NATIONAL FORUM OF TEACHER EDUCATION JOURNAL
VOLUME 20, NUMBER 3, 2010

The Multidimensional Measure of Conceptual Complexity: Elucidating Misconceptions Research in Science

Gerald J. Calais
Department of Teacher Education
Burton College of Education
McNeese State University
Lake Charles, Louisiana

ABSTRACT

This article describes a novel instrument, the Multidimensional Measure of Conceptual Complexity, which is predicated on contemporary conceptual change literature, or misconceptions, and also on the advantages of Rasch measurement, designed to gauge one's ability to conceptually understand the nature of chemical equilibrium. The instrument itself is portrayed as occupying a two-dimensional space of conceptual complexity, reflecting hierarchical continua manifesting distinct degrees of conceptual breadth and conceptual depth. Evidence for various types of validity and reliability are discussed; implications for conceptual learning are provided, and other current research techniques that also focus on understanding conceptual change are discussed.

Introduction

Given that successful reform is contingent upon the pivotal development of appropriate assessments, it is crucial that assessment instruments match current attempts to augment students' conceptual, rather than factual or procedural, understanding of science. It is with this in mind that Brown (2005) developed the Multidimensional Measure of Conceptual Complexity (MMCC).

Rationale for Developing the MMCC

According to Brown (2005) the literature discusses numerous instruments, typically called conceptual diagnostic inventories, claiming to assess one's ability to conceptually understand distinct scientific disciplines. These instruments rely mainly upon the misconceptions literature, which has classified thousands of students' scientific misconceptions.

The items associated with these instruments (e.g., Force Concept Inventory [Hestenes, Wells, & Swackhamer, 1992]; Chemical Concepts Inventory [Mulford & Robinson, 2002]) generally share a common two-part format: Initially, using a simple experiment, the student predicts its outcome, usually via a multiple-choice format. Frequently, the student then either provides a brief written rationale for the answer or chooses an explanation derived from a second list of potential explanations. Each item assesses the extent to which a student understands a specific concept and includes distractors representing common misconceptions associated with the concept.

Brown (2005) reminds us that, unfortunately, two intrinsically detrimental factors vitiate the utility of conceptual diagnostic inventories. First, these inventories rely upon views of conceptual change--referred to in the literature, perhaps too flippantly, as misconceptions research--that are basically outdated as well as problematic. At the very least, the highly prolific misconceptions research was instrumental in highlighting one of the tenets of constructivism: students do not function at the outset of instruction as blank slates devoid of any pre-existing conceptions. However, Smith, diSessa, and Roschelle (1993) and Strike and Posner (1992) readily acknowledge the excessively rational and excessively simplistic applications of this research, also known as confront-and-replace approaches, in instructional settings. Hence, even when students' misconceptions can be successfully discerned by conceptual diagnostic inventories, the instructional utility of such results are indeterminable.

Second, these inventories basically reflect an all-or-nothing view of learning which contradicts one of the main tenets of constructivism: our previous knowledge builds and constrains all of our knowledge, because this view rules out students' incomplete yet fruitful steps converging on a path leading to complete and correct understanding. Furthermore, locating students along this continuum is impossible if we cannot conceptualize the steps along this path.

According to Wilson (2004) and Wright (1997), Rasch measurement, a branch of item response theory, is uniquely adept at addressing the two factors that impair these inventories, because it describes the various potential hierarchical and qualitative levels underlying latent variables—hence, the need for a complete theory of one's latent variable to be measured. Likewise, these measurements permit us to locate a student's understanding of a concept along these continua by providing us with individuals' meaningful, construct-referenced measures.

It is with the aforementioned in mind that Brown (2005) pilot-tested and analyzed his novel instrument, the Multidimensional Measurement of Conceptual Complexity (MMCC), which is based upon contemporary conceptual change literature, incorporates the advantages of Rasch measurement, and specifically measures students' conceptual understanding of chemical equilibrium.

Constructs: A Two-Dimensional Space of Conceptual Complexity

In this section, the theoretical background of Brown's (2005) MMCC is initially discussed. Then, the qualitative levels of conceptual depth and the qualitative levels of conceptual breadth are discussed, respectively. Next, conceptual structure vs normatively-correct/normatively-incorrect misconceptions are discussed.

Theoretical background

The MMCC is theoretically based in extensive conceptual change literature focusing on scientific domains. This literature does not reflect a consensus, however; instead, it rationalizes various theories that describe the conceptual structures that undergird students' understanding. Indeed, one of the MMCC's goals is that of furnishing a common framework to locate these distinctive perspectives.

Conceptual understanding, according to diSessa (1988, 1993), begins as phenomenological primitives, or p-prims, i.e., a collection of superficial and indistinct fragments. P-prims are superficial precisely because concepts lack articulated mechanisms to support them. For example, your deriving more of a result is often due to your working harder. In this case, the p-prim "greater effort spawns more result" merely posits this causal connection. Moreover, the application of p-prims' range of phenomena is frequently significantly different than established scientific domains. For example, the previously mentioned p-prim could easily apply to situations like pushing tables and carts but clearly not to pushing a foundation. Yet this range of application is too limiting from a physicist's perspective because all of them entail the application of a force. In addition, the identical p-prim has numerous applications beyond the realm of physics, such as convincing someone to hear your side of an argument or faithfully following a diet.

The conceptual understanding literature is expressed in forms other than that of knowledge based on p-prims. For example, Gopnik and Wellman (1994) and Vosniadou (1994) assert that student understanding corresponds to simplistic theories, whose application range more closely parallels scientific domains. Chi, Slotta, and de Leeuw (1994) and Slotta, Chi, and Joram (1995) emphasize the ontological contrast between two types of reasoning—matter-based and dynamic systems, with the latter necessitating a broader conceptual structure involving more internal elements. On the other hand, coordination classes, which are even more complicated structures, demand that numerous separate elements be coordinated and manifestly contribute to expert understanding (diSessa & Sherin, 1998).

These various forms of conceptual understanding can be perceived as occupying a two-dimensional space of conceptual complexity distinguished by hierarchical continua manifesting various degrees of conceptual depth and conceptual breadth. Brown's (2005) MMCC was designed to situate students within this space by gauging their positions along both constructs.

Qualitative levels of conceptual depth

The depth construct (Figure 1) differentiates six levels of conceptual depth. Hence, scientific phenomena are described by these levels from the perspective of hierarchically increasing deep structures of understanding; (0) the nonexistence of understanding; (1) acausal understanding whereby phenomena are not viewed as needing any justification; (2) elemental understanding predicated on a single causal element lacking justification; (3) justified understanding contingent upon a single causal element containing justification; (4) multiple causal understanding involving multiple causal elements, all essential yet independent; (5) emergent understanding also involving multiple causal elements interacting within a system, ultimately generating an emergent phenomenon.

Figure 1. The six qualitative levels of the depth construct.

Levels	Student's Perception of Phenomena	Description of Student's Responses
5 Emergent	The phenomenon is perceived as a system's emergent property, composed of interacting components. The observed effect is ultimately generated as the system progresses over time.	X_1 and X_2 occur. Their interaction and evolution over time ultimately results in the occurrence of Y. X_1 and X_2, in the meantime, do not cease to occur.
4 Multiple	The phenomenon is perceived as an effect generated via multiple causal elements. Since all are needed, the removal of one nullifies the effect.	The simultaneous occurrence of X_1 and X_2 causes Y.
3 Justified	The phenomenon is perceived as an effect generated via a single causal element. However, the phenomenon requires a mechanism or justification.	X simply causes Y, based on my explanation.
2 Elemental	The phenomenon is perceived as an effect generated via a single causal element. Neither a mechanism nor justification is entailed.	X simply causes Y.
1 Acausal	The phenomenon is perceived to be a mere representation of reality. Hence, no need for a cause.	The occurrence of Y is simply due to the nature of things.
0 Absent	The phenomenon astonishes the student. Apparently, no explanation is seen as plausible.	Actually, I'm unable to elucidate the occurrence of Y.

Justified versus unjustified understanding and understanding comprising single versus multiple causal elements are two rather independently distinct types of understanding that are incorporated in the depth construct. Based on the current depth construct (Fig. 1), accordingly, the Multiple level exhibits greater depth of conceptual understanding than does the Justified level. The rationale for this ordering is based on the hypothesis that students functioning at the Elemental level on a specific item are more apt to furnish justification involving a single causal element, when confronting an easier item, than to combine a second element. Similarly, students functioning at the Emergent level on a specific item are more apt to revert to Multiple understanding, when confronting a more complex item, than to revert to Justified understanding. Brown (2001a) based this hypothesis on teaching experience and prior research involving undergraduate chemistry students studying chemical equilibrium.

Qualitative levels of conceptual breadth

The breadth construct (Fig. 2) incorporates four hierarchical levels of conceptual breadth, with each level describing a wider range of applicability. This wider range of applicability occurs when the same causal element explains progressively distinct paired phenomena: (0) the absence of phenomena; (1) phenomena sharing the same set of actors, i.e., the objects associated with the phenomena; (2) phenomena sharing the same process, i.e., the kind of change (or absence of it) involving the actors; and (3) phenomena sharing neither actors nor processes. The hierarchical nature of these levels is likewise based on teaching experience and prior research focusing on chemical equilibrium.

Figure 2. The four qualitative levels of the breadth construct.

Levels	Student's Perceptions of Phenomena	Description of Student's Responses
1 Subsurface Consistent	Phenomena are perceived as being caused by the identical thing, despite their sharing neither a type of change (process) nor the identical objects (actors).	The identical causal element is used to interpret phenomena X and Y, both of which share a different process and actors.
2 Process Consistent	The same thing can be perceived to cause phenomena sharing a type of change (process), despite the objects (actors) being different.	The identical causal element is used to interpret phenomena X and Y, both of which implicate the same process but dissimilar actors.
1 Actor Consistent	Only phenomena implicating identical objects (actors) are perceived to be caused by the identical thing.	The identical causal element is used to interpret phenomena X and Y, both of which share the same set of actors.
0 Distinct	The causes of phenomena are seen as due to various things, even when the same objects (actors) are implicated.	Various causal elements are used to interpret phenomena X and Y, both of which implicate the identical actors.

Conceptual structure vs normatively-correct/normatively-incorrect misconceptions

The depth constructs and breadth constructs collectively depict the Multidimensional Measure of Conceptual Complexity (MMCC). Significantly, the MMCC neither specifies nor designates distinct concepts that are normatively-correct or misconceptions that are normatively-incorrect. Similarly, although the MMCC establishes what determines more and less depth or breadth, the MMCC does not assume that better understanding is represented by higher levels of either construct. Specific scientific concepts, for example, neither automatically involve multiple causal elements (depth) nor do they relate to all potential worldly paired phenomena (breadth). Hence, for any given classroom topic, the instructional goal will be a point situated somewhere within the MMCC's two-dimensional space; however, selecting this point is contingent upon the classroom teacher's specific topic and goal, whose specification depends upon this content and pedagogical knowledge.

In order to avoid focusing on learning from an all-or-nothing perspective, an approach characteristic of previous instruments, the MMCC deliberately concentrates on conceptual understanding via its fundamental structural characteristics. Consequently, the MMCC inherently enables teachers to utilize the immense literature on pedagogical strategies (e.g., bridging analogies [Clement, 1993], epistemological methods employing simplistic theories [Chinn & Brewer,1993], representations utilizing dynamic systems reasoning [Frederiksen, White, & Gutwill, 1999]) matching the various levels of conceptual understanding.

Design of the Instrument's Items

According to Brown (2005), previous item types associated with misconceptions research instruments frequently necessitated that students initially offer predictions (first question) and then justify their predictions (second question), i.e., the second question was designed to permit students to articulate the understanding they applied in responding to the first question. However, Brown suggests three reasons why this assumption's validity is questionable: First, students themselves are by no means the best judge regarding how they generated their own prediction, particularly if they are unable to articulate their conceptual understanding. Second, rather than relying on their own pertinent conceptual understanding to formulate their prediction, students may instead have depended upon memorized information, problem statement cues, or guessing tactics. With this approach, the student's prediction and basic conceptual understanding may have little in common. Third, when provided with enumerated potential explanations, students' own explanation may be either discarded or forgotten, selecting instead among the best-sounding explanations provided. Here, too, the student's prediction and basic conceptual understanding may have little in common.

Based on the above, Brown (2005) designed his MMCC's items with two specific goals in mind: First, it was necessary to avoid confounding potentially distinct prediction activities and their corresponding explanations. Second, it was also essential to avoid temptations associated with distractor-laden choices. Hence, the MMCC, which converges on the topic of chemical equilibrium, incorporates nine questions that are both open-ended and free-response and that are designed to prompt students to rationalize the occurrence of a specific chemical phenomenon. Each of these nine separate questions function as a depth construct item. On the other hand, each breadth construct item is represented by 36 discrete paired questions derived from the nine questions (e.g., 1-2, 1-3…1-9; 2-3, 2-4…2-9; 3-4, 3-5…3-9). To ensure a substantial quantity of three types of paired questions (paired questions sharing actors/objects, paired questions sharing processes, and paired questions sharing neither actors/objects nor processes) for the breadth construct, a 3x3 matrix consisting of three actors (the dissolution of solids in water, the evaporation of liquids in air, and the dissociation of acids in water) and three processes (reactions terminating at specific points, reactions terminating at disparate points due to temperature changes, and reactions terminating at disparate points due to changes in substances) was used. The selection of all nine phenomena associated with these nine questions was contingent upon their importance relative to the general chemistry curriculum and involved the following pivotal topics: solution, phase, and acid-base equilibrium.

Outcome Spaces

Brown's (2005) instrument was administered to 103 university students enrolled in chemistry classes at UC Berkeley. When students completed the instrument, they were then asked to take a survey which provided demographic information. Their responses to the nine questions were codified as depth codes and breadth codes.

Depth codes

The depth construct functioned as a guide for coding the 103 subjects' responses to the nine questions. Each student's response was designated as one of six possible levels contingent upon the general descriptions and distinct criteria in Figure 3. Each of the six levels (absent, acausal, elemental, justified, multiple, emergent) in the depth code column has a sample student response, which reflects that specific level, along with a rationale explaining why said response reflects that level. All items were based on the general coding criteria. Interestingly, the depth constructs' levels do not designate the utilization of specific causal elements; accordingly, responses entailing various causal elements occupy unique categories at each level.

Breadth codes

The specific causal element(s) have no impact on the depth codes; in contrast, they constitute the breadth codes' foundation. Accordingly, each of the 103 subject's responses to the 36 distinct paired phenomena was designated as one of two levels contingent upon the criteria in Figure 3. The first level focuses on a paired phenomenon such that both responses depended on the same causal element, despite both responses' depth. The second level focuses on a paired phenomenon whose responses did not depend on the same causal element, despite both responses' depth.

Figure 3. Outcome spaces for depth codes and breadth codes.

Outcome Spaces			
Depth Codes		Breadth Codes	
Emergent	"Salt dissolves in the water, and as it builds up in the water it starts to turn back into salt again, which happens faster and faster until it's happening at the same speed the salt is dissolving at" (Brown, 2005, p.14) would be coded as Emergent because it described the nature of the interaction occurring between/amongst the elements.	1	A code of 1 was assigned to each of the 36 discrete paired phenomena if both responses depended on the same causal element, despite both responses' depth . For example, if both phenomena's explanations for one of the 36 paired phenomena relied on water's restricted capacity for amassing dissolved salt particles, then this paired phenomena would be coded as 1, regardless if both explanations reflected an Elemental depth, or if one explanation were Elemental while the other was Justified or any other depth codes combination.
Multiple	"The rate at which salt dissolves in the water is equal to the rate at which it becomes salt again" (Brown, 2005, p. 14) would be coded as Multiple because the response integrated several independent causal elements.	0	A code of 0 was assigned to each of the 36 discrete paired phenomena if both responses did not depend on the same causal element, despite both responses' depth.
Justified	Responses coded as Justified included support for claims like the one in the previous statement coded as Elemental. "Because more salt would be too heavy" (Brown, 2005, p. 14) functions as support for the aforementioned claim and would be coded as Justified, in spite of whether or not said support is either correct or justified.		
Elemental	"The water can't dissolve any more salt" (Brown, 2005, p. 13) statement is Elemental instead of Justified because additional information would be needed to support this claim. Consequently, "Why can't the water dissolve any more salt?" (Brown, 2005, p. 13) remains an unanswered question.		
Acausal	Students' responses that did not assign a cause to any particular actor/object were coded as Acausal (e.g., "No more salt dissolves in the water" (Brown, 2005, p. 13) is merely a descriptive statement because no object is attached to the verb dissolve. However, "the water can't dissolve any more salt" (Brown, 2005, p. 13) statement implicitly assigns cause to the water, not salt or any other actor/object because salt is the object of the verb dissolve.) The latter statement reflects an Elemental code.		
Absent	Students' responses asserting the inability to explain the phenomenon were coded as Absent.		

Validity and Reliability Summary

This section discusses issues regarding several types of the instrument's validity and reliability, thereby demonstrating the utility of the MMCC.

Validity and Reliability Evidence

According to Brown (2005), the MMCC exhibits compelling evidence for construct validity, content validity, external validity, consequential validity, absence of bias differential item functioning, and reliability.

Construct Validity

The analysis of the MMCC's two articulated constructs, conceptual depth and conceptual breadth, manifested data fitting the model that applies both constructs in specifying each level of the hierarchical nature of the responses. More specifically, the parameter fit statistics for each item exhibited a symmetrical distribution devoid of outliers, suggesting that the data successfully matched the model's assumptions.

Content Validity

Chemistry content experts, including chemistry professors, judged the instrument's content to be valid yet difficult. Since subjects encountered no difficulty when responding to the instrument's nine questions, the experts' classifying the content as difficult reflects a low floor and a high ceiling for the items , such that both novices and experts were able to generate meaningful responses.

External Validity

There has thus far not been any direct comparison between the MMCC and other assessments claiming to gauge conceptual understanding that students have of general chemistry.

Consequential Validity

The design of the MMCC's open-ended item format renders the memorization of correct responses very difficult. Rather, the instrument's questions compel the students to clarify their conceptualization of scientific phenomena--the questions' desired behavior.

Bias and differential item functioning. The MMCC did not generate gender effects, nor was there any gender related differential functioning associated with the individual items. Ethnicity, as an external variable, was not investigated, however.

Consequential Validity

Determining if the function of an instrument’s utility can be subverted via techniques that augment students’ performance while not enhancing the latent ability that the instrument is designed to quantify is a major attribute of validity. With this in mind, the MMCC’s open-ended item format almost neutralizes students’ reliance on memorization for deriving correct responses. Of greater significance, however, each question directly induces the sought after behavior, i.e. elucidating scientific phenomena, instead of functioning as indirect indicators.

Reliability

According to Brown (2005), the instrument demonstrated moderate separation reliability (Wright & Masters, 1982); however, said reliability turned out to be essentially high when comparing the standard error of measurement to the construct’s quantified level. Consequently, this latter finding facilitates the reliable location of students within the construct’s single levels. At this point, however, there has been no investigation of inter-rater reliability.

Implications

The extensive conceptual change (misconceptions) literature has relied heavily on outdated and pedagogically questionable instruments designed to gauge conceptual understanding. The Multidimensional Measure of Conceptual Complexity (Brown, 2005), which incorporates the benefits of Rasch measurement--a branch of item response theory which describes the various potential hierarchical and qualitative levels underlying latent variables and permits us to locate a student’s understanding of a concept along these continua--adds enormous clarity to the current assessment of conceptual understanding by bringing order out of chaos. And though this instrument focuses only on the topic of chemical equilibrium at the university level, the philosophy underpinning it could be applied for assessing K-16+ learners’ conceptual understanding in virtually any topic in science, mathematics, social studies, literature, as well as in other disciplines.

This achievement, therefore, has immediate implications for classroom instruction at all levels because teachers need to ascertain learners’ misconceptions on any topic during all three phases of instruction: before, during, and post. If learners’ misconceptions on any topic are not addressed or understood, neither learning nor transfer of learning can materialize. Hence, the ability to successfully pinpoint where learners’ understanding of a concept lies substantially increases the odds of resolving this problem.

It should be noted that the impact of Brown’s (2005) research parallels Barnett and Ceci’s (2002) research on transfer of learning. The latter researchers, likewise, brought order out of chaos relative to the transfer of learning research literature by designing a comprehensive taxonomy for gauging the degree of one’s transfer of learning. Because of their taxonomy, it is now significantly easier to identify and locate learners’ degrees of transfer of learning along various dimensions

It would be most interesting and fruitful methodologically to combine Brown’s (2005) work on assessing conceptual understanding with that of Siegler’s (1995, 1996) overlapping waves theory and microgenetic analysis of conceptual change. The latter researcher employed his overlapping waves theory to determine which strategies children used over time to gain conceptual understanding; on the other hand, microgenetic analysis was used to examine the changes as they occurred, thereby identifying and explaining the underlying mechanisms of change itself.

Finally, Brown's (2005) work could likewise be combined methodologically with another aspect of Siegler's (2002) research: the use of self-explanations to enhance learners' academic performance. With this research focus, learners are asked to explain why a researcher's correctly solved problems are indeed correct and why incorrectly solved problems are wrong. At least four distinct mechanisms were identified that enabled self-explanations to generate their effects.

References

Barnett, S.M., & Ceci, S.J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*, 612-637.

Brown, N. J. S. (2001a). Heterogeneous equilibrium: A reformulation. Unpublished paper. University of California, Berkeley.

Brown, N.J.S. (2005). *The multidimensional measure of conceptual complexity* (Tech. Rep. No. 2005-04-01). Berkeley, California: University of California, BEAR CENTER.

Chi, M. T. H., Slotta, J. D., & de Leeuw, N. A. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, *4*, 27-43.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, *63*, 1-49.

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching*, *30*, 1241-1257.

diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. Pufal (Eds.), *Constructivism in the Computer Age* (49-70). Hillsdale, NJ: Lawrence Erlbaum.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*, 105-225.

diSessa, A. A., & B. L. Sherin (1998). What changes in conceptual change? *International Journal of Science Education*, *20*(10), 1155-1191.

Frederiksen, J. R., White, B. Y., & Gutwill, J. (1999). Dynamic mental models in learning science: The importance of constructing derivational linkages among models. *Journal of Research in Science Teaching*, 36, 806-836.

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257-293). Cambridge, England: Cambridge University Press.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*, 141-158.

Mulford, D. R., & Robinson, W. R. (2002). An inventory for alternate conceptions among first semester general chemistry students. *Journal of Chemical Education*, *79*, 739-744.

Siegler, R.S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, *28*, 225-273.

Siegler, R.S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Siegler, R. S. (2002). Microgenetic studies of self-explanations. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31-58). New York: Cambridge University.

Slotta, J. D., Chi, M. T. H., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction, 13*, 373-400.

Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, *3*, 115-163.

Strike, K. A. & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of science, cognitive psychology, and educational theory and practice* (pp. 147-176). Albany, NY: State University of New York Press.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, *4*, 45-69.

Wilson, M. (2004). Constructing measures: An item response modeling approach. Mahwah, NJ: Lawrence Erlbaum.

Wright, B. D. (1997). A history of social science measurement. *Educational Measurement: Issues and Practice*, *16*, 33-45, 52.

Wright, B. D. & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Chicago, IL: MESA Press.