups and then applied CFA/MCFA methods to test the fit of these data with the tests' theoretical models. A finding of parallel factor structures would validate both the tests as well as CHC theory, whereas findings of larger differences between younger than older groups suggests differences in the way cognitive abilities are organized in young D/HOH children as compared to their normal-hearing peers, suggesting developmental discontinuities with the application of CHC theory to D/HOH children.

Hypotheses

Previous research separating children into relatively homogeneous age groups suggests differences in the structure of cognitive abilities between D/HOH and hearing children of younger ages. However, for older children, few differences are found between D/HOH and normal-hearing samples. Measurement invariance testing (to determine whether intelligence tests are measuring the same construct among the two groups) will be conducted only if the structure of cognitive abilities is similar for both groups (i.e., factor models adequately fit the data from normal-hearing and D/HOH people). Therefore, the predictions for the proposed study reflect these findings:

1. The factor structures that describe the normative data for younger (preschool through age 10) normal-hearing children will not fit the sample data from D/HOH children of the same ages (according to a variety of fit statistics, including the CFI, NNFI, RMSEA, and SRMR). More specifically, the CFI and NNFI will be less than .90, the RMSEA will be greater than .08, and SRMR will be greater than .10. Thus, measurement invariance testing will not be conducted.
2. The factor structures describing the normative data for older (age 11 through adulthood) normal-hearing children/adults will demonstrate an adequate fit with the sample data from D/HOH children/adults of the same ages (according to a variety of fit statistics, including the CFI, NNFI, RMSEA, and SRMR). More specifically, the CFI and NNFI will be greater than or equal to .90, the RMSEA will be less than or equal to .08, and SRMR will be less than or equal to .10.

3. A weak level of invariance (i.e., configural, metric) will be obtained across older (age 11 through adulthood) normal-hearing children/adults and D/HOH children/adults of the same ages. More specifically, to establish configural invariance, the p-value associated with χ^2 will be greater than or equal to .05, the CFI and NNFI will be greater than or equal to .90, the RMSEA will be less than or equal to .08, and the SRMR will be less than or equal to .10. To establish metric invariance, $\Delta\chi^2$ will be nonsignificant at the $p < .05$ level.

Chapter 2

Methods

Bibliographic Sample

I first collected those studies included in Braden's (1994) research synthesis. I then identified additional studies through the following means: (a) searching computerized databases consisting of published and unpublished studies (i.e., PsycINFO, ERIC, Dissertations Abstracts International) for investigations examining intelligence with D/HOH people (using *deaf, hard of hearing, hearing-impaired, intelligence, intelligence testing, and cognitive testing* as keywords), (b) reviewing the reference lists of identified studies for additional studies, (c) entering all identified studies into the Social Sciences Citation Index (SSCI) search engine to discover other works citing those studies that might be eligible for inclusion, and (d) soliciting unpublished data from psychologists serving D/HOH populations via deafness, disability, and psychology-related Listservs (i.e., the National Association of School Psychologists [NASP], the NASP D/HOH Interest Group, the American Psychological Association [APA] Divisions 16 and 22, the North Carolina School Psychology Association [NCSPA], the New York Association of School Psychologists [NYASP], the Council of Directors of School Psychology Programs [CDSPP]), and direct contact with authors who have published one or more eligible study in the past 10 years. If the same dataset was published in multiple sources (e.g., dissertation, book chapter, peer-reviewed journal article), I selected the peer-reviewed source to include in this study. However, I still entered the other citations into SSCI and reviewed their reference lists for additional sources.

The studies selected through these methods were retained for further analysis if they met the following criteria: (a) use of a multi-test intelligence battery with deaf, hard of hearing, or hearing-impaired people, (b) sample(s) comprised of participants either 10 and younger or 11 and older, (c) inclusion of a correlation or covariance matrix based on the primary subtests of the intelligence test battery from the D/HOH sample, (d) the study was published in English, and (e) I had access to a correlation or covariance matrix based on the primary subtests of the intelligence test battery from either the normative sample or a comparative sample of normal-hearing peers.

Out of an initial pool of 894 studies, a total of 6 studies matched the specified criteria and were included in the final analyses. These 6 studies yielded 8 samples. Table 2 outlines descriptions of each sample according to the following characteristics: (a) age, (b) gender, (c) ethnicity, and (d) degree of hearing loss.

Table 2

Descriptive Statistics by Sample

Sample	D/HOH Sample	N	Age in Years	Gender	Ethnicity	Degree of Hearing Loss
1	Bolton (1978)	559	2.6 – 10.5	Not specified	Not specified	Not specified
2	Bolton (1978)	520	10.6 – 17.5	Not specified	Not specified	Not specified

Table 2 (continued)

3	Brinich (1981)	40	5 – 6.6	Not specified	White	>80dB loss in the better ear, loss before 12 months of age
4	Ensor & Phelps (1989)	185	16.1 – 21.3	105 males, 80 females	124 white, 59 black, 2 Hispanic	Moderate to profound, prelingually deaf
5	Ensor & Phelps (1989)	185	16.1 – 21.3	105 males, 80 females	124 white, 59 black, 2 Hispanic	Moderate to profound, prelingually deaf
6	Hanson, Hancock, & Kopra (1969)	122	5.8 – 10.10	68 males, 54 females	61 white, 24 black, 33 Hispanic, 4 unidentified	27 congenital endogenous, 32 congenital exogenous, 50 prelingual undetermined, 13 unidentified

Table 2 (continued)

7	Johnson (1966)	82	6 – 9	Not specified	Not specified	≥60dB in the better ear, loss before 12 months of age
8	Murphy (1957)	300	6.6 – 10.6	150 males, 150 females	Not specified	Not specified

To increase the total number of datasets included in the present study, I decided to collect those studies that met all criteria with the exception of the age delineation. A total of 9 additional studies were obtained, which yielded 13 "mixed" samples. These mixed samples were segregated from the non-mixed samples for analysis.

Instruments

Hiskey-Nebraska Test of Learning Aptitude (HNTLA). The HNTLA was designed to assess the cognitive skills of deaf children and adolescents and is intended for use with those aged 3 to 17. Five subtests (Bead Patterns, Memory for Color, Picture Identification, Picture Association, Paper Folding) are specific to those aged 3 to 10, four subtests (Memory for Digits, Puzzle Blocks, Picture Analogy, Spatial Reasoning) are specific to those aged 11 to 17, and three subtests (Visual Attention Span, Block Patterns, Completion of Drawings) extend across these age ranges. It is recommended that examiners use pantomimed instructions with hearing-impaired individuals (Hiskey, 1966).

Illinois Test of Psycholinguistic Abilities (ITPA). The ITPA was “developed for the purpose of identifying psycholinguistic abilities and disabilities in children between the ages of two and one-half and nine” (McCarthy & Kirk, 1963, p. 1). The individually administered test battery consists of nine subtests, which are based on three dimensions of a communication model: (1) channels of communication (auditory input/vocal output and visual input/motor output), (2) levels of organization (automatic-sequential and representational), and (3) psycholinguistic processes (decoding, association, and encoding). At the representational level, the decoding subtests include Auditory Decoding and Visual Decoding, the association subtests include Auditory-Vocal Association and Visual-Motor Association, and the encoding subtests include Vocal Encoding and Motor Encoding. At the automatic-sequential level, tests are not subdivided into processes but are instead “whole-level” due to a lack of relevant literature at the time the ITPA was devised. The automatic subtest is Auditory-Vocal Automatic and the sequencing subtests are Auditory-Vocal Sequencing and Visual-Motor Sequencing. No “suitable” visual-motor subtest at the automatic level could be identified so McCarthy and Kirk did not include one in the ITPA (McCarthy & Kirk, 1963, p. 12).

Kaufman Assessment Battery for Children (K-ABC). The K-ABC is an individually administered measure of intelligence and achievement that is intended for children aged 2 ½ through 12 ½ years. The K-ABC is comprised of 16 subtests; however, only 13 may be given to any one child. The K-ABC is based on a theoretical foundation that distinguishes between sequential and simultaneous processing. This foundation is intended to reflect Luria’s neuropsychological model. The Sequential Processing Scale focuses on the serial or temporal order of stimuli during problem solving, whereas the Simultaneous Processing Scale requires concurrent integration of stimuli to solve problems. The three subtests that make up the Sequential Processing Scale are Hand Movements, Number Recall, and Word Order and the seven subtests that contribute to the Simultaneous Processing Scale are Magic Window, Face Recognition, Gestalt Closure, Triangles, Matrix Analogies, Spatial Memory, and Photo Series. These two Mental Processing Scales minimize the role of language in performance. The Achievement Scale is made up of six subtests (Expressive Vocabulary, Faces & Places, Arithmetic, Riddles, Reading/Decoding, Reading/Understanding), which measure global or verbal intelligence and school achievement (Kaufman & Kaufman, 1983).

Universal Nonverbal Intelligence Test (UNIT). The UNIT is a measure of intelligence that is obtained nonverbally. It is intended for use with children and adolescents aged 5 through 17 years and is based on a hierarchical model of intelligence in which "*g*" is defined as "the ability to solve problems using memory and reasoning" (Bracken & McCallum, 1998, p. 12). The UNIT is comprised of six subtests, three of which assess memory (Symbolic Memory, Object Memory, Spatial Memory) and three of which assess reasoning (Analogic Reasoning, Cube Design, Mazes). These subtests are also grouped into those that are symbolic and nonsymbolic. Symbolic subtests (Symbolic Memory, Object Memory, Analogic Reasoning) require the use of concrete and abstract symbols to understand the environment, whereas nonsymbolic subtests (Spatial Memory, Cube Design, Mazes) are symbol-free and require the ability to perceive and make judgments about physical relationships in the environment. The UNIT's conceptual model is based on Wechsler's (1939) distinction between symbolic (verbal) and nonsymbolic performance, Jensen's (1980) notion that intelligence is characterized by memory and reasoning, and the *Gf-Gc* Model of fluid and crystallized abilities (Bracken & McCallum, 1998; McCallum, 2003).

Wechsler Intelligence Scale for Children (WISC). The Wechsler intelligence scales are individually administered tests that are designed to assess the intellectual functioning of children and adults. They are the most frequently used, as well as the most highly researched, intelligence tests (Zhu & Weiss, 2005). Wechsler's first test was the Wechsler-Bellevue I, an adult intelligence scale published in 1939. Since that time, Wechsler has continued to develop tests according to the notion that intelligence in global in nature. The Verbal and Performance scales that comprise the Wechsler tests act as two modes in which general intelligence may be expressed; they are not representative of distinct abilities (Kamphaus, 2005). For the present study, mainly the Performance tests were examined.

The WISC is intended for use with children and adolescents aged 5 to 15. The WISC is comprised of five main Verbal subtests (Information, Similarities, Arithmetic, Vocabulary, Comprehension) and one supplementary Verbal subtest (Digit Span), as well as five Performance subtests (Picture Completion, Picture Arrangement, Object Assembly, Block Design, Coding) and one supplementary Performance subtest (Mazes). These subtests combine to yield Verbal, Performance, and Full Scale IQ scores (Wechsler, 1949).

Wechsler Preschool and Primary Scale of Intelligence (WPPSI). The WPPSI is similar to the WISC in both form and content; however, it was developed to measure the intelligence of children aged four to six and one-half because "this age period constitutes a well-defined landmark in the young child's mental development" (Wechsler, 1967, p. 1). The WPPSI consists of 11 subtests, 5 Verbal (Information, Vocabulary, Arithmetic, Similarities, Comprehension), 5 Performance (Animal House, Picture Completion, Mazes, Geometric Design, Block Design), and 1 supplementary Verbal (Sentences). Three of these tests (Sentences, Animal House, Geometric Design) are specific to the WPPSI and replace the Digit Span, Picture Arrangement, Object Assembly, and Coding subtests of the WISC (Wechsler, 1967).

Wechsler Intelligence Scale for Children-Revised (WISC-R). The WISC-R is the revised version of the WISC and includes changes in standardization procedures, item content, administration, and scoring. Additionally, the age range of the battery is intended for those aged 6 to 16 (as opposed to 5 to 15) to reduce the overlap with the WPPSI and allow for more high schoolers to be assessed. However, all subtests from the WISC were retained in the WISC-R, roughly 83% of the items remain completely unchanged, and subtests continue to combine to yield Verbal, Performance, and Full Scale IQ scores (Doppelt & Kaufman, 1977; Wechsler, 1974).

Wechsler Adult Intelligence Scale-Revised (WAIS-R). The WAIS-R is the revised version of the WAIS and includes changes in the standardization sample, item content, administration, and scoring. However, it is composed of the same 11 subtests as the WAIS; *viz*., 6 Verbal (Information, Comprehension, Arithmetic, Similarities, Digit Span, Vocabulary) and 5 Performance (Digit Symbol, Picture Completion, Block Design, Picture Arrangement, Object Assembly). Like the other Wechsler tests, these scores combine to yield Verbal, Performance, and Full Scale IQ scores. The WAIS-R is appropriate for use with individuals aged 16 to 90 (Wechsler, 1981).

*Wechsler Intelligence Scale for Children-Third Edition (WISC-III).*The WISC-III is the revised version of the WISC-R. The WISC-III maintains the basic structure and content of its predecessor (with 80% of the items remaining unchanged) but includes updated norms, as well as minor changes to subtest content, administration, and scoring rules. Like the WISC-R, the WISC-III yields Verbal, Performance, and Full Scale IQ scores; however, the WISC-III also yields four factor indexes, namely Verbal Comprehension, Perceptual Organization, Freedom from Distractibility, and Processing Speed. The WISC-III Verbal Scale consists of five mandatory subtests (Information, Similarities, Arithmetic, Vocabulary, Comprehension) and one supplementary subtest (Digit Span). The Performance Scale consists of five mandatory subtests (Picture Completion, Picture Arrangement, Block Design, Object Assembly, Coding) and two supplementary subtests (Mazes, Symbol Search). All of these subtests were retained from the WISC-R except for Symbol Search, which is a new addition (Prifitera & Saklofske, 1998; Wechsler, 1991).

Analyses

For the present study, I used maximum likelihood estimation (MLE) to conduct CFA (with hierarchical models) and MCFA across samples of D/HOH and normative examinees. I originally proposed that I would implement MLE in AMOS 16.0; however, this program required covariance matrices as opposed to correlation matrices. Because I could not access the means and standard deviations for each matrix, I opted to use LISREL 8.8. The first step in these analyses was to identify which broad (stratum II) cognitive abilities were intended to be represented on each cognitive battery by "mapping" subtests onto the ability they tap. For the purposes of my study, I assumed the eight broad CHC abilities to include the following:

fluid reasoning (*Gf*), crystallized intelligence (*Gc*), auditory processing (*Ga*), visual processing (*Gv*), short-term memory (*Gsm*), long-term retrieval (*Glr*), processing speed (*Gs*), and occasionally math/quantitative knowledge (*Gq*; McGrew, 2005).

To assign subtests to broad abilities, I used those classifications for select major test batteries outlined in Table 3. This summary is a reproduction of classifications compiled by Alfonso, Flanagan, and Radwan (2005) and Flanagan and Mascolo (2005; also reproduced in Flanagan, Ortiz, Alfonso, & Dynda, 2008), and is based on the extant literature and primary sources. Importantly, each subtest was assigned to one and only one broad ability factor. For those cognitive batteries that are not reflected in Table 3, but for which data were obtained, I solicited assistance from two colleagues (doctoral students enrolled in the North Carolina State University school psychology program) to assign tests to ability factors. First, I assigned subtests from each of the intelligence batteries to a broad ability factor using the test manual's description of the abilities it intends to measure, and descriptions of the eight CHC broad abilities (McGrew, 2005). Please refer to Table 4 for these broad ability descriptions and to Appendices A through C for each test battery's coding worksheet. A graduate student colleague conducted the same analysis without knowledge of how I assigned tests to factors. I reviewed our assignments and determined six total inconsistencies (a 25% disagreement rate). I asked a second colleague to make a third assignment for those subtests on which the first rater and I disagreed. For four of these ratings, the second rater agreed with either me or the first rater. Thus, I assigned the subtest according to majority rating (i.e., to whichever broad ability received two "votes"). For two of these ratings, there

Table 3

Mapping of Subtests from Select Cognitive Batteries onto CHC Broad Abilities

Battery	Gf	Gc	Gv	Gsm	Glr	Ga	Gs
WISC-III	-----	Vocabulary, Information, Similarities, Comprehension	Block Design, Object Assembly, Picture Completion, Picture Arrangement	Digit Span	-----	-----	Symbol Search, Coding
WISC-IV	Matrix Reasoning, Picture Concepts, Arithmetic	Vocabulary, Information, Similarities, Comprehension, Word Reasoning	Block Design, Picture Completion	Digit Span, Letter-Number Sequencing	-----	-----	Symbol Search, Coding, Cancellation

Table 3 (continued)

Battery	Gf	Gc	Gv	Gsm	Glr	Ga	Gs
WAIS-III	Matrix Reasoning	Vocabulary, Information, Similarities, Comprehension	Block Design, Object Assembly, Picture Completion, Picture Arrangement	Digit Span, Letter-Number Sequencing	-----	-----	Symbol Search, Digit-Symbol Coding
WPPSI-R	-----	Vocabulary, Information, Similarities, Comprehension	Block Design, Object Assembly, Picture Completion, Mazes, Geometric Design	Sentences	-----	-----	Animal Pegs
WPPSI-III	Matrix Reasoning, Picture Concepts	Vocabulary, Information, Similarities, Comprehension, Receptive Vocabulary, Picture Naming, Word Reasoning	Block Design, Object Assembly, Picture Completion	-----	-----	-----	Coding, Symbol Search

Table 3 (continued)

Battery	Gf	Gc	Gv	Gsm	Glr	Ga	Gs
KAIT	Mystery Codes, Logical Steps	Definitions, Famous Faces, Auditory Comprehension, Double Meanings	Memory for Block Designs	-----	Rebus Learning, Rebus Delayed Recall, Auditory Delayed Recall	-----	-----
K-ABC	Matrix Analogies	-----	Triangles, Face Recognition, Gestalt Closure, Magic Window, Hand Movements, Spatial Memory, Photo Series	Number Recall, Word Order	-----	-----	-----

Table 3 (continued)

Battery	Gf	Gc	Gv	Gsm	Glr	Ga	Gs
KABC-II	Pattern Reasoning, Story Completion	Expressive Vocabulary, Verbal Knowledge, Riddles	Triangles, Gestalt Closure, Rover, Block Counting, Conceptual Thinking, Face Recognition	Number Recall, Word Order, Hand Movements	Atlantis, Rebus, Atlantis Delayed, Rebus Delayed	-----	-----
CAS	-----	-----	Figure Memory, Verbal Spatial Relations, Nonverbal Matrices	Word Series, Sentence Repetition, Sentence Questions	-----	-----	Matching Numbers, Receptive Attention, Planned Codes, Number Detection, Planned Connections, Expressive Attention

Table 3 (continued)

Battery	Gf	Gc	Gv	Gsm	Glr	Ga	Gs
DAS	Matrices, Picture Similarities, Sequential and Quantitative Reasoning	Similarities, Verbal Comprehension, Word Definitions, Naming, Vocabulary	Pattern Construction, Block Building, Copying, Matching Letter-Like Forms, Recall of Designs, Recognition of Pictures	Recall of Digits	Recall of Objects	-----	Speed of Information Processing

Table 3 (continued)

Battery	Gf	Gc	Gv	Gsm	Glr	Ga	Gs
WJ-R	Concept Formation, Analysis-Synthesis	Oral Vocabulary, Picture Vocabulary, Listening Comprehension, Verbal Analogies	Spatial Relations, Picture Recognition, Visual Closure	Memory for Words, Memory for Sentences, Numbers Reversed	Memory for Names, Visual-Auditory Learning, Delayed Recall: Memory for Names, Delayed Recall: Visual-Auditory Learning	Incomplete Words, Sound Blending, Sound Patterns	Visual Matching, Cross Out

Table 3 (continued)

Battery	Gf	Gc	Gv	Gsm	Glr	Ga	Gs
WJ-III	Concept Formation, Analysis-Synthesis	Verbal Comprehension, General Information	Spatial Relations, Picture Completion	Memory for Words, Numbers Reversed, Auditory Working Memory	Visual-Auditory Learning, Visual-Auditory Learning Delayed, Retrieval Fluency, Rapid Picture Naming	Incomplete Words, Sound Blending, Auditory Attention	Visual Matching, Decision Speed
SB-IV	Matrices, Equation Building, Number Series	Verbal Relations, Comprehension, Absurdities, Vocabulary	Pattern Analysis, Bead Memory, Copying Memory for Objects, Paper Folding and Cutting	Memory for Sentences, Memory for Digits	-----	-----	-----

Table 3 (continued)

Battery	Gf	Gc	Gv	Gsm	Glr	Ga	Gs
SB5	Nonverbal Fluid Reasoning, Verbal Fluid Reasoning, Nonverbal Quantitative Reasoning, Verbal Quantitative Reasoning	Nonverbal Knowledge, Verbal Knowledge	Nonverbal Visual-Spatial Processing, Verbal Visual-Spatial Processing	Nonverbal Working Memory, Verbal Working Memory	-----	-----	-----
RIAS	Odd-Item Out	Guess What, Verbal Reasoning	What's Missing	Verbal Memory, Nonverbal Memory	-----	-----	-----

Table 4

CHC Broad Ability Factors

Broad Ability Factor	Description
Fluid Intelligence/reasoning (*Gf*)	The ability to use inductive and deductive reasoning to solve novel problems using such mental operations as drawing inferences, concept formation, classification, identifying relations, comprehending implications, and transforming information.
Crystallized intelligence/knowledge (*Gc*)	A person's breadth and depth of acquired knowledge of the language, information, and concepts of a specific culture and his/her ability to apply this knowledge. This knowledge is obtained through formal and informal experiences within the mainstream culture.
Auditory processing (*Ga*)	The ability to analyze, manipulate, comprehend, and synthesize auditory (sound-based) information.

Table 4 (continued)

Visual-spatial abilities (*Gv*)	The ability to generate, store, retrieve, and transform visual patterns and spatial configurations.
Short-term memory (*Gsm*)	The ability to hold information in mental awareness and use it within a minute or so.
Long-term storage and retrieval (*Glr*)	The ability to store and retrieve information through association.
Processing speed (*Gs*)	The ability to perform relatively simple cognitive tasks automatically, particularly when attention and concentration are required.
Quantitative knowledge (*Gq*)	An individual's store of mathematical knowledge (not one's ability to reason with this knowledge) that is primarily acquired through formal educational experiences.

did not exist a majority vote (i.e., the third rater assigned the subtest to a different factor from either the first rater or me). Here, I asked my dissertation chair (Dr. Jeffery Braden) to conduct an independent rating. If he agreed with any of the first three ratings, his rating was final. I originally proposed that if he did not agree with any of the first three ratings, those subtests for which there was no agreement would be omitted from analyses. However, no subtests were omitted as a result of rater disagreement. I did exclude the Bead Pattern subtest of the HNTLA from the analyses because, although there exist three versions of this subtest on the test battery (Stringing Beads, Copying Bead Patterns, and Memory for Bead Patterns), they are combined into one on the intercorrelation table, and the independent CHC ratings conducted as part of this study suggest each version taps a different broad ability. Also, although the WISC, WISC-R, and WPPSI are not included in Table 3, the more updated versions (i.e., the WISC-III and the WPPSI-R) of these tests do appear on the table. Because of the high overlap in item content (with at least 60% of the items remaining unchanged) and the lack of change in required response processes from one test version to the next, I decided to use the subtest mappings already included on the table. (Please note that Animal Pegs is Animal House renamed.) Additionally, I was forced to drop four datasets obtained for cognitive batteries that required ratings but for which a test manual was unavailable and three others as a result of transcription errors.

The following paragraphs describe how I analyzed the collected data, and these steps are displayed in Figure 1. First, I generated factor models for each of the cognitive batteries for which a covariance/correlation matrix was obtained. Each model was higher-order in nature and had two components: (a) one third-order factor (*g*), and (b) as many second-order factors as there are broad CHC abilities represented by the test. I designed models so that subtests exclusively loaded (1.0) on their assigned second-order abilities, and these second-order abilities loaded (freely) on *g* (Keith, 2005). In cases where subtests may have been administered to the normative group but not to the D/HOH sample (e.g., omission of verbal subtests, omission of supplemental tests), I analyzed only those subtests administered to both groups. Please note that, for most of the factor models included in this study, no subtests loaded onto Gc (or many of the other broad ability factors). Therefore, the range of CHC factors represented within these models is restricted relative to more elaborate CHC models. Please refer to Appendices D through K for all factor models.

Following the specification of the latent models, the second step was to test each model for the normative sample using the CFI, NNFI, RMSEA, and SRMR to determine if the expected model fit the normative data (per the recommendation of Hu and Bentler, 1999). Importantly, if a test manual offered correlation matrices disaggregated by age group, I averaged the matrices that allowed for the greatest overlap in age range with the D/HOH sample data. If an adequate fit was found for the normative sample data, I applied this same model to the D/HOH sample in the third step. If the anticipated model did not adequately fit the normative sample data, I respecified the model by examining the modification indices for those parameter estimates constrained to equal zero and freely estimating the parameter with the largest modification index. If, after freely estimating this parameter, the model still did not adequately fit the data, I then freed the parameter with the next highest modification index, and so on (Jö reskog, 1993). The modification index "estimates the amount by which the overall model chi-square statistic…would decrease if a particular fixed-to-zero path were freely estimated" (Kline, 2005, p. 148). The larger the index, the better the predicted

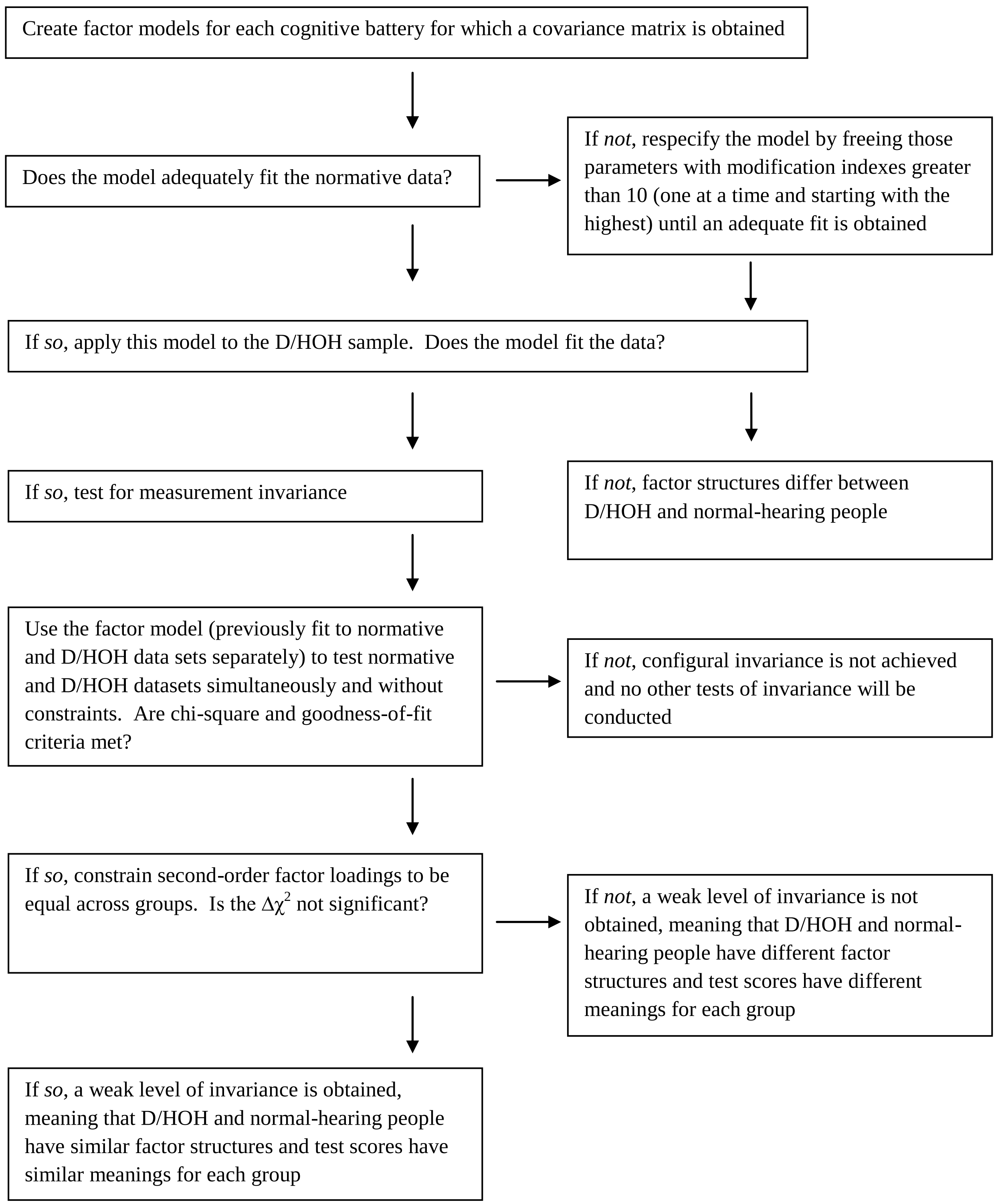


Figure 1. A Step-by-Step Plan of Analysis.

improvement in model fit. Generally, modification indices should be examined if they are greater than 10 (P. J. Curran, personal communication, September 27, 2010), so subtests with indices greater than 10 were freely estimated (one by one), starting with the highest, for the normal-hearing group onto the specified ability factor until an adequate fit was obtained. In cases where the model did not provide a reasonable fit for the D/HOH group, I determined that factor structures differed between the two groups, and I did not conduct measurement invariance testing. If I found an adequate fit, I tested for measurement invariance. The process for testing invariance is discussed below.

Testing for measurement invariance involves multiple steps, each of which requires the application of more constraints; thus, the models are nested within one another and the likelihood ratio test (i.e., $\Delta\chi^2$) is used to compare models. A significant $\Delta\chi^2$ (here, $p < .05$) demonstrates that additional constraints caused a significant change in model fit and thus there exists a lack of invariance between samples. The χ^2, CFI, NNFI, and RMSEA were used to assess the fit of single models as opposed to competing nested models. Although there are many ways to test invariance (i.e., equality of covariance matrices, equality of factor loadings, equality of the number of common factors, equality of factor variances, equality of factor intercorrelations, and equality of unique/error variances), group comparisons require the investigator to establish a baseline model that has the same factor pattern across groups (Meredith, 1993). Doing so establishes configural invariance, suggesting equality in the number of salient factor loadings (but not necessarily in the magnitude of these loadings; Schaie, Maitland, Willis, & Intrieri, 1998) between groups. Thus, the configural invariance model was identical to the specified higher-order factor model (previously fit to the normal-hearing and D/HOH datasets separately), but was fit to the normal-hearing and D/HOH datasets simultaneously and without constraints. Failure to achieve this level of invariance precluded further invariance testing. Therefore, if I was unable to establish a model that met criteria for goodness-of-fit across both the normal-hearing and D/HOH samples, I did not apply any additional tests of model invariance. If the baseline model did not differ between the two samples, I conducted the second test, that of metric invariance, by constraining the second-order factor loadings (i.e., loadings from subtests onto broad abilities) to be equal across the two groups.

Because subsequent tests of invariance become increasingly more stringent, Meredith (1993) delineates three levels of factorial invariance depending on the level achieved: (a) weak, which exists when factor loadings are constrained to be equal across groups; (b) strong, which holds when item intercepts, in addition to factor loadings, are equivalent across groups; and (c) strict, which occurs when item uniquenesses (i.e., unique/error variances), in addition to factor loadings and item intercepts, are equivalent across groups. Although more stringent tests of invariance can be conducted, they were not included as part of the current study for the following reasons: (a) raw data was not be solicited, and therefore invariance at the item intercept level could not be determined, (b) more complex levels of invariance are not theoretically meaningful, and (c) most importantly, only configural and metric invariance are necessary for unambiguous interpretation of invariance (here, factor structure similarity) between groups. That is, metric invariance provides “support for a hypothesis of measurement invariance” and “is a reasonable ideal for research in the behavioral sciences” (Horn, 1991, p. 124).

For each comparison, confirmatory factor analyses determined whether cognitive abilities are organized similarly (i.e., factor structures are similar) in D/HOH and normal-hearing people. Factor structures were deemed similar if the identified factor model adequately fit the data from both the normal-hearing and D/HOH samples according to the fit statistics (CFI, NNFI, RMSEA, and SRMR). I argue that D/HOH and normal-hearing people have different factor structures if the factor model fitting the normative sample data did not adequately fit the D/HOH sample data. If factor structures were deemed similar, measurement invariance testing determined whether intelligence tests were measuring the same construct among D/HOH and normal-hearing people. I propose that tests are measuring the same construct among the two groups if a weak level of invariance (i.e., configural, metric) was achieved. Configural invariance was determined according to the following fit statistics: χ^2, CFI, NNFI, and RMSEA. Metric invariance was determined according to the $\Delta\chi^2$ statistic; however, I also examined the other goodness-of-fit statistics (χ^2, CFI, NNFI, and RMSEA) for descriptive purposes. (Please note that the SRMR was not available for the metric invariance model.) Conversely, tests are not measuring the same construct among these two groups if a weak level of invariance was not obtained. Prediction 1 was supported if the majority of comparisons suggested differences in the factor structures of D/HOH and normal-hearing children aged 10 and younger. Prediction 2 was supported if the majority of comparisons suggested similarities in the factor structures of D/HOH and normal-hearing children/adults aged 11 and older. Prediction 3 was supported if the majority of comparisons yielded a weak level of invariance across older normal-hearing and D/HOH individuals. I attempted to identify patterns of consistency (e.g., some tests simply do not adhere to the predicted model) or inconsistency (e.g., some tests appear to adhere in one sample but not in another) in subsequent discussion to describe the factors that support or undermine my hypotheses.

Chapter 3

Results

Hypotheses

The three hypotheses in this study concerned the interaction between age and hearing status as it relates to factor structures on cognitive test batteries. Hypothesis One predicted that factor structures derived from CHC theory will describe the normative sample data for younger normal-hearing children (preschool through age 10), but will not describe the data for D/HOH children of the same ages according to various goodness-of-fit statistics. In contrast, Hypothesis Two predicted that factor structures derived from CHC theory will describe the normative sample data for older normal-hearing children/adults (age 11 through adulthood) and will describe the data for D/HOH children/adults of the same ages according to the same fit statistics. Hypothesis Three predicted that a weak level of invariance (defined here as configural and metric) will be obtained across older (age 11 through adulthood) normal-hearing children/adults and D/HOH children/adults of the same ages. To test the first two hypotheses, I compared the normative sample data for each test battery to its respective CHC hierarchical factor model and subjected the models to CFA. I then examined four goodness-of-fit indices to determine the fit of the models (i.e., CFI, NNFI, RMSEA, SRMR). Tables 5 and 6 outline the results of these CFAs for normal-hearing younger children and older children/adults respectively.

All CHC-based models, with the exception of the model applied to the HNTLA, provided an excellent fit to the normative data for the younger group, with fit indices falling well within their criterion values (i.e., CFI and NNFI > .90; RMSEA ≤ .08; SRMR ≤ .10).

Table 5

Results of Confirmatory Factor Analyses for Normal-Hearing Children of Preschool Age through Age 10

Test Battery	Normative Sample (Sample Portion Used)	*N*	CHC Factors	CFI	NNFI	RMSEA	SRMR
HNTLA	Hiskey (1966) (3 – 10 years)	617	*Gv*, *Gsm*, *Gf*	.90	.82	.25	.08
ITPA	McCarthy & Kirk (1963) (5.5 – 9 years)	400	*Gc*, *Gf*, *Gv*	1.00	1.07	.00	.00
ITPA	McCarthy & Kirk (1963) (6 – 9 years)	350	*Gc*, *Gf*, *Gv*	1.00	1.00	.00	.00
WISC	Wechsler (1949) (7.5 – 10.5 years)	400	*Gv*, *Gs*	.99	.98	.05	.03
WPPSI	Wechsler (1967) (5 – 6.5 years)	800	*Gv*, *Gs*	1.00	.99	.04	.02

The model applied to the HNTLA evidenced a poor fit to the data. Although the CFI and SRMR were acceptable, the NNFI and RMSEA fell considerably outside the adequate range. Thus, I examined the modification indices to determine which parameters might improve model fit. Overall, three modification indices revealed significant values (i.e., were greater than 10; Memory for Color onto *Gv***,** Visual Attention Span onto *Gv*, and Block Patterns onto *Gsm*). As proposed, these parameters were freed one-by-one starting with the highest; however, in all cases, results were worse than those of the original model; thus, this dataset was deleted from subsequent analyses.

Table 6

Results of Confirmatory Factor Analyses for Normal-Hearing Children/Adults Aged 11 through Adulthood

Test Battery	Normative Sample (Sample Portion Used)	*N*	CHC Factors	CFI	NNFI	RMSEA	SRMR
HNTLA	Hiskey (1966) (11 – 17 years)	460	*Gv, Gsm, Gf*	.99	.98	.05	.03
WAIS-R	Wechsler (1981) (16 – 24 years)	600	*Gv*, *Gs*	.97	.94	.10	.04

For the older normal-hearing group, the CHC model fit the HNTLA normative data well. The CHC model fit the WAIS-R normative data well by all standards except the RMSEA (.10). The extant literature provides a range of acceptable cutoff points for each of the fit indices, but cautions against strict adherence to these numbers, and supports evaluating fit according to a variety of indices (Hu & Bentler, 1999; Kline, 2005; Vandenberg & Lance, 2000). Hu and Bentler (1999) suggest that an RMSEA value less than or equal to .06 indicates good fit, while Browne and Cudeck (1993) recommend that “a value of the RMSEA of about .05 or less would indicate a close fit of the model in relation to the degrees of freedom," and that "the value of about .08 or less for the RMSEA would indicate reasonable error of approximation and we would not want to employ a model with a RMSEA greater than .1" (p.144). Thus, .10 is the upper limit of acceptability according to Browne and Cudeck, and many researchers argue for lower values. However, the other fit indices indicated adequate fit and there does not exist a theoretically defensible alternative; therefore, I decided not to respecify the model.

In the next step, I applied the aforementioned CHC factor models to the D/HOH sample data, subjected these models to CFA, and evaluated the resulting goodness-of-fit indices. Tables 7 and 8 present the results of these CFAs for younger and older D/HOH samples respectively. The first column in these tables lists the test battery administered and the portion of the normative sample from that test battery that most closely matches the ages of those individuals included in the D/HOH study. For those factor models that did not provide a reasonable fit for the D/HOH sample (i.e., CFI and NNFI < .90, RMSEA > .08, SRMR >.10), I determined that normal-hearing and D/HOH factor structures differed and discontinued testing. For those factor models offering an adequate fit, I tested for measurement invariance.

Results from the application of the CHC factor models to the younger D/HOH data are shown in Table 7. Although the CHC models offered an adequate fit to the Hanson, Hancock, and Kopra ITPA data, as well as the Johnson ITPA data according to the CFI and SRMR, the NNFI and RMSEA values suggested a very poor fit. Thus, I did not include these models in any further testing and determined that the normal-hearing and D/HOH factor structures' differed in these cases. However, the CHC models provided an excellent fit

Table 7

Results of Confirmatory Factor Analyses for D/HOH Children of Preschool Age through Age 10

Test Battery (Sample Portion Used)	D/HOH Study	*N*	CHC Factors	CFI	NNFI	RMSEA	SRMR
ITPA (5.5 – 9 years)	Hanson, Hancock, & Kopra (1969)	122	*Gc*, *Gf*, *Gv*	.91	.47	.18	.05
ITPA (6 – 9 years)	Johnson (1966)	82	*Gc*, *Gf*, *Gv*	.93	.59	.22	.06
WISC (7.5 – 10.5 years)	Murphy (1957)	300	*Gv*, *Gs*	.99	.99	.06	.02
WPPSI (5 – 6.5 years)	Brinich (1981)	39	*Gv*, *Gs*	1.00	1.06	.00	.05

to the Murphy WISC and Brinich WPPSI datasets. As shown in Table 8, all CHC models fit the data from the older D/HOH children/adults well; thus, Hypothesis Two was supported and all models were tested for measurement invariance (Hypothesis Three). That is, I ran multisample CFAs to establish invariance across the normal-hearing and D/HOH groups. First, I examined whether the baseline (configural invariance) model of interest fit across the

Table 8

Results of Confirmatory Factor Analyses for D/HOH Children/Adults Aged 11 through Adulthood

Test Battery (Sample Portion Used)	D/HOH Study	*N*	CHC Factors	CFI	NNFI	RMSEA	SRMR
HNTLA (11 – 17 years)	Bolton (1978)	529	*Gv*, *Gsm*, *Gf*	1.00	1.00	.00	.02
WAIS-R (16 – 24 years)	Ensor & Phelps (1989), Females	80	*Gv*, *Gs*	.98	.96	.09	.04
WAIS-R (16 – 24 years)	Ensor & Phelps (1989), Males	105	*Gv*, *Gs*	1.00	1.00	.02	.03

normal-hearing and D/HOH groups by testing the two datasets simultaneously but without constraints. If this model proved to be well-fitting, I continued with metric invariance testing, in which like factor loadings were constrained to be equal across the two groups (e.g., path of subtest 1 to normal-hearing = path of subtest 1 to D/HOH). Here, I constrained the paths from subtests to second-order factors to be equal. Results of these invariance tests are shown on the next page in Table 9.

Table 9

Results from the Tests of Invariance for D/HOH Children/Adults Aged 11 through Adulthood

Test Battery	Normative Sample (Sample Portion Used)	D/HOH Study	Model	$X^2(df)$	*P*	CFI	NNFI	RMSEA	$\Delta\chi^2(\Delta df)$	*p*
HNTLA	Hiskey (1966)	Bolton (1978)	Configural	32.07(24)	.13	1.00	.99	.03	-----	-----
	(11 – 17 years)		Metric	38.55(30)	.14	1.00	1.00	.02	6.48(6)	NS
WAIS-R	Wechsler (1981)	Ensor &	Configural	43.35(10)	.00	.97	.94	.10	-----	------
	(16 – 24 years)	Phelps (1989),	Metric	46.59(13)	.00	.97	.95	.09	3.24(3)	NS
		Females								
WAIS-R	Wechsler (1981)	Ensor &	Configural	40.11(10)	.00	.97	.95	.09	-----	-----
	(16 – 24 years)	Phelps (1989),	Metric	40.43(13)	.00	.98	.96	.08	0.32(3)	NS
		Males								

The CHC configural invariance model offered an excellent fit to the HNTLA normal-hearing data provided by Hiskey and the HNTLA D/HOH data provided by Bolton. All other configural invariance models were rejected statistically ($p < .05$), but not on the basis of the criterion values for the remaining fit indices. As noted previously, because the value of p associated with the χ^2 statistic is highly affected by sample size, large samples may “lead to rejection of the model even though differences between observed and predicted covariances are slight” (Kline, 2005, p. 136). That is, large samples may yield statistically significant results (i.e., unlikely outcomes), but these results are not necessarily of a sufficient magnitude to be important in a practical sense. Therefore, I deemed the configural invariance models plausible and proceeded with the next test, that of metric invariance. Similarly, the metric models were rejected statistically ($p < .05$), but all other fit indices indicated adequate fit. Most importantly, the nonsignificant differences in χ^2 ($\Delta\chi^2$) between the models supported the notion that the application of invariance constraints imposed by the metric model did not worsen the fit of the configural model, thereby supporting Hypothesis Three.

Supplemental Analyses

In addition to testing each of the proposed hypotheses, I conducted supplemental analyses to further examine the relationship between age, hearing status, and cognitive test battery factor structure. More specifically, as with the older samples, I continued with measurement invariance testing with those younger samples for which evidence suggested similar factor structures between normal-hearing and D/HOH individuals. Results are shown in Table 10. Additionally, I performed the same sequence of analysis with mixed-age

Table 10

Results from the Tests of Invariance for D/HOH Children of Preschool Age through Age 10

Test Battery	Normative Sample (Sample Portion Used)	D/HOH Study	Model	$X^2(df)$	*P*	CFI	NNFI	RMSEA	$\Delta\chi^2(\Delta df)$	*p*
WISC	Wechsler (1949)	Murphy	Configural	20.38(10)	.03	.99	.98	.06	-----	-----
	(7.5 – 10.5 years)	(1957)	Metric	21.51(13)	.06	.99	.99	.04	1.13(3)	NS
WPPSI	Wechsler (1967)	Brinich	Configural	19.89(10)	.03	.99	.98	.05	-----	-----
	(5 – 6.5 years)	(1981)	Metric	16.06(13)	.25	1.00	1.00	.02	3.83(3)	NS

samples as I did with the younger and older samples (CFAs followed by invariance testing). These results appear in Tables 11 through 13.

Results of invariance testing with the younger samples are shown in Table 10. Here, the CHC configural models (fit to the two WISC datasets and to the two WPPSI datasets) were rejected statistically ($p < .05$), but the other fit indices indicated excellent fit. Thus, I imposed metric invariance models onto the data, which resulted in nonsignificant $\Delta\chi^2$s. Additionally, all other standards evidenced an excellent fit. Thus, the models were supported and I could conclude that they were invariant across the normal-hearing and D/HOH groups. Because the majority of comparisons between normal-hearing and D/HOH children aged 10 and younger did not suggest differences in factor structure, Hypothesis One was not supported. More specifically, the CHC models that adequately fit the normative data proved ill-fitting when applied to the D/HOH sample data for only two of four younger samples (Hanson, Hancock, & Kopra ITPA; Johnson ITPA). (CHC models adequately fit the D/HOH data for the Murphy WISC and Brinich WPPSI samples.)

Table 11

Results of Confirmatory Factor Analyses for Normal-Hearing Children/Adults of Mixed Ages

Test Battery	Normative Sample (Sample Portion Used)	*N*	CHC Factors	CFI	NNFI	RMSEA	SRMR
ITPA	McCarthy & Kirk (1963) (6 – 9 years)	350	*Gc, Gf, Gv*	1.00	1.08	.00	.00
K-ABC	Kaufman & Kaufman (1983) (5 – 12.5 years)	1500	*Gv, Gsm, Gf*	.98	.97	.06	.04
UNIT	Bracken & McCallum (1998) (Entire sample)	2100	*Gv, Gf*	.99	.96	.09	.02
UNIT	Bracken & McCallum (1998) (Entire sample)	2100	*Gv, Gsm, Gf*	.98	.96	.07	.03
WISC	Wechsler (1949) (Entire sample)	600	*Gv, Gs*	.99	.96	.08	.03
WISC-III	Wechsler (1991) (Entire sample)	2200	*Gv, Gsm, Gs*	.99	.98	.05	.03

Table 11 (continued)

WISC-R	Wechsler (1974) (Entire sample)	2200	*Gv, Gs*	.99	.98	.05	.02
WISC-R	Wechsler (1974) (7.5 – 16.5 years)	2000	*Gv, Gs*	1.00	.99	.03	.02
WISC-R	Wechsler (1974) (7.5 – 13.5 years)	1400	*Gv, Gs*	.99	.98	.06	.03

All CHC models proved to be generally well-fitting when applied to the normative data; thus, I applied them to the D/HOH sample data. Results appear in Table 12.

As shown in Table 12, the CHC models that adequately fit the normative sample data were ill-fitting, according to a variety of fit indices, when applied to these mixed-age D/HOH samples: (1) Hughes, Sapp, and Kohler UNIT and WISC-III, (2) Hirshoren, Hurley, and Hunt WISC-R, (3) Sisco (females) WISC-R, and (4) Welch WISC-R. These ill-fitting models provide evidence of differing factor structures between the two groups. Although the fit of the CHC models to the Ulissi, Brice, and Gibbins K-ABC sample and Phelps and Ensor (males) WISC-R sample yielded RMSEAs that fell well outside the acceptable range, their sample sizes were quite small ($N < 100$), which can lead to greater error and therefore higher values. Therefore, I decided to subject these two models to invariance testing. CHC models fit the following D/HOH samples generally well: (1) Espeseth ITPA, (2) Krivitsi, McIntosh, Rothlisberg, and Finch UNIT, (3) Brooks and Riggs WISC, (4) Brooks and Riggs WISC-R,

(5) Phelps and Ensor (females) WISC-R, and (6) Sisco (males) WISC-R. Therefore, the models were subjected to invariance testing.

Table 12

Results of Confirmatory Factor Analyses for D/HOH Children/Adults of Mixed Ages

Test Battery (Sample Portion Used)	D/HOH Study	*N*	CHC Factors	CFI	NNFI	RMSEA	SRMR
ITPA (6 – 9 years)	Espeseth (1965)	36	*Gc, Gf, Gv*	1.00	1.16	.00	.03
K-ABC (5 – 12.5 years)	Ulissi, Brice, & Gibbins (1989)	50	*Gv, Gsm, Gf*	.94	.91	.13	.07
UNIT (Entire sample)	Hughes, Sapp, & Kohler (2006)	32	*Gv, Gf*	.73	.19	.39	.10
UNIT (Entire sample)	Krivitski, McIntosh, Rothlisberg, & Finch (2004)	39	*Gv, Gsm, Gf*	.98	.95	.09	.07
WISC (Entire sample)	Brooks & Riggs (1980)	40	*Gv, Gs*	1.00	1.07	.00	.04

Table 12 (continued)

WISC-III (Entire sample)	Hughes, Sapp, & Kohler (2006)	32	*Gv*, *Gsm*, *Gs*	.86	.74	.19	.09
WISC-R (Entire sample)	Brooks & Riggs (1980)	40	*Gv*, *Gs*	1.03	1.09	.00	.02
WISC-R (7.5 – 13.5 years)	Hirshoren, Hurley, & Hunt (1977)	59	*Gv*, *Gs*	.94	.88	.15	.06
WISC-R (Entire sample)	Phelps & Ensor (1987), Females	60	*Gv*, *Gs*	.99	.99	.06	.03
WISC-R (Entire sample)	Phelps & Ensor (1987), Males	65	*Gv*, *Gs*	.96	.93	.12	.06
WISC-R (Entire sample)	Sisco (1982), Females	536	*Gv*, *Gs*	.97	.95	.12	.04
WISC-R (Entire sample)	Sisco (1982), Males	669	*Gv*, *Gs*	.98	.97	.09	.03
WISC-R (7.5 – 16.5 years)	Welch (1992)	87	*Gv*, *Gs*	.94	.88	.14	.07

Table 13 shows the results of the invariance tests for the mixed-age samples. All configural models were rejected statistically ($p < .05$) but were not rejected according to the remaining fit indices. As previously noted, the chi-square statistic is affected by sample size; therefore, I continued with metric invariance testing in all cases. This application of constraints resulted in a statistically significant change according to the $\Delta\chi^2$ criterion for the following CHC models: (1) ITPA, (2) WISC-R (as applied to the normative and Phelps and Ensor [females] samples), and (3) WISC-R (as applied to the normative and Sisco [males] samples). Because the $\Delta\chi^2$ served as the primary statistic by which measurement invariance was evaluated, I deemed these models as not invariant across normal-hearing and D/HOH samples.

When the CHC model was applied to the normal-hearing ITPA data (yielded from those in the normative group aged six to nine) and the D/HOH ITPA data (yielded from the Espeseth sample), the factor loadings for the Visual Decoding subtest varied significantly between the two groups. When the CHC model was applied to the normal-hearing WISC-R data (yielded from the entire normative sample) and both of the following D/HOH WISC-R samples (Phelps and Ensor [females] and Sisco [males]), the factor loadings for Block Design varied significantly between the normal-hearing group and each D/HOH group. Additionally, the factor loadings for Picture Arrangement varied significantly between normal-hearing and D/HOH individuals when the CHC model was applied to the WISC-R normative group and the Sisco (males) WISC-R sample. For the mixed-age samples, I deemed the remaining CHC models invariant across the normal-hearing and D/HOH groups according to nonsignificant $\Delta\chi^2$s: (1) K-ABC, (2) UNIT (as applied to the normative and

Krivitski, McIntosh, Rothlisberg, and Finch samples), (3) WISC, (4) WISC-R (as applied to the normative and Brooks and Riggs samples), and (5) WISC-R (as applied to the normative and Phelps and Ensor [males] samples).

A summary of all test results is provided in Table 14. Overall, the same number of comparisons between normal-hearing and D/HOH children aged 10 and younger show differences in factor structure as show similarities; thus, results are inconclusive in finding factor structure differences at younger ages. However, all of the comparisons between normal-hearing and D/HOH children/adults aged 11 and older show similarities in the factor structures of these groups suggesting factor structure differences do not exist at older ages. Additionally, a weak level of invariance was obtained across all older samples of normal-hearing and D/HOH individuals indicating that test scores have similar meanings for these groups.

Table 13

Results from the Tests of Invariance for D/HOH Children/Adults of Mixed Ages

Test Battery	Normative Sample (Sample Portion Used)	D/HOH Study	Model	$X^2(df)$	*p*	CFI	NNFI	RMSEA	$\Delta\chi^2(\Delta df)$	*p*
ITPA	McCarthy & Kirk	Espeseth (1965)	Configural	.56(2)	.76	1.00	1.09	.00	-----	-----
	(1963)		Metric	6.81(3)	.08	.95	.82	.08	6.25(1)	<.05
	(6 – 9 years)									
K-ABC	Kaufman &	Ulissi, Brice, &	Configural	158.29(36)	.00	.98	.97	.07	-----	-----
	Kaufman (1983)	Gibbins (1989)	Metric	166.61(43)	.00	.98	.97	.06	8.32(7)	NS
	(5 – 12.5 years)									
UNIT	Bracken &	Krivitski,	Configural	107.11(16)	.00	.98	.96	.07	-----	------
	McCallum (1998)	McIntosh,	Metric	112.76(21)	.00	.98	.97	.06	5.65(5)	NS
	(Entire sample)	Rothlisberg, &								
		Finch (2004)								

Table 13 (continued)

WISC	Wechsler (1949)	Brooks & Riggs	Configural	11.54(4)	.02	.99	.96	.08	-----	-----
	(Entire sample)	(1980)	Metric	12.62(6)	.05	.99	.98	.06	1.08(2)	NS
WISC-R	Wechsler (1974)	Brooks & Riggs	Configural	14.94(4)	.00	.99	.98	.05	-----	-----
	(Entire sample)	(1980)	Metric	15.30(6)	.02	1.03	.99	.04	.36(2)	NS
WISC-R	Wechsler (1974)	Phelps & Ensor	Configural	39.05(10)	.00	.99	.98	.05	-----	-----
	(Entire sample)	(1987), Females	Metric	50.74(13)	.00	.99	.99	.05	11.69(3)	<.01
WISC-R	Wechsler (1974)	Phelps & Ensor	Configural	42.59(10)	.00	.99	.98	.05	-----	------
	(Entire sample)	(1987), Males	Metric	44.49(13)	.00	.99	.99	.05	1.90(3)	NS
WISC-R	Wechsler (1974)	Sisco (1982),	Configural	62.55(10)	.00	.99	.98	.06	-----	-----
	(Entire sample)	Males	Metric	72.43(13)	.00	.99	.98	.06	9.88(3)	<.05

Table 14

Results of the Invariance Tests for the Factor Structures of Normal-Hearing and D/HOH Samples

Age (Sample Portion Used)	Test Battery	Normative Sample	D/HOH Sample	Results
Young (3 – 10 years)	HNTLA	Hiskey (1966)	Bolton (1978)	[Not Tested]
Young (5.5 – 9 years)	ITPA	McCarthy & Kirk (1963)	Hanson, Hancock, & Kopra (1969)	Not invariant
Young (6 – 9 years)	ITPA	McCarthy & Kirk (1963)	Johnson (1966)	Not invariant
Young (7.5 – 10.5 years)	WISC	Wechsler (1949)	Murphy (1957)	Invariant
Young (5 – 6.5 years)	WPPSI	Wechsler (1967)	Brinich (1981)	Invariant
Old (11 – 17 years)	HNTLA	Hiskey (1966)	Bolton (1978)	Invariant

Table 14 (continued)

Old (16 – 24 years)	WAIS-R	Wechsler (1981)	Ensor & Phelps, Females (1989)	Invariant
Old (16 – 24 years)	WAIS-R	Wechsler (1981)	Ensor & Phelps, Males (1989)	Invariant
Mixed (6 – 9 years)	ITPA	McCarthy & Kirk (1963)	Espeseth (1965)	Not invariant
Mixed (5 – 12.5 years)	K-ABC	Kaufman & Kaufman (1983)	Ulissi, Brice, & Gibbins (1989)	Invariant
Mixed (Entire sample)	UNIT	Bracken & McCallum (1998)	Hughes, Sapp, & Kohler (2006)	Not invariant
Mixed (Entire sample)	UNIT	Bracken & McCallum (1998)	Krivitski, McIntosh, Rothlisberg, & Finch (2004)	Invariant

Table 14 (continued)

Mixed (Entire sample)	WISC	Wechsler (1949)	Brooks & Riggs (1980)	Invariant
Mixed (Entire sample)	WISC-III	Wechsler (1991)	Hughes, Sapp, & Kohler (2006)	Not invariant
Mixed (Entire sample)	WISC-R	Wechsler (1974)	Brooks & Riggs (1980)	Invariant
Mixed (7.5 – 13.5 years)	WISC-R	Wechsler (1974)	Hirshoren, Hurley, & Hunt (1977)	Not invariant
Mixed (Entire sample)	WISC-R	Wechsler (1974)	Phelps & Ensor (1987), Females	Not invariant
Mixed (Entire sample)	WISC-R	Wechsler (1974)	Phelps & Ensor (1987), Males	Invariant
Mixed (Entire sample)	WISC-R	Wechsler (1974)	Sisco (1982), Females	Not invariant

Table 14 (continued)

Mixed (Entire sample)	WISC-R	Wechsler (1974)	Sisco (1982), Males	Not Invariant
Mixed (7.5 – 16.5 years)	WISC-R	Wechsler (1974)	Welch (1992)	Not invariant

Chapter 4

Discussion

The major purpose of the present study was to better understand whether the cognitive abilities of D/HOH and normal-hearing people are organized similarly throughout the lifespan, from preschool through adulthood. This study built on previous factor analytic research in two distinct ways. First, I employed updated methodology, which promises less ambiguous and more direct, objective, and complete techniques to resolve competing claims. More specifically, I used CFA with hierarchical models to examine whether cognitive abilities were similar for both groups (i.e., factor models adequately fit the data from normal-hearing and D/HOH people) and subsequent MCFA to assess for measurement invariance (i.e., to determine whether intelligence tests are measuring the same construct in both groups). Previous research incorporates EFA methodology, which does not align with current intelligence theory because it is not hierarchical in nature and does not allow for a specified model to be fit to a given dataset. That is, it is not a theory-driven approach. Second, I applied this methodology to relatively homogeneous groups of D/HOH and normal-hearing individuals because evidence suggests that younger D/HOH children may show greater differences in cognitive structure from their normal-hearing peers than do older D/HOH children and adults.

Hypothesis One

I generated three hypotheses to clarify questions and conflicts raised by previous research. In Hypothesis One, I postulated that the factor structures that describe the normative sample data for younger (preschool through age 10) normal-hearing children *would not* fit the sample data for D/HOH children of the same ages according to the criterion values of various goodness-of-fit statistics (i.e., CFI and NNFI $\geq .90$; RMSEA $\leq .08$; SRMR $\leq .10$). I argued that Hypothesis One would be supported if the majority of comparisons identified different factor structures for samples of younger D/HOH children relative to samples of same-age normal-hearing peers. Given that the CHC models that adequately fit the normal-hearing data proved ill-fitting when applied to the D/HOH sample data for only two of the four younger samples, Hypothesis One was not supported. That is, the same number of comparisons showed differences as showed similarities so data were not consistent with the hypothesis. Results are therefore inconclusive with regard to whether factor structure differences exist between younger D/HOH and normal-hearing children (i.e., they organize and approach nonverbal cognitive tasks differently) and thus the degree to which there are developmental discontinuities with the application of CHC theory.

Although differences in test reliability characteristics between groups could contribute to differences in cognitive factor structure, test reliability tends to be similar or higher with those who are D/HOH as compared to those who are normal hearing (Kostrubala, 1998; Krouse, 2008). Examination of the factor loadings for the CHC models for each group provides further evidence that reliability differences do not account for the differences between D/HOH and normal-hearing children, as factor loadings were generally similar in magnitude (but not necessarily in patterns) for each group. Factor loadings are listed in Appendix L.

It is important to note that it was the CHC model applied to the ITPA that lacked factorial similarity across the younger D/HOH and normal-hearing groups. To date, no research has been conducted that examines the ITPA in relation to the intellectual structure of either normal-hearing or D/HOH individuals; however, the ITPA factor model provided an excellent fit to the normal-hearing data from multiple study samples. The poor fit of the ITPA model with the younger D/HOH children indicates that this test is measuring different attributes in this group as compared to those who are normal hearing and, therefore, that test scores do not have similar meanings for each group. Although this may be evidence of different intellectual structures between the groups, it is also possible that, because this test intends to measure the acquisition and development of language, construct-irrelevant variance was introduced thereby threatening the construct validity of the instrument. Although the auditory-vocal subtests were intentionally eliminated, it is possible that some of the administered items required linguistic mediation and, as noted earlier, D/HOH children score substantially lower than their normal-hearing peers on verbal intelligence tests. Ideally, I would further investigate this issue by excluding those subtests most likely to act as verbal/vocal confounders (i.e., those that map onto *Gc*) to determine whether the remaining model provided an adequate fit to the data. However, the remaining model would include only two broad abilities with one subtest mapping onto each and therefore could not be tested because all paths would have to be fixed for the model to be identified. As a result, it is not possible to conclude whether the difference in the fit of the ITPA model between normal-hearing and D/HOH children is due to differences in factor structures or construct-irrelevant variance.

Hypotheses Two and Three

As noted, previous research suggests that the differences that exist in factor structure between D/HOH and normal-hearing children (i.e., those under age 11) tend to diminish in samples of older D/HOH children and adults. Therefore, in Hypothesis Two, I postulated that the factor structures that describe the normative sample data for older (age 11 through adulthood) normal-hearing children *would* fit the sample data for D/HOH children of the same ages. I argued that Hypothesis Two would be supported if 50% or more of the comparisons suggested similarities in factor structures for samples of older D/HOH children/adults relative to samples of same-age normal-hearing peers. This hypothesis was supported according to the same fit statistics as were used to evaluate Hypothesis One (i.e., CFI and NNFI $\geq .90$; RMSEA $\leq .08$; SRMR $\leq .10$). More specifically, 100% (three out of three) of the comparisons between the older groups indicated similar factor structures. Thus, these results support previous research and suggest that older D/HOH individuals and their same-age normal-hearing peers organize and approach cognitive tasks similarly. That is, they do not differ in their flexible, nonverbal problem-solving abilities.

Additionally, I proposed that a weak level of invariance would be obtained across the older D/HOH and normal-hearing individuals. This hypothesis (Hypothesis Three) was also supported, as 100% of the comparisons yielded nonsignificant chi-square difference tests (at the $p < .05$ level). The presence of invariance indicates that test scores have similar meanings for these groups, thereby supporting the validity of these intelligence tests and the viability of CHC theory with older D/HOH children and adults. However, it is important to note that, because it is best practice to measure D/HOH people's intelligence with nonverbal instruments, researchers tend to systematically delete those tasks that tap factors dependent upon language, such as crystallized intelligence. As a result, these abilities are not included in many of the CHC factor models used here making these models much more limited than CHC models used in other research.

Supplemental Analyses

To increase the total number of datasets included in the current study, I also collected studies with heterogeneous (i.e., mixed-age) samples. Previous factor-analytic research examining the intellectual structure within the D/HOH population using nonverbal tests and heterogeneous samples has found normal-hearing and D/HOH people to have similar factor structures. However, the supplemental analyses conducted as part of this study with mixed-age samples revealed a lack of invariance between normal-hearing and D/HOH individuals in the majority of comparisons. This finding bolsters support for the notion that there exists an interaction between age, hearing status, and factor structure on cognitive test batteries.

A more thorough review of invariance patterns in these mixed-age samples shows Block Design in particular varies between the two groups, which suggests that scores on this subtest may have different meanings for each group. The only other study that has evaluated the invariance of a Wechsler test across normal-hearing and D/HOH samples also found Block Design (among other subtests) to lack invariance across groups with a mixed-age population (Maller & Ferron, 1997). Because D/HOH individuals tend to score higher on performance tasks (such as Block Design) as compared to nonverbal cognitive tasks that are motor-free, it is possible that response characteristics accounted for the differences in factor loadings between the two groups (Braden, Kostrubala, & Reed, 1994). More specifically, it is likely that the Block Design task allowed D/HOH children to invoke psychomotor speed strategies that aided in their solution of the subtest items. Consequently, the CHC abilities D/HOH individuals used to solve this task are likely to be different than those used by people who have normal hearing.

Implications for Practice with D/HOH Individuals

Two practical recommendations follow from the results of this research study. The first is that the ITPA should not be used to assess cognitive abilities with those who are D/HOH. Although the ITPA subtests could be mapped reliably onto CHC constructs, the instrument lacked invariance across normal-hearing and D/HOH groups with both younger and mixed samples suggesting that scores have different meanings for each group. The fact that the ITPA is an outdated test, and is no longer in wide use even within normal-hearing populations, makes this recommendation somewhat obsolete, but I make the recommendation nonetheless in the event practitioners still use the instrument. The second recommendation is that language-reduced multi-test intelligence batteries be administered to D/HOH individuals since these measures produced similar cognitive factor structures for those who were D/HOH and their normal-hearing peers. Therefore, there does not exist evidence that the use of these batteries with those who are D/HOH is not a valid practice.

Theoretical Implications

The cumulative deficit theory posits that environmental factors have a cumulative effect on cognitive abilities, meaning individuals in either enriched or deprived environments have similar IQs early in life, but over time, the effects of the environment cause these IQ distributions to diverge so that those with less stimulating environments fall further behind peers in enriched environments (Braden, 1994; Jensen, 1974; 1977). As noted earlier, D/HOH individuals experience language deprivation; thus, it is likely that the cumulative effect of this deprivation would result in D/HOH children diverging from their normal-hearing peers as they age. Previous research (Allen, 1986; Karchmer & Mitchell, 2003; Vernon & Koh, 1970) suggests this pattern occurs with regard to academic achievement, but there is no evidence that D/HOH children experience a cumulative deficit in nonverbal IQ (Braden, 1994). Similarly, the results of the current study suggest that D/HOH and normal-hearing individuals' cognitive factor structures do not diverge over time, which is unexpected if environmental differences would be expected to exert a cumulative effect on the organization of cognitive abilities. To parallel these findings using a different area of research, studies examining the impact of heredity on IQ have consistently found that not only is the correlation between adopted children's IQ and their biological parents' IQ greater than the correlation between the children's IQ and their adoptive parents' IQ, but that the association between the children's IQ and their biological parents' IQ actually gets stronger as the children age (Phillips & Fulker, 1989; Plomin et al., 1997). Again, this is not consistent with the notion of a cumulative deficit, or cumulative effects of environment on intellectual development. Therefore, future research should seek to explore this phenomenon more closely to help elucidate the influence of environmental effects on cognitive abilities.

Limitations and Future Directions

Although the results of the current study suggest insights into the relationship between age, hearing status, and cognitive factor structure, there exist a number of issues important to consider when generalizing these results, and when planning future research.

First, despite a thorough search of the extant literature, comparisons used to examine the study hypotheses were based on a limited number of studies. Therefore, results should be considered tentative, as they are based on a small number of available data points, which limits the confidence with which results can be generalized. Thus, future research should produce more studies, and more homogeneous samples (i.e., with those either 10 and younger or 11 and older) to test whether these results generalize beyond a limited number of studies.

Second, the studies included in this investigation likely varied according to sample demographic characteristics. Although age was the primary focus of this research, variables such as gender, race, age at onset of hearing loss, severity of hearing loss, type of school the participants attended, additional disabilities, parental hearing status, and cognitive test administration procedures may have affected the obtained results. However, it is difficult to assess the impact of these variables, even qualitatively, because different authors provide different demographic information. Braden (1994) notes that many of these factors (i.e., gender, age at onset of hearing loss, severity of hearing loss) have no appreciable effect on mean IQ. (Racial differences are similar to those found in the normal-hearing population.) However, type of school, presence of additional disabilities, parent hearing status, and test administration procedures are associated with mean IQ. To date, no research has examined the relationship between these variables and cognitive factor structures that underlay mean scores. Thus, future research may wish to isolate these variables to better understand which, if any, are related to the structure of cognitive abilities in D/HOH people, and its measurement.

Third, findings for this study are based largely on goodness-of-fit statistics. However, the following factors contribute to the complexity of using goodness-of-fit statistics to interpret results: (1) there are many model fit statistics described in the literature, (2) it is recommended that multiple measures of fit be examined in conjunction with one another since each focuses on a different aspect of fit, (3) there exist a wide range of cutoff values for each index, and (4) strict adherence to specific cutoff values is not suggested. As a result, I may have garnered different findings had I used alternate goodness-of-fit statistics or more conservative cutoff values. To prevent selective reporting and assist in comparing results across studies, future research should seek to streamline the guidelines regarding the use of fit indices.

Conclusions

The results yielded from this study do not support the notion that differences exist in the developmental progression of nonverbal cognitive abilities for normal-hearing and D/HOH individuals. More specifically, the evidence regarding younger, D/HOH children (i.e., those age 10 and younger) does not exhibit reliable differences in cognitive structure relative to their normal hearing peers; however, the evidence is split down the middle, so it is not wise to draw strong conclusions regarding differences or similarities. The evidence regarding older D/HOH children/adults (i.e., those age 11 and older) suggests they clearly do not have a different factor structure from their normal-hearing peers and therefore suggests that they organize and approach nonverbal tasks similarly (i.e., they do not differ in their flexible, nonverbal problem-solving abilities). These findings support the validity of the use of nonverbal cognitive test batteries with D/HOH children/adults and contribute to the viability of CHC theory with this population. This study serves as the first to evaluate the relationship between age, hearing status, and cognitive factor structure using CFA (with hierarchical models) and invariance testing (via MCFA) with homogeneous age groups, and as such, serves to clarify the mixed results obtained in previous factor-analytic research.

References

Akamatsu, C. T., Mayer, C., & Hardy-Braz, S. (2008). Why considerations of verbal aptitude are important in educating deaf and hard-of-hearing students. In M. Marschark & P. C. Hauser (Eds.), *Deaf cognition: Foundations and outcomes* (pp. 131-169). New York: Oxford University Press.

Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 185-202). New York: Guilford Press.

Allen, T. E. (1986). Patterns of academic achievement among hearing-impaired students: 1974 and 1983. In A. N. Schildroth & M. A. Karchmer (Eds.), *Deaf children in America* (pp. 161-206). San Diego, CA: College Hill Press.

Bickley, P.G., Keith, T. Z., & Wolfe, L. M. (1995). The three-stratum theory of cognitive abilities: Test of the structure of intelligence across the life span. *Intelligence, 20,* 309-328.

Bolton, B. (1978). Differential ability structure in deaf and hearing children. *Applied Psychological Measurement, 2,* 147-149.

Bracken, B. A., & McCallum, R. S. (1998). *Universal Nonverbal Intelligence Test.* Itasca, IL: Riverside Publishing.

Braden, J. P. (1984). The factorial similarity of the WISC-R performance scale in deaf and hearing samples. *Journal of Personality and Individual Differences, 5,* 403 – 409.

Braden, J. P. (1985). The structure of nonverbal intelligence and deaf and hearing subjects. *American Annals of the Deaf, 130*, 496-501.

Braden, J. P. (1992). Intellectual assessment of deaf and hard of hearing people: A quantitative and qualitative research synthesis. *School Psychology Review, 21*, 82-94.

Braden, J. P. (1994). *Deafness, deprivation, and IQ.* New York: Plenum Press.

Braden, J. P. (2001). The clinical assessment of deaf people's cognitive abilities. In M. D. Clark, M. Marschark, & M. Karcher (Eds.), *Context, cognition, and deafness* (pp. 14-37). Washington, DC: Gallaudet University Press.

Braden, J. P. (2005). Hard-of-hearing and deaf clients: Using the WISC-IV with clients who are hard-of-hearing or deaf. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 351-376). San Diego, CA: Elsevier Academic Press.

Braden, J. P., & Athanasiou, M. S. (2005). A comparative review of nonverbal measures of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 557-577). New York: Guilford Press.

Braden, J. P., Kostrubala, C. E., & Reed, J. (1994). Why do deaf children score differently on performance vs. motor-reduced nonverbal intelligence tests? *Journal of Psychoeducational Assessment, 12*, 357-363.

Brinich, P. M. (1981). Relationship between intellectual functioning and communicative competence in deaf children. *Journal of Communication Disorders, 14*, 429-434.

Brooks, C. R., & Riggs, S. T. (1980). WISC-R, WISC, and reading achievement relationships among hearing-impaired children attending public schools. *Volta Review, 82*, 96-102.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Beverly Hills, CA: Sage.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment.* Beverly Hills, CA: Sage.

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

Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 69-76). New York: Guilford Press.

Domino, G., & Domino, M. L. (2006). *Psychological testing: An introduction* (2nd ed.). New York: Cambridge University Press.

Doppelt, J. E., & Kaufman, A. S. (1977). Estimation of the differences between WISC-R and WISC IQs. *Educational and Psychological Measurement, 37*, 417-424.

Ensor, A., & Phelps, L. (1989). Gender differences on the WAIS-R Performance Scale with young deaf adults. *Journal of the American Deafness and Rehabilitation Association, 22*, 48-52.

Espeseth, V. K. (1965). *An investigation of visual sequential memory in the deaf child.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Database. (AAT 6514875).

Farrant, R. H. (1964). The intellective abilities of deaf and hearing children compared by factor analyses. *American Annals of the Deaf*, *109*, 306-325.

Flanagan, D. P., & Mascolo, J. T. (2005). Psychoeducational assessment and learning disability diagnosis. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 521-544). New York: Guilford Press.

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Dynda, A. M. (2008). Best practices in cognitive assessment. In A. Thomas & J. Grimes (Eds.), *Best Practices in school psychology V: Vol. 2* (pp. 633-660). Bethesda, MD: National Association of School Psychologists.

Gallaudet Research Institute. (2006). *Regional and national summary report of data from the 2006-2007 annual survey of deaf and hard of hearing children and youth.* Washington, DC: GRI, Gallaudet University.

Gustafsson, J., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research, 28*, 407-434.

Hanson, G. H., Hancock, B.B., & Kopra, L. L. (1969). *Relationships among audiological status, linguistic skills, visual-motor perception, and academic achievement of deaf children* (U.S. Department of Health, Education, and Welfare Grant No. OEG 4-7-002353-2051). Austin: University of Texas.

Hine, W. D. (1970). The abilities of partially hearing children. *British Journal of Educational Psychology, 40*, 171-178.

Hirshoren, A., Hurley, O. L., & Hunt, J. T. (1977). The WISC-R and the Hiskey-Nebraska Test with deaf children. *American Annals of the Deaf, 122*, 392-394.

Hiskey, M. S. (1966). *Hiskey-Nebraska Test of Learning Aptitude.* Lincoln, NE: College View Printers.

Horn, J. L. (1991). Discussion of the issues of factorial invariance. In L. M. Collins & J. L. Horn (Eds.), *Best Methods for the analysis of change* (pp. 114-125). Washington, DC: American Psychological Association.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*, 117-144.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. A. Harrison (Eds.), *Contemporary intellectual assessment: Theories and issues* (2nd ed., pp. 53-91). New York: Guilford Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hughes, D., Sapp, G. L., & Kohler, M. P. (2006). *Issues in the intellectual assessment of hearing impaired children.* Birmingham: University of Alabama. (ERIC Document Reproduction Service No. ED493819)

Jensen, A. R. (1974). Cumulative deficit: A testable hypothesis. *Developmental Psychology, 10*, 996-1019.

Jensen, A. R. (1977). Cumulative deficit in IQ of Blacks in the rural South. *Developmental Psychology, 13*, 184-191.

Jensen, A. R. (1980). *Bias in mental testing.* New York: The Free Press.

Johnson, G. B. (1966). *A comparison of scores earned by certain groups of residential school deaf children and the standardization sample of the Illinois Test of Psycholinguistic Abilities on four selected subtests.* (Master's thesis). Retrieved from ProQuest Dissertations and Theses Database. (AAT 6706206).

Jö reskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.). *Testing structural equation models* (pp. 294-316). Newbury Park, CA: Sage.

Kamphaus, R. W. (2005). *Clinical assessment of child and adolescent intelligence* (2nd ed.). New York: Springer Science and Business Media.

Karchmer, M. A., & Mitchell, R. E. (2003). Demographic and achievement characteristics of deaf and hard-of-hearing students. In M. Marschark & P. E. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (pp. 21-37). Oxford: Oxford University Press.

Kaufman, A.S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children, Second Edition.* Circle Pines, MN: American Guidance Service.

Keith, T. Z. (1999). Effects of general and specific abilities on student achievement: Similarities and differences across ethnic groups. *School Psychology Quarterly, 14,* 239-262.

Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 581-614). New York: Guilford Press.

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher-order multi-sample confirmatory factor analysis of the Wechsler Intelligence Scale for Children– Fourth Edition: What does it measure? *School Psychology Review, 35*, 108-127.

Keith, T. Z., Fugate, M. H., DeGraff, M., Diamond, C. M., Shadrach, E. A., & Stevens, M. L. (1995). Using multi-sample confirmatory factor analysis to test for construct bias: An example using the K-ABC. *Journal of Psychoeducational Assessment*, *13*, 347-364.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Kostrubala, C. (1998). *Translation of the WAIS-II from English to ASL.* Unpublished master's thesis, University of Wisconsin, Madison.

Krivitski, E. C., McIntosh, D. E., Rothlisberg, B., & Finch, H. (2004). Profile analysis of deaf children using the Universal Nonverbal Intelligence Test. *Journal of Psychoeducational Assessment*, *22*, 338-350.

Krouse, H. E. (2008). *The reliability and validity of the WISC-IV with deaf and hard-of-hearing children.* Unpublished master's thesis, North Carolina State University, Raleigh. Retrieved from http://www.lib.ncsu.edu/resolver/1840.16/2149

Maller, S. J. (2003a). Intellectual assessment of deaf people: A critical review of core concepts and issues. In M. Marschark & P. Spencer (Eds.), *Handbook of deaf studies, language, and education* (pp. 451-463). New York: Oxford University Press.

Maller, S. J. (2003b). Best practices in detecting bias in nonverbal tests. In R. S. McCallum(Ed.), *Handbook of nonverbal assessment* (pp. 23-48). Needham Heights, MA: Allyn & Bacon.

Maller, S. J., & Ferron, J. (1997). WISC-III factor invariance across deaf and standardization samples. *Educational and Psychological Measurement, 57*, 987-994.

Maller, S. J., & French, B. F. (2004). Universal Nonverbal Intelligence Test factor invariance across deaf and standardization samples. *Educational and Psychological Measurement, 64*, 647-660.

Maller, S. J., & Immekus, J. C. (2005). The psychometric properties of intelligence tests when used with deaf and hard of hearing individuals: Practices and recommendations. In J. L. Mounty & D. S. Martin (Eds.), *Assessing deaf adults: Critical issues in testing and evaluation* (pp. 75-89). Washington, DC: Gallaudet University Press.

Marschark, M. (1997). *Raising and educating a deaf child.* Oxford: Oxford University Press.

Marschark, M., Lang, H. G., & Albertini, J. A. (2002). *Educating deaf students: From research to practice.* New York: Oxford University Press.

McCallum, R. S. (2003). The Universal Nonverbal Intelligence Test. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment* (pp. 87-111). Needham Heights, MA: Allyn & Bacon.

McCarthy, J. J., & Kirk, S. A. (1961). The Illinois Test of Psycholinguistic Abilities: An approach to differential diagnosis. *American Journal of Mental Deficiency, 66*, 399-412.

McCarthy, J. J., & Kirk, S. A. (1963). *The construction, standardization, and statistical characteristics of the Illinois Test of Psycholinguistic Abilities.* Madison, WI: Photo Press.

McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 136-181). New York: Guilford Press.

Meredith, W. (1993). Measurement invariance, factor analysis, and factor invariance. *Psychometrika*, *58*, 525-543.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.

Moore, K. L., & Dalley, A. F. (1999). *Clinically oriented anatomy* (4th ed.). Baltimore, MD: Lippincott Williams & Wilkins.

Murphy, L. J. (1957). Tests of abilities and attainments: Pupils in schools for the deaf aged six to ten. In A. W. Ewing (Ed.), *Educational guidance and the deaf child* (pp. 213-251). Manchester, UK: Manchester University Press.

Northern, J. L., & Downs, M. P. (2002). *Hearing in children* (5th ed.). Baltimore, MD: Lippincott Williams & Wilkins.

Ong, A. D., & van Dulmen, M. (Eds.). (2006). *The Oxford handbook of methods in positive psychology*. New York: Oxford University Press.

Phelps, L.A., & Ensor, A. (1987). The comparison of performance by sex of deaf children on the WISC-R. *Psychology in the Schools, 24*, 209-214.

Phelps, L. A., McGrew, K. S., Knopik, S. N., & Ford, L. (2005). The general (*g*), broad, and narrow CHC stratum characteristics of the WJ-III and WISC-III tests: A confirmatory cross-battery investigation. *School Psychology Quarterly, 20*, 66-88.

Phillips, K., & Fulker, D. W. (1989). Quantitative genetic analysis of longitudinal trends in adoption designs with application to IQ in the Colorado Adoption Project. *Behavior Genetics, 19*, 621-658.

Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 years: A parent-offspring adoption study. *Psychological Science, 8*, 442-447.

Prifitera, A, Weiss, L. G., & Saklofske, D. H. (1998). The WISC-III in context. In A. Prifitera & D. Saklofske (Eds.), *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 1-38). San Diego, CA: Elsevier Academic Press.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales, Fifth Edition*. Itasca, IL: Riverside Publishing.

Sattler, J., & Hoge, R. D. (2006). *Assessment of children: Behavioral, social, and clinical foundations* (5th ed.). San Diego, CA: Sattler Publishing Company.

Schaie, K. W., Maitland, S. B., Willis, S. L., & Intrieri, R. L. (1998). Longitudinal invariance of adult psychometric ability factor structures across seven years. *Psychology and Aging, 13,* 8-20.

Sisco, F. H. (1982). Sex differences in the performance of deaf children on the WISC-R Performance Scale. *Dissertation Abstracts International, 43*, 6A. (University of Microfilms No. DA8226432).

Sullivan, P. M., & Montoya, L. A. (1997). Factor analysis of the WISC-III with deaf and hard-of-hearing children. *Psychological Assessment, 9*, 317-321.

Sullivan, P. M., & Schulte, L. E. (1992). Factor analysis of the WISC-R with deaf and hard-of-hearing children. *Psychological Assessment, 4,* 537-540.

Ulissi, S. M., Brice, P. J., & Gibbins, S. (1989). Use of the Kaufman-Assessment Battery for Children with the hearing impaired. *American Annals of the Deaf, 134*, 283-287.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4-70.

Vernon, M. (1967/2005). Fifty years of research on the intelligence of deaf and hard-of-hearing children: A review of literature and discussion of implications. *Journal of Deaf Studies and Deaf Education, 10*, 225-231.

Vernon, M., & Koh, S. D. (1970). Early manual communication and deaf children's achievement. *American Annals of the Deaf, 115*, 527-536.

Wasserman, J. D., & Tulsky, D. S. (2005). A history of intelligence assessment. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 3-22). New York: Guilford Press.

Wechsler, D. (1939). *Measurement of adult intelligence.* Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York: The Psychological Corporation.

Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence.* New York: The Psychological Corporation.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised.* New York: The Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children: Third Edition.* New York: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children: Fourth Edition*. San Antonio, TX: The Psychological Corporation.

Welch, J. A. (1992). *Predicting academic achievement of hearing impaired students using the Wechsler Performance Scale and the Planning, Attention, Simultaneous, and Successive (PASS) Model of Cognitive Processing.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Database. (AAT 9227400).

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Zhu, J., & Weiss, L. (2005). The Wechsler Scales. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 297-324). New York: Guilford Press.

Zweibel, A., & Mertens, D. M. (1985). A comparison of intellectual structure in deaf and hearing children. *American Annals of the Deaf, 130*, 27-31.

Appendices

Appendix A

Coding Worksheet for the Hiskey-Nebraska Test of Learning Aptitude (HNTLA)

Name: _______________

Date: _______________

Beside each subtest description, please note the broad ability factor that you believe the subtest most closely measures. For your reference, the 8 broad ability factors from which you can choose are described on the second page.

Subtest	Description	Broad Ability Factor and any Notes
Bead Patterns (Stringing Beads)	The examinee is asked to string beads as rapidly as possible	
Bead Patterns (Copying Bead Patterns)	The examinee is asked to reproduce a string of beads, matching the round, square, and rectangular beads	
Bead Patterns (Memory for Bead Patterns)	The examinee is asked to reproduce a string of beads from memory	
Memory for Color	The examinee selects, from memory, one or more color chips to match those presented by the examiner	
Picture Identification	The examinee selects one of several pictures that matches the target picture	
Picture Association	The examinee selects the picture that “goes with” the	

	other pictures	
Paper Folding	The examinee is to replicate, from memory, paper folding sequences the examiner makes	
Visual Attention Span	The examinee reproduces a series of pictures from memory	
Block Patterns	The examinee constructs block patterns to match the target picture	
Completion of Drawings	The examinee draws in the missing parts of various pictures of shapes/objects	
Memory for Digits	The examinee reproduces a series of visually presented numerals from memory	
Puzzle Blocks	The examinee puts together puzzle pieces to create a solid block	
Picture Analogies	The examinee completes visually presented analogies by selecting from five pictures	
Spatial Reasoning	The examinee selects which of four groups of geometric figures could be assembled to form the target design	

Note. The information in columns 1 and 2 is from *Psychological Testing: An Introduction, 2nd Edition* (p. 243), by G. Domino and M. L. Domino, 2006, New York, NY: Cambridge University Press; "A new performance test for young deaf children," by M.S.Hiskey, 1941, *Educational and Psychological Measurement, 1,* pp. 221-22.

Appendix B

Coding Worksheet for the Illinois Test of Psycholinguistic Abilities (ITPA)

Name: _______________

Date: _______________

Beside each subtest description, please note the broad ability factor that you believe the subtest most closely measures. For your reference, the 8 broad ability factors from which you can choose are described on the second page.

Subtest	Description	Broad Ability Factor and any Notes
Visual Decoding	The examinee is to select, from memory, which of four pictures is semantically (not physically) identical to the target picture (e.g., a silver knife and a jack knife)	
Visual-Motor Association	The examinee completes visually presented analogies by selecting which of four pictures "goes with" the stimulus picture	
Motor Encoding	The examinee is shown an object or a picture and is asked to show "what we do with this"	
Visual-Motor Sequential Ability	The examinee duplicates, from memory, a series of pictures of geometrical designs	

Note. The information in columns 1 and 2 is from "The Illinois Test of Psycholinguistic Abilities: An approach to differential diagnosis," by J.J. McCarthy and S.A. Kirk, 1961, *American Journal of Mental Deficiency, 66*, pp 404-405.

Appendix C

Coding Worksheet for the Universal Nonverbal Intelligence Test (UNIT)

Name: _______________

Date: _______________

Beside each subtest description, please note the broad ability factor that you believe the subtest most closely measures. For your reference, the 8 broad ability factors from which you can choose are described on the second page.

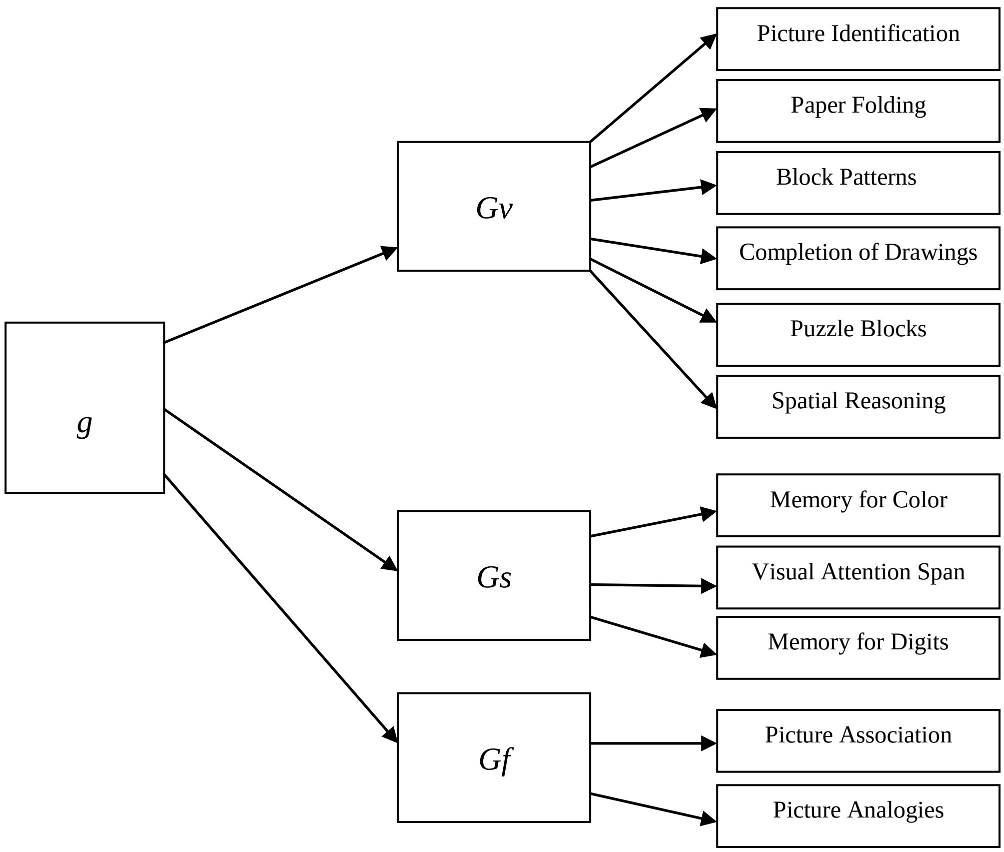
Subtest	Description	Broad Ability Factor and any Notes
Symbolic Memory	The examinee reproduces a series of visually presented symbols (e.g., green boy, black woman) from memory	
Spatial Memory	The examinee reproduces the placement of black and/or green chips on a grid from memory	
Object Memory	The examinee is shown a visual array of common objects (e.g., shoe, tree) for five seconds and is then asked to select which objects were shown from a larger array of objects	
Cube Design	The examinee uses between one and nine green and white blocks to construct three-dimensional designs	
Analogic Reasoning	The examinee completes a	

	matrix analogies task using common objects (e.g., hand : glove : foot : __?) and novel geometric figures	
Mazes	The examinee traces a path from the center of each maze to an exit	

Note. The information in columns 1 and 2 is from *Handbook of Nonverbal Assessment* (p.89), by R.S. McCallum, 2003, Needham Heights, MA: Allyn & Bacon.

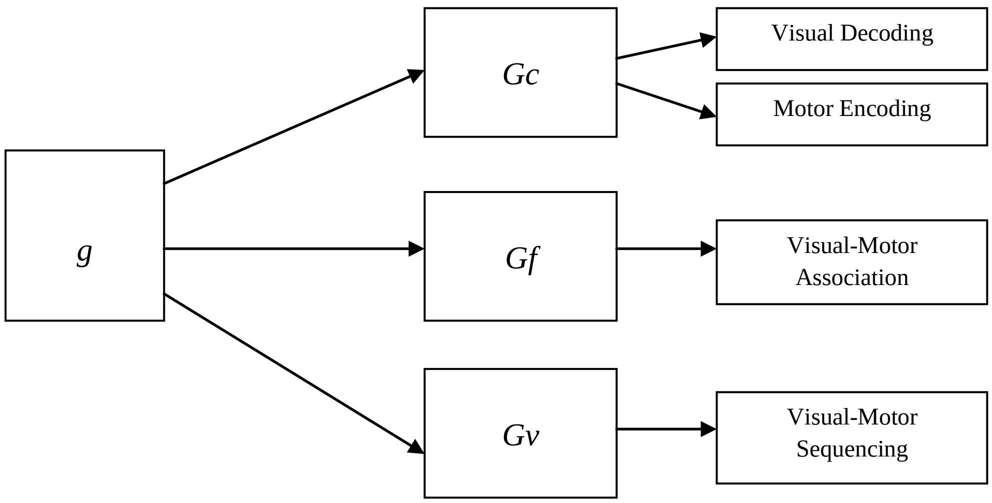
Appendix D

Factor Model for Hiskey-Nebraska Test of Learning Aptitude (HNTLA)



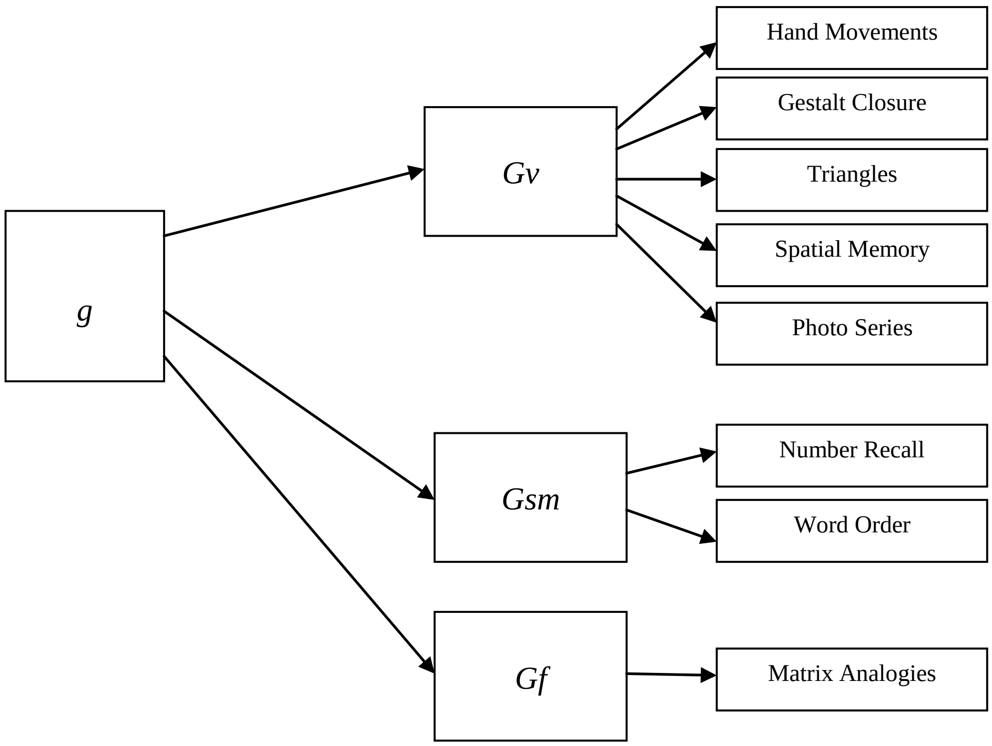
Appendix E

Factor Model for the Illinois Test of Psycholinguistic Abilities (ITPA)



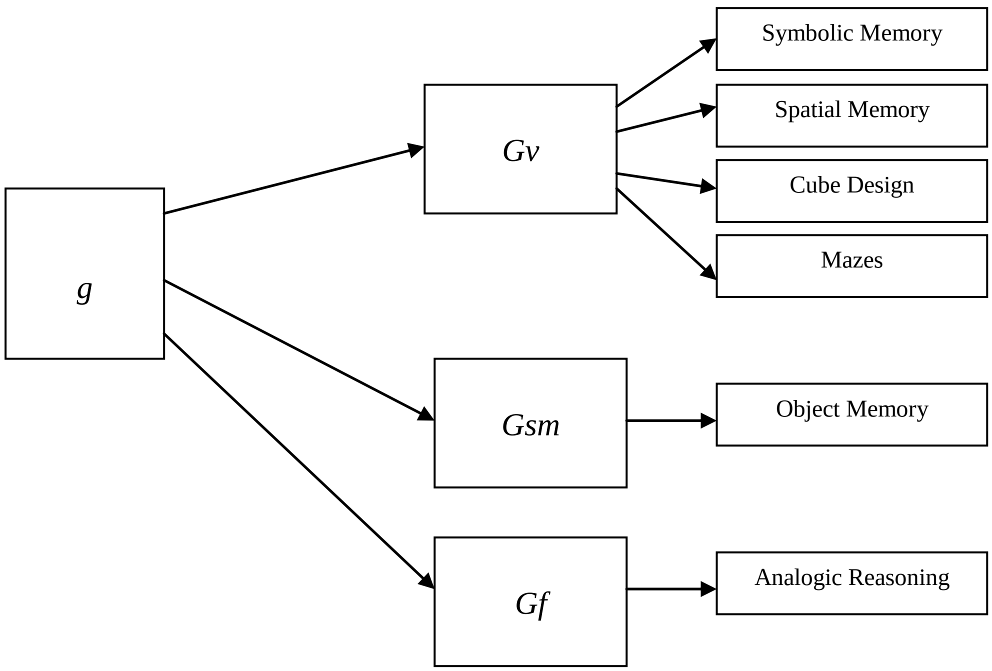
Appendix F

Factor Model for the Kaufman Assessment Battery for Children (K-ABC)



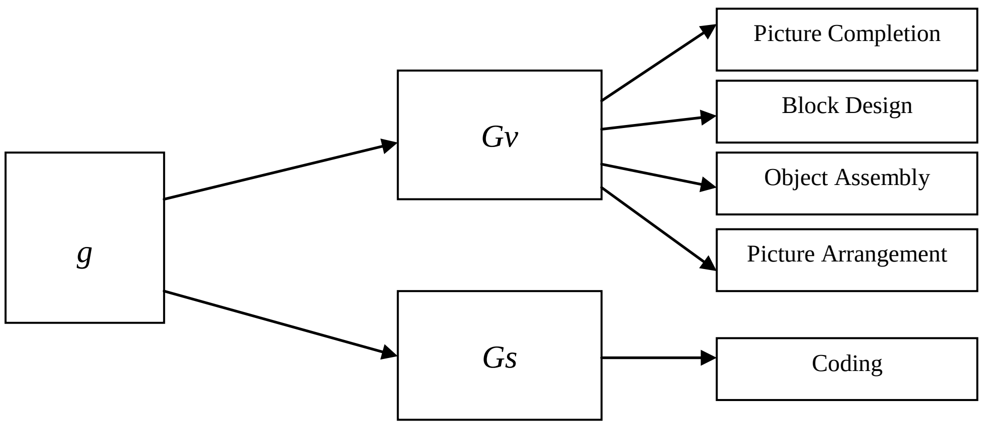
Appendix G

Factor Model for the Universal Nonverbal Intelligence Test (UNIT)



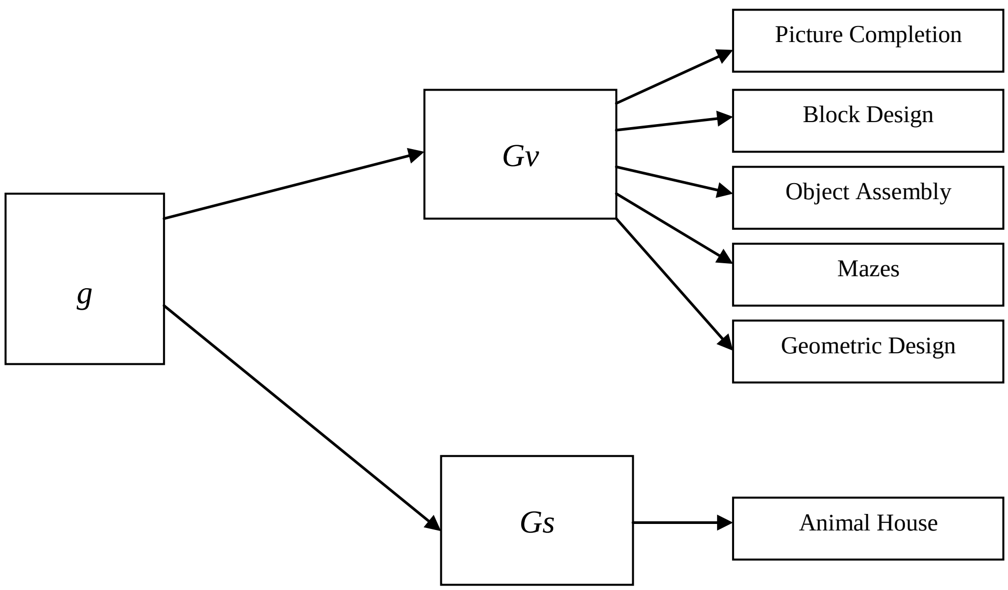
Appendix H

Factor Models for the Wechsler Intelligence Scale for Children (WISC) and the Wechsler Intelligence Scale for Children-Revised (WISC-R)



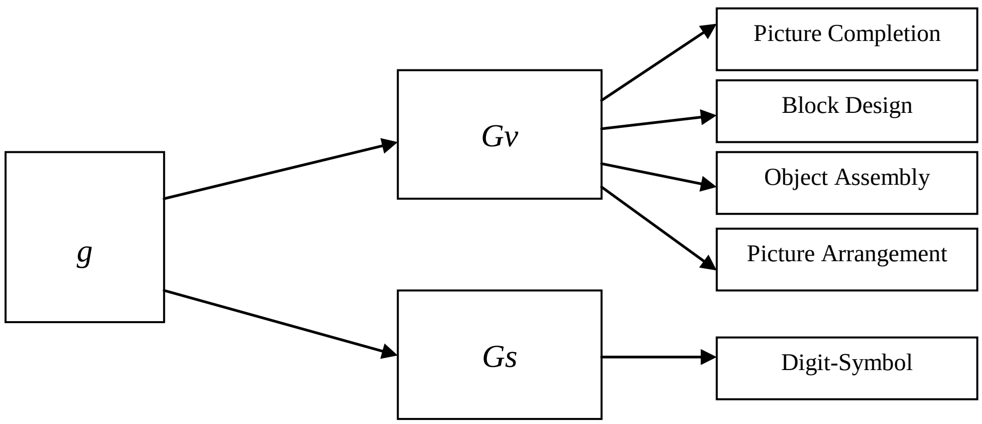
Appendix I

Factor Model for the Wechsler Preschool and Primary Scale of Intelligence (WPPSI)



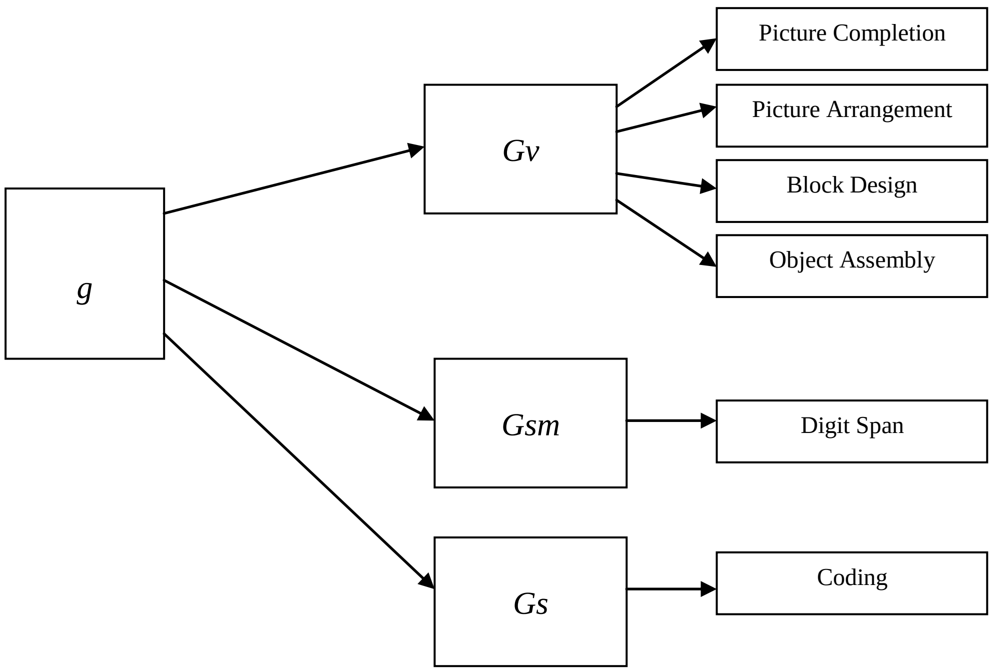
Appendix J

Factor Model for the Wechsler Adult Intelligence Scale-Revised (WAIS-R)



Appendix K

Factor Model for the Wechsler Intelligence Scale for Children-Third Edition (WISC-III)



Appendix L

CHC Model Factor Loadings for D/HOH Children of Preschool Age through Age 10 and their Normal-Hearing Peers

Model	Pathway	D/HOH Loading	Normal-Hearing Loading
ITPA (Hanson, Hancock, & Kopra; 1969)	*Gf – g*	.94	1.06
	Gv – g	.89	1.53
	VD – Gc	.86	.64
ITPA (Johnson, 1966)			
	Gf – g	.84	.96
	Gv – g	.70	1.16
	VD – Gc	1.24	.55
WISC (Murphy, 1957)			
	PA – Gv	.79	.69
	PC – Gv	.85	.84
	BD – Gv	1.05	1.09
WPPSI (Brinich, 1981)	*PC – Gv*	1.09	.94
	GD – Gv	.92	1.10
	BD – Gv	1.40	1.08