This is the submitted version, which underwent peer review and was accepted and published December 20, 2020.

The doi is https://doi.org/10.1111/1467-9817.12335

The citation is: Goodwin, A. P., Petscher, Y., & Tock, J. (2021). Multidimensional Morphological Assessment for Middle School Students. *Journal of Research in Reading, 44*(1), 70-89. Retrieved from https://doi.org/10.1111/1467-9817.12335

The final accepted version can be accessed at (or you can contact the first author at amanda.goodwin@vanderbilt.edu) https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9817.12335?casa_token=u2W6ENTjSa0AAAAA:slGuIxP4CjhUKbVlet1Q7ymzu5IjNAuuCa9zZQdcC9cPTM5djq106Yb5hIx5To-ZlQbuj4h4Z7EqYQ

Running Head: Multidimensional Morphological Assessment

Multidimensional Morphological Assessment for Middle School Students

Paper Submitted to Journal of Research in Reading

For Special Issue on Morphology

December, 2019

Authors: Amanda Goodwin[1], Yaacov Petscher [2], & Jamie Tock[2]

Author Note:
[1]Department of Teaching and Learning at Vanderbilt University's Peabody College, Nashville, TN.
[2]Florida Center for Reading Research, Florida State University, Tallahassee, FL.

Contact information: (Corresponding Author)
Amanda Goodwin, Associate Professor, Department of Teaching and Learning, Vanderbilt University, 230 Appleton Place, Vanderbilt University, Nashville, TN 37203, USA
Amanda.Goodwin@Vanderbilt.edu, Phone: 615-936-8301

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150199 to Vanderbilt University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Abstract

Background: Middle school students use the information conveyed by morphemes (i.e., units of meaning such as prefixes, root words, and suffixes) in different ways to support their literacy endeavors, suggesting the likelihood that morphological knowledge is multidimensional. This has important implications for assessment.

Methods: The current study investigates the dimensionality of morphological knowledge considering the performance of 3,214 fifth through eighth graders on a range of morphological tasks (N=14 across the project's development and 10 for dimensionality analyses) and items (N=491) using multiple-group item response modeling. It then presents validation evidence related to performance of 1140 fifth through eighth graders on a gamified, computer-adaptive, multidimensional assessment of morphological knowledge which consists of seven morphological tasks and 181 items that make four morphological skills.

Results: Results indicate morphological knowledge is multidimensional and best represented via a bifactor model of four skills as well as task-related variance. These skills are Skill 1: Morphological Awareness, Skill 2: Use of Syntactic Morphological Knowledge, Skill 3: Use of Semantic Morphological Knowledge, and Skill 4: Use of Phonological and Orthographic Morphological Knowledge. This assessment designed after this model, called Monster, PI, was shown to be both reliable and valid with each morphological knowledge skill explaining additional variance in standardized reading vocabulary.

Conclusions: Findings suggest that morphological skills play an additive role in language and literacy outcomes. This indicates the importance of conceptualizing and assessing morphological knowledge as multidimensional. Implications for theory, research, policy, and practice are considered.

Implications for Practice

A) What is already known about the topic:

- There has been much evidence lately that morphological knowledge is likely multidimensional although this is still being studied. In general, models have conceptualized morphological knowledge (or morphological awareness) as being multidimensional and this multidimensionality looks different in different studies (i.e., different components and even different links to other constructs like vocabulary)
- The field has established that morphological knowledge links to different literacy performances, and studies that connect dimensionality with performance indicate performance in certain aspects of morphological knowledge connect more to certain aspects of literacy. In other words, most studies looks at the role of general morphological knowledge as assessed by a single or multiple tasks in predicting a literacy outcome. Recent research is suggesting that we need much more fine-grained tools to really understand the role of morphemes in reading and spelling development. To get that fine-grained analysis, we need to think about the multidimensionality of morphological knowledge and develop and empirically test models and their operationalization within assessments that can provide more fine-grained details about these roles.
- There are few studies that have looked at morphological knowledge in this fine-grained manner and there are no morphological assessments that assess the breadth and depth of morphological knowledge needed for these fine grained analyses. Additionally, most morphological assessments are researcher designed and the few standardized

morphological assessments are not readily available for use in research and practice. Hence, it is rare that these fine-grained analyses occur in either research or practice

B) What this paper adds

- This study is the largest study of morphological knowledge that we know of in terms of range of morphological assessments (14 across the project's development, 7 ultimately retained in the assessment), items (491 across the project's development, 181 delivered by the CAT in this study) and number of participants assessed (more than 3000 across the project with an analytical sample of 1140 students taking the final CAT). Because of this scope, we are able to present a novel conceptualization of morphological knowledge. We are also able to operationalize this model in a valid and reliable assessment that can move research and practice related to morphology forward.

C) implications for theory, policy, or practice.

- In terms of theory, this emphasis on multidimensionality and teachable skills is critical in terms of broadening the focus on what and how morphological knowledge supports literacy. It is the fine-grained tool that can lead to fine-grained understandings that is needed
- In terms of policy, morphological knowledge is not assessed as part of school district's yearly assessments. The operationalization of this model into a gamified, computer-adaptive measure makes the collection of such data now possible at scale and in a fun and motivating manner. Participants rated the game highly, so even in districts where too much testing is occurring, our participants did not see this as one more test but rather a fun game. Teachers saw it as a game their students enjoyed but which could provide them

with nuanced data regarding their morphological knowledge that they could then put into instructional practice.

- In terms of practice, both this theoretical model and the assessment can provide teachers with the needed data to inform their instruction. Additionally, this fine-grained lense can inform morphological intervention research to really think about how to best teach morphological knowledge as related to the different areas of literacy. This can lead to future work in identifying instructional profiles, better interventions, and actually getting morphological instruction into the classroom.

This special issue aims to build understanding of morphological processing. The current paper adds to this by investigating a multidimensional model of morphological knowledge and then operationalizing this model into a computer adaptive, gamified assessment for upper elementary and middle school students.

What is Morphological Knowledge? Connections to Theory

We define morphological knowledge as an umbrella term (Bowers, Kirby, & Deacon, 2010; Nagy, Carlisle, Goodwin, 2014) that encompasses different ways that adolescents use information conveyed by morphemes (i.e., units of meaning such as prefixes, root words, and suffixes) to support literacy endeavors. This includes the different automatic and strategic processing conveyed via terms like morphological processing, awareness, decoding, analysis, and problem-solving (see Table 1 in Nagy et al., 2014).

Our definition gets at the different information conveyed by morphemes and the different ways that such information is used. Related to processing, we are guided by reading theories like Perfetti’s (1988) verbal efficiency theory, which highlights the role of strategic actions of problem solving meanings based on parts as well as the less conscious processing of words more quickly “because of stronger, redundant links between the form of the word [including its parts] and its meaning” (Nagy et al., p. 4). Related to content, we draw on Perfetti’s Lexical Quality hypothesis (2007) to emphasize the role of different types of morphological knowledge. For example, a morpheme carries phonological information, or how the prefix, root word, or suffix is pronounced. It also carries orthographic/spelling, semantic/meaning, and syntactical/grammatical class information. Research suggests this knowledge is contained within a readers lexicon and that different aspects of morphological knowledge—and specifically root word knowledge—support lexical representations (Goodwin, Gilbert, Cho, & Kearns, 2014) and word reading efforts (Goodwin, Gilbert, & Cho, 2013). Hence, we use the term morphological knowledge to convey these different types of morphological content sources. We note here that the more word knowledge a readers has of a word, the more efficiently that lexical representation is accessed and used for comprehension. This links to ideas of constituent binding, stability and synchronocity, which suggest accessing the full representation rather than modular content. From these ideas, we consider the general morphological lexical representation as well as the specific content that makes it up.

Why Morphological Knowledge? Connections to Research

Morphemes convey the code of language, and hence we need to develop an assessment that can convey what students know about this code. As Nunes and Bryant (2006, p. 157) suggest, “Some of the most important correspondences between spoken and written language are at the level of the morpheme…The system of morphemes, therefore, is a powerful resource for those learning literacy.” We see this via the connection between morphological knowledge and word reading (e.g., Carlisle & Stone, 2005; Goodwin, Gilbert, & Cho, 2013), spelling (e.g., Deacon & Bryant, 2006; Nunes, Bryant, & Bindman, 2006), vocabulary (e.g., Anglin, 1993; Nagy, Berninger, & Abbott, 2006), and reading comprehension (e.g., Nagy et al., 2006; Tyler & Nagy, 1990).

Morphemes are important because they are combined to form words that express different meanings and serve different syntactic roles (e.g., *tasteful, distaste*). Yet their combinations often involve shifts in sound, spelling, grammatical role, and meaning (e.g., *finance, financial, financing, refinance*) that present challenges to students learning to read and write. For example, young spellers likely spell magician “magition,” because they do not distinguish the suffixes *tion* and *ian,* which sound alike but have different meanings (Nunes &

Bryant, 2006). We emphasize this link to the written orthography as languages like English are morphophonemic in that the spelling of the morpheme is privileged, conveying the link to meaning via it's representation in the written orthography. For example, *missed* is spelled with the past tense morpheme *ed* rather than */t/* to highlight the meaning conveyed within the suffix.

For adolescents, morphological knowledge is particularly important because it plays a major role in students' acquisition of academic language (e.g., Nagy & Townsend, 2012; Pacheco & Goodwin, 2013). About 60-80% of words in the academic texts that middle school students read are morphologically complex (Anglin, 1993); many are challenging low frequency derived words like *nationalistic* (Nagy & Anderson, 1984) that can be figured out from parts. With that said, what makes academic vocabulary so hard is that many are long morphologically complex words with multiple affixes (Nagy & Townsend, 2012), many of which are not learned until adolescence (Tyler & Nagy, 1990). Additionally, many of the complex syntactical structures through which academic vocabulary words are communicated relate to morphology (Nagy & Townsend, 2012). For example, nominalization is difficult (Nippold & Sun, 2008). This is where a verb or adjective is turned into a noun to head a noun phrase (which is a way of adding information density) via adding a suffix and changing the position of the word. Students who are aware of the syntactic role of suffixes are more likely to be able to comprehend the information in such phrases.

Multidimensionality & Assessment

The vast majority of work establishing the role of morphological knowledge in literacy primarily operationalizes morphological knowledge as a unidimensional construct, mostly via using a single measure of morphological knowledge to predict a literacy outcome. Yet the field has begun to move beyond this unitdimensional conceptualization to consider morphological knowledge as potentially multidimensional, and this has important relevance to assessment leading to the need for the current study.

Recent studies of morphological research highlight measurement problems (Apel, 2014; Apel, Diehm, & Apel, 2013; Bowers et al., 2010; Carlisle, 2010; Nagy, Carlisle, & Goodwin, 2014). Many researcher-developed measures were developed at a time when less emphasis was placed on reporting the validity, reliability, and other technical characteristics of measures. Most were developed to answer a particular research question and therefore focus on one aspect of morphological awareness (e.g., suffix use) rather than examining the multiple dimensions that capture the relation of morphological awareness and literacy (Goodwin, Petscher, Carlisle, Mitchell, 2017).

The three standardized morphological assessments that we know of (e.g., Berninger, 2007; Foorman, Petscher, & Bishop, 2012; Newcomer & Hammill, 2008) represent many of these challenges, and as a result, are used rarely within research and practice. Each has serious challenges as they are either not readily available, are not computer-adaptive/group administered, and/or provide limited data about students' morphological awareness. For example, Berninger's test (2007) is multidimensional, but must be individually administered, introducing time demands and possible measurement error. This test is also hard to find (i.e., not listed on the publisher's website). The other standardized measures, Foorman et al.'s (2012) CAT of morphological knowledge for grades 3 through 12 and the Morphological Completion subtest of the Test of Language Development (TOLD-P:4, Newcomer & Hammill, 2008) both assess a single component of morphological knowledge. Hence, the need for the current study.

To be clear, our emphasis on the need for a validated, multidimensional morphological knowledge assessment reflects the state of the field today as recent research has provided evidence for multidimensionality of morphological knowledge, and as we have shown above, theory supports this multidimensional conceptualization. Findings depend, though, on how multidimensionality is conceptualized and modeled as Spencer et al. (2015) found evidence for a unidimensional construct when considering differences between the format of assessment tasks (i.e., oral vs. written, multiple-choice vs. oral) and when using certain modeling structures (i.e., confirmatory factor analysis). In contrast, when considering differences in the morphological content measured and when considering a bifactor model, Goodwin et al. (2017) showed best fit of a multidimensional model with a general factor of morphological knowledge (created via a bifactor model involving overlap of seven morphological knowledge tasks) and specific factors for each task (representing the different content of each task). Such models provide evidence for considering the multidimensional content over format and also deepens understanding of how morphological knowledge contributes to literacy. For example, in the above study, Goodwin et al. (2017) showed that the general morphological knowledge factor and also the specific factor of morphological meaning processing predicted reading comprehension and reading vocabulary, controlling for the other factors. Also, the factor for generating morphologically complex words predicted reading vocabulary, controlling for the other factors. This indicates these morphological skills may be particularly important in literacy endeavors compared to the other skills considered.

Levesque and teams similarly showed the importance of conceptualizing morphological knowledge as multidimensional. In a 2019 study, Levesque et al. showed the contribution of morphological knowledge depended on the type of morphological knowledge being measured. Here, morphological analysis made a significant contribution to growth in reading comprehension whereas morphological awareness did not when controlling for prior performance. In a 2017 study, Levesque, Kieffer, & Deacon provided further evidence of different ways different types of morphological knowledge supported reading comprehension. This time, they reported that morphological awareness contributed to different mediators—morphological decoding and morphological analysis—which then predicted reading comprehension. This study showed that it matters whether one assesses morphological awareness, morphological decoding, or morphological analysis in terms of unraveling relationships to reading comprehension. Overall, recent research is highlighting consideration of various morphological skills to better understand the role of morphological knowledge in literacy. For example, Lam Chen, and Deacon (2020) showed that consideration of suffixes elucidated a way that morphological knowledge transferred from one language to another. These authors concluded that "awareness of cross-language suffix correspondences [i]s a novel aspect of morphological awareness that is specifically useful in second-language reading comprehension"(p. 9). This study again confirms the need to develop an assessment of morphological knowledge that taps into different morphological skills.

Current Study

The current study explores and operationalizes the multidimensionality of morphological knowledge into an assessment that can inform theory, research, and practice. Overall, the complexity of morphological knowledge has been seen as a challenge in terms of defining, assessing, and identifying the role of morphological knowledge in literacy outcomes, specifically reading comprehension (Nagy, Carlisle, & Goodwin, 2014). In contrast, Goodwin et al. (2018) suggests that this complexity can be an asset in that when channeled via the appropriate modeling technique, the complexity of morphological knowledge can build deeper understandings of the construct, its relationship to outcomes, and links to instruction. This can be done by identifying construct-relevant versus construct-irrelevant variance which related to morphology would involve morphological skills that could be instructionally relevant vs formats like multiple choice which are less relevant. Our research questions are as follows:

1) What is the dimensionality of morphological knowledge?
2) Does our multidimensional, gamified, computer-adaptive assessment assess morphological knowledge in a valid and reliable way?

Methods

Participants

Across the three years of the study, we worked in an urban district in the Southeastern United States. To answer our first question regarding dimensionality, we used our full sample of 3,214 fifth through eighth graders (*N*=1,026 fifth graders, 742 sixth graders, 715 seventh graders, and 731 eighth graders) learning in the classrooms of 15, 38, and 37 teachers in Years 1, 2, and 3, respectively. Based on demographic data from the district for *N*=3,041 of the sample, participants were 53.4% female; the race and ethnicity breakdown was as follows: 33.3% Black, 40.4% White, 21.9% Latino, 4.3% Asian, and <1% American Indian. Fifty percent of students were classified as economically disadvantaged, and 8% were identified as ELLs.

To answer our second question regarding the reliability and validity evidence for the final CAT, we consider our Year 3 sample. This consists of 1140 fifth through eighth graders (*N*=447 fifth graders, 258 sixth graders, 198 seventh graders, and 237 eighth graders) who took the CAT. Demographics of this group suggest the sample was 53% female; 43% White, 31% Black, 22% Latinx, and 4% Asian; 29% of students were eligible of free or reduced price lunch, and 5% of students were identified as an English Language Learners within the past two years at the time of participation of the study. With that said, many students spoke a language other than English at home (Spanish N=256, Arabic N=65, Amharic N=22, Vietnamese N=11, and Kurdish N=11 with thirty total languages reported being spoken). Note that for both the full and Year 3 sample, the way that students were classified by the districts grouped White students with students of Middle Eastern descent, many of whom spoke Arabic or languages other than English at home.

Assessment Development

Tasks & Items & Format

During Year 1, we piloted 12 morphological tasks either adapted from the literature or created by a panel of experts. These include 8 tasks retained (described below in the measures section--not *Sentence Sense* and the *Morphological Coherence Task* which were piloted Year 2) and four discarded because of poor performance due to ceiling effects, poor discrimination, lack of range, or high correlations with another task. The discarded tasks were from the literature (*Comes From*, Mahony, 1994; *Nonword Suffix*, Singson et al., 2000) and developed by our expert panel (*Karate Chop* where participants choose the best way of splitting a word into morphemes; *Passage task* where participants identified definitions for morphologically complex words within larger paragraphs). We also explored whether 1) reading items aloud with the goal of minimizing word reading demands or 2) order of presentation influenced performance, randomly assigning participants to 1) completing a task read aloud or only in writing and 2) to two different orders. Results indicated no effect of either reading manipulation (i.e., Cohen's d = -0.23) or order effects (i.e., average Cohen's d = -0.16; range d = -0.37 to 0.08).

Year 2 involved our Full Scale Experimental Linking Study where we administered the ten tasks listed in the measures section with linking items to connect Year 1 and 2 items. Also, we used a planned-missing design where students were assigned a set of items by grade level with linking items shared across grade levels.

Year 3 involved our validation study for the CAT, and also administering an additional set of items for the CAT. Students took the CAT in the first session and then fixed items in a second session with linking items connecting Years 1- 3 items. Hence, Research Question 1 investigating the dimensionality of morphological knowledge looks across performance on items from Years 1-3. Research Question 2 that explores performance on the CAT involves just data from Year 3.

Technology and Gaming System

Year 1 questions were administered on iPads using a traditional-testing format, but responses from students and teachers suggested that test-fatigue was occurring. Hence, we revisited ways to increase engagement and ultimately gamified the assessment, placing traditional questions within a larger storyline as part of .Monster, PI. (see Author et al., in press a, b and Author, 2019 for technical information on full system assessing vocabulary, syntax, and morphological knowledge). We focus on the morphological knowledge assessment in this paper.

Monster, PI features a mischievous monster causing chaos within different areas of a city like a school, museum, library, sports arena, and amusement park. Students take on the role of detective where they solve word puzzles—which are items from the assessment-- to earn clues to catch the monster! Clues are dependent on completion, not correctness. After completing an area (i.e., completing a few tasks), participants play a 30-60 second minigame to help refocus attention. In Year 2 and for the fixed items in Year 3, students completed the items assigned to them. In the Year 3 CAT version, which used a 2-parameter logistic item response adaptive set of algorithms, students took just the items relevant to their ability. Our informal data suggests students enjoyed the gamification. Of the students (N=2033) who provided quantitative feedback on the gamified version of Monster, PI, 93% rated it as 3 stars or higher (out of 5), 74% rated it 4 stars or higher, and 33% rated it as 5 stars. Additionally, 1100 students provided written feedback with student responses most often featuring “make”, “clues”, “end”, “enjoyed” and “fun”.

Measures

We organize the measures by the morphological skills they assess and then present the standardized vocabulary assessment used for validity purposes at the end. For each morphological task, we varied items in terms of transparency, frequency (of morpheme and derived word), length (number of morphemes) and family size (Carlisle & Katz, 2006; Goodwin & Cho, 2016). We also selected words that students would be likely to encounter by using academic word lists. Starred measures were retained in the CAT. We consider reliability of the full assessment rather than each task as the tasks are not analyzed separately in our models.

Skill 1- Morphological Awareness. The ability to reflect upon and manipulate morphemes was assessed via:

Odd Man Out (OMO).* Adapted from Ku and Anderson (2003), students were given a set of three similar-looking words with two sharing meaningful parts. Students identified the word did not fit. Examples include *corner, farmer, swimmer* and *season, seashore, seaweed*. We varied whether words overlapped in suffixes, prefixes, or root words.

Meaning Puzzles (MP).* Designed by the research team based on intervention work (Goodwin, 2016) and piloting (Pacheco & Goodwin, 2013), students identified words that shared morphemes versus words that only overlapped orthographically. They were given a target word (*astronomy*) and selected the answer choice that shared a morpheme (*astronaut*) from four answer choices (*fast, strong, as, astronaut*). Distractors had visual overlap with the target word.

Skill 2: Use of Syntactic Morphological Knowledge. Knowledge of how morphemes shift words' parts of speech was assessed via:

Real Word Suffix (RWS).* Adapted from Tyler and Nagy (1989), students read a sentence with one word missing—but the missing word's base word was presented. Students then had to select the syntactically correct form of the base word from four or five choices. An example was *It was a* **____________** *occasion when Barack Obama was elected president.* Students had to choose from choices of *historic, historian, history, historically.*

Making it Fit (MIF).* Adapted from Carlisle (1988), students were asked to read a low-context sentence with a missing word. Given the base, they typed the form of the word that fit. An example is as follows: *Some people argue that the* **______** *[sense] thing for Rosa Parks to do on her historic bus ride would have been to give up her seat, but instead, Rosa Parks stood up for what was right and started a movement.*

Skill 3: Use of Semantic Morphological Knowledge. Participant's ability to use the semantic information in morphemes to figure out the meanings of words was assessed via:

Word Detectives (WD).* Adapted from reading vocabulary tasks and Anglin's (1993) wordsolving work as well as Tyler and Nagy (1989), students were presented with a sentence containing a morphologically complex target word and four options of the target word's meaning. To increase motivation, the task was framed as a detective activity, with students encouraged to look for clues in the word and sentence to help figure out the word's meaning. An example was *There was <u>movement</u> in the upper body. The upper body was: A) staying still, B) full of bones, and C) changing position.*

Sentence Sense (SS). Students were presented with a morphologically complex word in a sentence. They determined the meaning and applied it to the full sentence to choose the inference that best fit the situation. For example, given '*After the hurricane, the city government set out to repair the streets only in* <u>*residential*</u> *areas.'* participants had to choose from 4 options to identify '*People needed to get to and from their homes safely.'* as the inference that applied the meaning of *residential* to the situation (*homes*). Distractor sentences were topically linked, but contained no other rhetorical (e.g. "In addition") or morphological connections to it.

Morphological Coherence (MC). Adapted from Uccelli, Galloway, Barr, Meneses, & Dobbs (2015), students matched a target word with the earlier text that referred to it. An example is "*Complex engineering problems* are often *impossible to manage* [correct answer] unless they are solved through teamwork. Construction workers, architects, and project managers *have to work together*. Often, projects that look <u>unmanageable</u> can be achieved through people working together."

Skill 4: Use of Phonological and Orthographic Morphological Knowledge. These sources of morphological knowledge were assessed via:

Morphological Spelling (SP).* Adapted from Carlisle (1988) and Nunes et al. (2012), to assess students' ability to spell morphologically complex words, students listened using headphones and were asked to spell the word heard (e.g., *selective*) using the iPad's keyboard.

Morphological Word Reading (WR).* Presented digitally, students listened to three pronunciations of a morphologically complex word and choose the correct pronunciation. Distractors were the mispronunciations were identified from pilot data where students had read a list of words aloud into an audio recorder. An example is _________, A) *selective,* B) *selecteyeve,* and C) *seelecteyeve.*

Word Hunt (WH). Modeled off of the test of silent word reading fluency (TOSWRF, Mather, Hammill, Allen, & Roberts, 2004) and piloted with good psychometric properties (Goodwin, 2016), 12 to 14 morphologically complex words were written on a single line without any spaces between the words. Participants tapped the screen where the word breaks should be (*camouflage/obstacle/diagram/indigo/shovel*).

Vocabulary *(Gates-MacGinitie Reading Vocabulary Assessment,* MacGinitie, MacGinitie, Maria, & Dreyer, 2000) was assessed at the end of the first testing session using Form S of the Level 5 through 8 Gates-MacGinitie Standardized Reading Vocabulary Assessment. Students answered 45 multiple-choice items where they read an underlined word within a phrase and choose the word or phrase that conveyed a similar meaning. This task is used extensively in research with strong reliability and validity. Extended scale scores were used.

Data Analysis

For Research Question 1 regarding the dimensionality of morphological knowledge, we used multiple-group item response modeling (MG-IRM). MG-IRM allowed for simultaneous testing of the factor structure for the items, the vertical equating of item difficulty, and the vertical scaling of person ability. We tested task, skill-level, and trait-level models. For skill and trait-level, we tested unidimensional, correlated, and bifactor models (see Author, 2019 for additional detail). For the final model, the ability of the individual and standard error associated with ability was estimated along with the item difficulty and item discrimination parameters.

For Research Question 2, we explored reliability and validity evidence for the final CAT. Marginal reliability (Sireci, Thissen, & Wainer, 1991) using the factor scores (i.e., student ability scores) and standard errors from the MG-IRT was estimated with

$$\bar{\rho} = \frac{\sigma_\theta^2 - \overline{\sigma_{e*}^2}}{\sigma_\theta^2}$$

where σ_θ^2 is the variance of ability score for the normative sample and $\overline{\sigma_{e*}^2}$ is the mean-squared error. For concurrent validity, we explored correlations amongst skills by grade level and then investigated links to a related standardized measure, standardized reading vocabulary using hierarchical multiple regression modeling. The goal here is to look at how this multidimensional CAT relates to standardized vocabulary, expecting a strong relationship (Nagy et al., 2006).

Results

Research Question 1: Dimensionality

Because data were collected over a three year period with a common-item, nonequivalent group design, data were specifically missing completely at random due to the planned missing data aspect of item deployment across samples. Table 1 includes descriptive statistics. Overall, developmental increases in the mean percent grade by grade level was noted for some (OMO, MIF, RWS, MCO, SSE, WD) but not others (MPU, WH, SP, and WR). This is expected prior to item culling and validation as creating stable linking items ensures a vertical scale.

To explore dimensionality, we analyzed various skill and trait-level MG-IRMs (see Table 2).[1] Models demonstrated excellent fit. As a reminder, we conceptualized the skills as follows: Skill 1- Morphological Awareness; Skill 2: Use of Syntactic Morphological Knowledge; Skill 3: Use of Semantic Morphological Knowledge; and Skill 4: Use of Phonological and Orthographic Morphological Knowledge.

[1] Note as unidimensionality of the tasks is assumed by the model, we first established task-level unidimensionality (See technical details at Authors, 2019). Also, the four-factor correlated model did not converge with the inclusion of higher convergence criteria and increased iterations.

For skill-level model results, the bifactor provided the best fit to the data for Skill 1, $\chi^2(117) = 177.65$, CFI = .95, TLI = .94, RMSEA = .031 (90% CI = .021, .039) and Skill 2, $\chi^2(75) = 73.75$, CFI = 1.00, TLI = 1.00, RMSEA = .000 (90% CI = .000, .023). For both Skills 3 and 4, each of the unidimensional, correlated trait, and bifactor models yielded acceptable fit with CFI and TLI at or above ~.95 and RMSEA < .05. For trait-level models, the bifactor model that included tasks as specific constructs and the four skills as uncorrelated, global constructs fit the data well, $\chi^2(13,666)$ =15,250, CFI = .97, TLI = .96, RMSEA = .015 (90% CI = .012, .025), as did a tri-factor model that included task-level factors, skill-level factors, and a global construct of morphological knowledge, $\chi^2(13,524)$ = 14,445, CFI = .98 TLI = .98, RMSEA = .011 (90% CI = .009, .017).

The viability of multiple trait-level models for morphology in the probit estimation (see Table 2 with TLI, CFI, and RMSEA values) led to an important consideration in the estimation of logit-based MG-IRM (see Table 3 for Log Liklihood, AIC, and BIC values); specifically, the balancing of test information (i.e., reliability) to be gained by the model with malleability of the factor. In other words, we considered additional information about the operationalization of the model. For example, a trifactor model would result in a global factor that would theoretically represent the most information about student performance, whereas for the bifactor model, the skill-level constructs would represent the most reliable portion of the data and yield malleable factors that facilitate the provision of instructional recommendations for teachers. Results of the logit-based MG-IRMs from flexMIRT software are reported in Table 3.

Overall, logit-based model results provided additional information to guide model selection. For skill-level models, results suggested the items were best represented by a global-type of construct whether unidimensional or bifactor. This is because as Table 3 shows, for each of the skill-level models, the correlated factors fit worse compared to either the 1factor or bifactor models with differences greater than at least 10 for the BIC which is considered to be practically important (Raftery, 1995). For Skills 1 and 2, AIC was in favor of the bifactor model and the BIC was in favor of the unidimensional model. For Skill 3, there were lower estimates of both AIC and BIC for the unidimensional model compared to bifactor, which contrasts to results for Skill 4, which had lower estimates of AIC and BIC for the bifactor model. Combining this information with earlier information led us to select the bifactor skill-level models for each morphology skill.

For trait-level models, the tri-factor provided the best relative fit compared to the bifactor model according to both the AIC (tri-factor AIC = 181596.42, 181687.88 vs. bifactor AIC = 186619.61, 186671.66) and the BIC (tri-factor BIC = 187977.83, 188069.29vs. bifactor BIC = 191346.36, 191398.41). With that said, a review of the test information for the tri-factor model demonstrated that little information about the skill factors was apparent (i.e., low or near 0 discrimination values) and only the global morphology factor yielded a reliable estimate of ability. Conversely, in the bifactor model, the skills yielded ranges of item discriminations that suggested good test information. This is where we had to balance the information from the model with malleability of the factor. With both of the bifactor and tri-factor models in the probit testing providing good fit to the test, combined with the bifactor model in the logit MG-IRM giving skill-level test information, we opted to select to the bifactor model for the purpose of reporting reliability and subsequently testing for validity of scores.

Research Question 2: Validation of Assessment

Next, we explored reliability and validity evidence for the final CAT. We started by investigating the precision of scores (i.e., marginal reliability). Although more details are contained in our technical manual (see Authors, 2019), results indicated strong reliability for most of the skills within the morphological CAT. Marginal reliability for Skill 1 (Morphological Awareness) was estimated at .90 compared to .93 for Skill 2 (Use of Syntactic Morphological Knowledge), and .92 for Skill 4 (Use of Phonological and Orthographic Morphological Knowledge). Reliability was .70 for Skill 3, which is acceptable, but lower than the other skills likely because fewer items and a single task was used to measure that skill. Overall, scores from the model were extracted as standardized scores with a mean of 0 and standard deviation of 1. In order to provide developmental scale scores for evaluating growth over time, the ability scores were transformed to have a mean of 500 and a standard deviation of 100.

For concurrent validity, correlations among the skill scores are reported in Table 4. Skills were consistently moderately correlated with each other across the grade levels with a lower bound of r = .53 between Skills 3 and 4 (representing Use of Semantic vs Phonological/Orthographic Morphological Knowledge) in grade 6 to an upper bound of r = .67 among several pairwise components (e.g., Grades 5-7: Skills 1 and 2—Morphological Awareness vs Use of Syntactic Morphological Knowledge; Skills 1 and 3- Morphological Awareness vs Use of Semantic Morphological Knowledge). In terms of connections to standardized reading vocabulary, results also suggest strong connections (see Table 5). Specific results for each model including the beta weights, standard errors, and *p*-values can be acquired by contacting the second author (Tables S1-S16). Overall, results showed that the combination of morphology skill ability scores resulted in 55%, 60%, 53%, and 50% of the variance in reading vocabulary in Grades 5, 6, 7, and 8 respectively, showing the expected strong link to vocabulary. It is important to note that the amount of variance explained in reading vocabulary increased substantially (Grade 5: 27 to 55%; Grade 6: 45 to 60%; Grade 7: 36 to 53%; Grade 8: 34 to 50%)

as the combination of skills was considered versus a single morphological skill, suggesting the importance of considering morphological knowledge in a multidimensional way.

Discussion

Research and theory have made clear the importance of morphological knowledge to literacy (Nagy et al., 2014). This has been done using different terms to represent slightly different conceptualizations of morphological knowledge often assessed via different unidimensional measures. In meta-analyses and synthesis studies that compare findings across a range of works, it often feels like apples and oranges are being compared. Our study suggests the reason synthesis work has been so difficult is because slightly different aspects of morphological knowledge are being assessed and taught. Hence, the relationship between morphological knowledge and the literacy outcome being studied (or even how morphological knowledge is being taught) depends on how morphological knowledge is being conceptualized and assessed.

This is where our study makes an important contribution to the literature. As part of developing our morphological assessment, we considered 14 morphological tasks, ultimately using 10 to provide dimensionality data and finally using seven to makeup our gamified CAT. Our findings provide further evidence that indeed morphological knowledge is multidimensional, consisting of four skills: *Skill 1- Morphological Awareness, Skill 2: Use of Syntactic Morphological Knowledge, Skill 3: Use of Semantic Morphological Knowledge, and Skill 4: Use of Phonological and Orthographic Morphological Knowledge*. These link to the different terms within the literature. For example, Skill 3 is similar to morphological analysis or morphological problem-solving. Skill 4 is similar to morphological decoding. These skills can be assessed reliably, which can later lead to unique instructional decisions and connections to specific literacy outcomes. Hence, assessing morphological knowledge in a multidimensional way is important. As part of this, studies need to communicate what conceptualization of morphological knowledge they are studying and assessing. Our findings suggest it is no longer enough to default to assuming morphological knowledge is a unidimensional construct assessed by a single measure. Multiple tasks of different content are needed to assess the range of content and ways that content can be used to support literacy endeavors.

In terms of connections to literacy, our study suggests that the role of morphological knowledge may be underestimated in many studies. As mentioned, most research studies use a single measure or consider a specific conceptualization of morphological knowledge. In our study, a single morphological skill explains substantially less variance than the four skills combined suggesting that these skills work together to support literacy tasks, which in our study involved reading vocabulary. In other words, if we had just explored the relationship between Skill 1 (Morphological Awareness) and reading vocabulary, we would have found between 27-45% of the variance in reading vocabulary explained depending on grade. But by considering all four skills, we observed a much stronger relationship (more than half, 50-60% of the variance explained). Clearly, understanding morphological knowledge as multidimensional deepens our understanding of the role of morphological knowledge in literacy outcomes.

From an assessment perspective, we have shown that you can assess multidimensional morphological knowledge in a valid and reliable manner—and in an efficient and fun manner as well as the assessment that we provided validation evidence for is group-administered on an iPad such that students take the items that are most relevant to their abilities. The assessment is framed as a game—and students and teachers seemed to enjoy it, which suggests that researchers and practitioners can build these understandings in engaging ways.

Overall, this study is the largest study of morphological knowledge that we know of in terms of range of morphological assessments (14 across the project's development, 7 ultimately retained in the assessment), items (491 across the project's development, 181 delivered by the CAT in this study) and number of participants assessed (more than 3000 across the project with an analytical sample of 1140 students in the CAT analysis). Because of this scope, we were able to explore various models of morphological knowledge (unidimensional, correlated traits, bifactor, and trifactor), and we ultimately arrived at one that identified teachable skills that can be associated with how readers apply morphological knowledge to different areas of literacy. We were also able to operationalize this model in a valid and reliable assessment that can move research and practice related to morphology forward.

In terms of theory, this emphasis on multidimensionality and teachable skills is critical in terms of broadening the focus on what and how morphological knowledge supports literacy. It is the fine-grained tool that can lead to fine-grained understandings that are needed. In terms of policy, morphological knowledge is not assessed as part of school district's yearly assessments. The operationalization of this model into a gamified, computer-adaptive measure makes the collection of such data now possible at scale and in a fun and engaging manner. Participants rated the game highly, so even in districts where too much testing is occurring, our participants did not see this as one more test but rather a fun game. Teachers saw it as a game their students enjoyed but which could provide them with nuanced data regarding their morphological knowledge that they could then put into instructional practice. In terms of practice, both this theoretical model and the assessment can provide teachers with the needed data to inform their instruction. Additionally, this fine-grained lense can inform morphological intervention research to really think about how to best teach morphological knowledge as related to the different areas of literacy. This can lead to future work in identifying instructional profiles, better interventions, and actually getting morphological instruction into the classroom.

While moving the field forward, there are certain limitations to take into consideration. First, we considered different models of dimensionality, but clearly there are other ways of thinking about the content and dimensionality of morphological knowledge. Future studies should explore this. Additionally, we considered dimensionality for the sample as a whole rather than for different groups of students. Hence future work should explore individual differences within this model.

Another challenge to note is our emphasis on the written orthography in designing our assessment. While we considered whether reading aloud versus presenting only in writing would result in differences, this was only considered in Year 1, so it is possible that word reading demands are confounded within some of our measures. We do have evidence that these are separate as Skill 4 (Use of Phonological/Orthographic Morphological knowledge) involved word reading, and this skills was only moderately correlated with the other skills, but future research should continue to unravel these relationships and ideally future studies will be able to collect many of the typical control variables in these studies like word reading.

Overall, this paper adds understanding of morphological processing by providing further evidence that morphological knowledge is multidimensional and suggesting a structure for that multidimensionality consisting of four skills, which themselves connect to the different types of information conveyed by morphemes. Additionally, the study operationalizes this model into a valid and reliable gamified assessment, which then suggests that when considering this broad conceptualization of morphological knowledge, the contribution to outcomes is larger and more meaningful than considering more unidimensional conceptualizations of the morphological knowledge.

References

Authors (2019).

Authors et al. (in press a, b).

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development, 58*(10), Serial # 238.

Apel, K. (2014). A comprehensive definition of morphological awareness: Implications for assessment. *Topics in Language Disorders*, *34*(3), 197-209.

Apel, K., Diehm, E., & Apel, L. (2013). Using multiple measures of morphological awareness to assess its relation to reading. *Topics in Language Disorders*, *33*(1), 42–56.

Berninger, V.W. (2007). Process assessment of the learner: PAL-II Diagnostic Assessment for Reading and Writing. TX: Pearson.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of educational research*, *80*(2), 144-179.

Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*, *9*(3), 247-266.

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading research quarterly*, *45*(4), 464-487.

Carlisle, J. F., & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. *Reading and Writing*, *19*(7), 669–693.

Carlisle, J. F., & Stone, C. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly*, *40*(4), 428–449.

Deacon, S.H., & Bryant, P.E. (2006). Getting to the root: young writers' sensitivity to the role of root morphemes in the spelling of inflected and derived words. *Journal of Child Language*, *33*, 401–417.

Foorman, B. R., Petscher, Y., & Bishop, M. D. (2012). The incremental variance of morphological knowledge to reading comprehension in grades 3-10 beyond prior reading comprehension, spelling, and text reading efficiency. *Learning and Individual Differences*, *22*(6), 792-798. doi:10.1016/j.lindif.2012.07.009.

Goodwin, A. P. (2016). Effectiveness of word solving: Integrating morphological problem-solving within comprehension instruction for middle school students. *Reading and Writing*, *29*(1), 91-116.

Goodwin, A. P., & Cho, S. J. (2016). Unraveling word learning: Reader and item-level predictors of word learning within comprehension instruction for fifth and sixth graders. *Scientific Studies of Reading*, *20*(6), 490-514. doi 10.1080/10888438.2016.1245734

Goodwin, A. P., Gilbert, J. K., & Cho, S. J. (2013). Morphological contributions to adolescent word reading: An item response approach. *Reading Research Quarterly*, *48*(1), 39-60. doi: 10.1002/rrq.037

Goodwin, A. P., Gilbert, J. K., Cho, S. J., & Kearns, D. M. (2014). Probing lexical representations: Simultaneous modeling of word and person contributions to multidimensional lexical representations. *Journal of Educational Psychology, 106*(2), 448-468. doi: 10.1037/a0034754

Goodwin, A. P., Petscher, Y., Carlisle, J. F., & Mitchell, A. M. (2017). Exploring the dimensionality of morphological knowledge for adolescent readers. *Journal of Research in Reading*, *40*(1), 91-117.

Goodwin, A., Petscher, Y., Reynolds, D., Lantos, T., Gould, S., & Tock, J. (2018). When complexity is your friend: Modeling the complex problem space of vocabulary. *Education Sciences*, *8*(4), 169.

Ku, Y. M., & Anderson, R. C. (2003). Development of morphological awareness in Chinese and English. *Reading and Writing*, *16*(5), 399-422.

Lam, K., Chen, X., & Deacon, S.H. (2020). The role of awareness of cross-language suffix correspondences in second language reading comprehension. *Reading Research Quarterly, 55*(1), 29-43.

Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2017). Morphological awareness and reading comprehension: Examining mediating factors. *Journal of experimental child psychology*, *160*, 1-20.

Levesque, K. C., Kieffer, M. J., & Deacon, S. H. (2019). Inferring meaning from meaningful parts: The contributions of morphological skills to the development of children's reading comprehension. *Reading Research Quarterly*, *54*(1), 63-80.

MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L. G. (2000). Gates-MacGinitie reading test (4th ed.). Itasca, IL: Riverside Publishing Company.

Mahony, D. L. (1994). Using sensitivity to word structure to explain variance in high school and college level reading ability. *Reading and Writing: An Interdisciplinary Journal*, *6*, 19–44.

Mather, N., Hammill, D. D., Allen, E.A., Roberts, R. (2004). Test of silent word reading fluency (TOSWRF). Austin, TX: Pro-Ed.

Nagy, W., & Anderson, R. (1984). The number of words in printed school English. *Reading Research Quarterly*, *19*, 304-330.

Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology*, *98*(1), 134.

Nagy, W., Carlisle, J. F., & Goodwin, A. P. (2014). Morphological knowledge and literacy acquisition. *Journal of Learning Disabilities*, 47(1), 3-12.

Nagy, W. E., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91-108.

Newcomer, P. L., & Hammill, D. D. (2008). *TOLD–P: 4 Test of Language Development Primary*. Austin, TX: Pro-Ed Inc.

Nippold, M.A., & Sun, L. (2008). Knowledge of morphologically complex words: A developmental study of older children and young adolescents. *Language, Speech, and Hearing Services in Schools*, *39,* 365–373. doi:10.1044/0161-1461(2008/034)

Nunes, T., & Bryant, P. (2006). *Improving literacy by teaching morphemes*. NY: Routledge.

Nunes, T., Bryant, P., & Barros, R. (2012). The development of word recognition and its significance for comprehension and fluency. *Journal of Educational Psychology*, *104*(4), 959-973. doi: 10.1037/a0027412

Nunes, T., Bryant, P., & Bindman, M. (2006). The effects of learning to spell on children's awareness of morphology. *Reading and Writing: An Interdisciplinary Journal*, *19*, 767-787.

Pacheco, M. B., & Goodwin, A. P. (2013). Putting Two and Two Together: Middle School Students' Morphological Problem Solving Strategies for Unknown Words. *Journal of Adolescent and Adult Literacy, 56*(7), 541-553.

Perfetti, C.A. (1988). Verbal efficiency in reading ability. In G.E. MacKinnon, T. G. Waller & M. Daneman (Eds.), *Reading research: Advances in theory and practice*, Vol. 6 (pp. 109-143). New York: Academic Press, Inc. Raftery, 1995

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383.

Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing: An Interdisciplinary Journal*, *12*(3–4), 219–252.

Sireci, S., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement JEM.*, *28*(3), 237–247. https://doi.org/10.1111/j.1745-3984.1991.tb00356.x

Spencer, M., Muse, A., Wagner, R., Foorman, B., Petscher, Y., Schatschneider, C., Tighe, E., Bishop, D. Examining the underlying dimensions of morphological awareness and vocabulary knowledge. *Reading and Writing*, *28*(7), 959–988.

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language, 28*, 649–667.

Tyler, A., & Nagy, W. E. (1990). Use of derivational morphology during reading. *Cognition, 36*, 17-34.

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary dcademic-language proficiency and its association with reading comprehension. *Reading Research Quarterly*, *50*(3), 337– 356.

Table 1.

Descriptive statistics for combined data (N = 3,240) over three years

	Grade 5			Grade 6			Grade 7			Grade 8		
	Nitems	M	SD	Nitems	M	SD	Nitems	M	SD	Nitems	M	SD
Combined Data (N = 3240)												
Morphology												
OMO	28	.59	.26	27	.66	.30	32	.67	.26	28	.68	.25
MPU	28	.48	.27	33	.60	.28	34	.64	.29	27	.60	.28
MIF	28	.29	.30	34	.41	.34	33	.44	.32	29	.45	.31
RWS	36	.54	.28	33	.56	.31	37	.63	.27	29	.67	.27
MCO	4	.66	.34	2	.73	.36	4	.73	.32	2	.76	.35
SSE	3	.48	.32	2	.49	.39	1	.57	.50	2	.62	.38
WH	5	.69	.31	7	.57	.26	7	.61	.25	4	.56	.28
WD	32	.44	.27	34	.56	.27	36	.58	.27	29	.62	.26
SP	13	.35	.30	12	.37	.31	12	.41	.31	13	.44	.30
WR	10	.71	.27	8	.66	.36	10	.86	.26	11	.79	.24

Table 2

Morphology Model Fit via Probit Link Skill-Level Models Trait-Level Models

Model Form	χ^2	df	RMSEA	LB	UB	CFI	TLI
Skill 1 - 1factor	300.44	135	0.047	0.040	0.054	0.78	0.75
Skill 1 - 2factor	230.28	134	0.036	0.028	0.044	0.87	0.86
Skill 1 – Bifactor*	177.65	117	0.031	0.021	0.039	0.95	0.94
Skill 2 - 1factor	140.22	104	0.025	0.013	0.035	0.93	0.92
Skill 2 - 2factor	136.86	103	0.024	0.011	0.035	0.93	0.93
Skill 2 – Bifactor*	73.75	75	0.000	0.000	0.023	1.00	1.00
Skill 3 - 1factor*	745.99	665	0.015	0.007	0.200	0.98	0.98
Skill 3 - 2factor*	742.62	664	0.014	0.006	0.020	0.98	0.98
Skill 3 – Bifactor$	658.17	627	0.009	0.000	0.016	0.99	0.99
Skill 4 - 1factor*	6549.41	2849	0.047	0.046	0.049	0.94	0.94
Skill 4 - 2factor*	6101.28	2848	0.044	0.043	0.046	0.95	0.95
Skill 4 – Bifactor$	5111.3	2772	0.038	0.037	0.040	0.97	0.97
Trait - 1factor	24182	13692	0.087	0.078	0.124	0.79	0.79
Trait- 4factor	-	-	-	-	-	-	-
Trait – Bifactor$	15250	13666	0.015	0.012	0.025	0.97	0.96
Trait – Tri-factor*	14445	13524	0.011	0.009	0.017	0.98	0.99

Note. *=Model showing acceptable fit to the data. $=Chosen model when two models showed acceptable fit. 1factor= all items associated with that skill or trait. 2factor= correlated with each factor reflective of task. Skill Bifactor=specific task factors and a global skill factor indicated by all items. Trait-Bifactor= task-level specific factors and skill-level global factors. Trait-Tri-factor= task-level factors, skill level factors, and a global factor for morphology indicated by all items. Trait 4factor model did not converge.

Table 3

Fit Indices for Models via Logit Link

	1factor	Correlated Factors 95% CI (LB, UB)	Bifactor 95% CI (LB, UB)	Tri-factor 95% CI (LB, UB)
Skill 1		*2factor*		
-2LL	64300.40	70878.65, 70881.45	63925.08, 63927.44	-----
AIC	64650.40	71242.65, 71245.45	64457.08, 64459.44	-----
BIC	65714.98	72349.82, 72352.62	66075.25, 66077.61	-----
Skill 2				
-2LL	58320.12	63885.40, 63891.05	57649.04, 57653.04	-----
AIC	58670.12	64249.40, 64255.05	58181.04, 58185.04	-----
BIC	59734.70	65356.57, 65362.22	59799.20, 59803.21	-----
Skill 3				
-2LL	41497.14	48866.84, 48871.09	41226.69, 41230.18	-----
AIC	41703.14	49108.84, 49113.09	41550.69, 41554.18	-----
BIC	42329.72	49844.92, 49849.18	42536.19, 42539.68	-----
Skill 4				
-2LL	33276.78	37246.97, 37251.36	31944.79, 31949.49	-----
BIC	34206.37	38322.05, 38326.44	33399.79, 33404.49	-----
AIC	33506.78	37512.97, 37517.36	32304.79, 32309.49	-----
Trait Level;		*4factor*	*Skill & Task*	*Trait, Skill, & Task*
-2LL	195442.49	-----	185065.61, 185117.66	179498.42, 179589.88
AIC	197562.81	-----	186619.61, 186671.66	181596.42, 181687.88
BIC	199671.24	-----	191346.36, 191398.41	187977.83, 188069.29

Table 4

Concurrent correlations among skills by grade level

Grade	Variable	Skill 1 Morphological Awareness	Skill 2 Syntactic Information	Skill 3 Semantic Information	Skill 4 Phon/Ortho Information
5	Skill 1	1.00			
	Skill 2	0.67	1.00		
	Skill 3	0.67	0.66	1.00	
	Skill 4	0.62	0.65	0.60	1.00
6	Skill 1	1.00			
	Skill 2	0.67	1.00		
	Skill 3	0.62	0.67	1.00	
	Skill 4	0.56	0.60	0.53	1.00
7	Skill 1	1.00			
	Skill 2	0.67	1.00		
	Skill 3	0.66	0.67	1.00	
	Skill 4	0.60	0.59	0.55	1.00
8	Skill 1	1.00			
	Skill 2	0.67	1.00		
	Skill 3	0.65	0.66	1.00	
	Skill 4	0.59	0.63	0.56	1.00

Table 5

R-squared values for hierarchical multiple regression by grade level

Grade	Model	Y3 Gates Voc
5	S1	0.27
	S1 + S2	0.51
	S1 + S2 + S3	0.53
	S1 + S2 + S3 + S4	0.55
6	S1	0.45
	S1 + S2	0.56
	S1 + S2 + S3	0.60
	S1 + S2 + S3 + S4	0.60
7	S1	0.36
	S1 + S2	0.47
	S1 + S2 + S3	0.52
	S1 + S2 + S3 + S4	0.53
8	S1	0.34
	S1 + S2	0.45
	S1 + S2 + S3	0.50
	S1 + S2 + S3 + S4	0.50