

Reliability and Validity of the Test of Morphological Structure

Hannah Krimm
Vanderbilt University

C. Melanie Schuele
Vanderbilt University School of Medicine

ABSTRACT

Rationale: To conduct a preliminary (i.e., small *n*) investigation of the reliability and validity of the Derivation and Decomposition subtests of the widely-used *Test of Morphological Structure* (Carlisle, 2000).

Method: The data that were analyzed included responses from 52 third-grade children who completed the TMS as part of a larger study. Analyses examined internal consistency, item difficulty, and item discrimination for each TMS subtest.

Results: The Derivation subtest demonstrated adequate internal consistency. However, internal consistency on the Decomposition subtest was low. Item difficulty was poorly distributed for both subtests. Item discrimination indices were low for both subtests.

Conclusions: Low reliability may relate to the syntactic and semantic features of the test items. These findings suggest opportunities for improvement in the measurement of morphological knowledge.

BACKGROUND AND METHOD

The *Test of Morphological Structure* (TMS; Carlisle, 2000) is a commonly-used research measure purported to assess morphological knowledge. It has been used with clinical populations (e.g., children with specific language impairment, children with learning disabilities) and non-clinical populations (e.g., children in general education). Notably, Carlisle (1985) originally proposed that the measure tapped a child's spoken and written production of specific morphological derivatives; it was not proposed as a measure of morphological knowledge. We hypothesize that the measure relies heavily on the respondent's semantic and syntactic knowledge. Hence, conclusions drawn about morphological knowledge from TMS results are confounded by children's semantic and syntactic proficiency. Thus, the purpose of this study was to provide preliminary evidence regarding the reliability and validity of the TMS using a small sample of third grade children.

Participants: Third-grade children (*n* = 52; mean age 8;11, *sd* = 5 months)

Measure: *Test of Morphological Structure* (TMS; Carlisle, 2000). Example items are shown below. Each item has a single correct answer provided in Carlisle (2000).

Derivation subtest (28 items)

E Farm. My uncle is a~ farmer.
C Farmer.

Decomposition subtest (28 items)

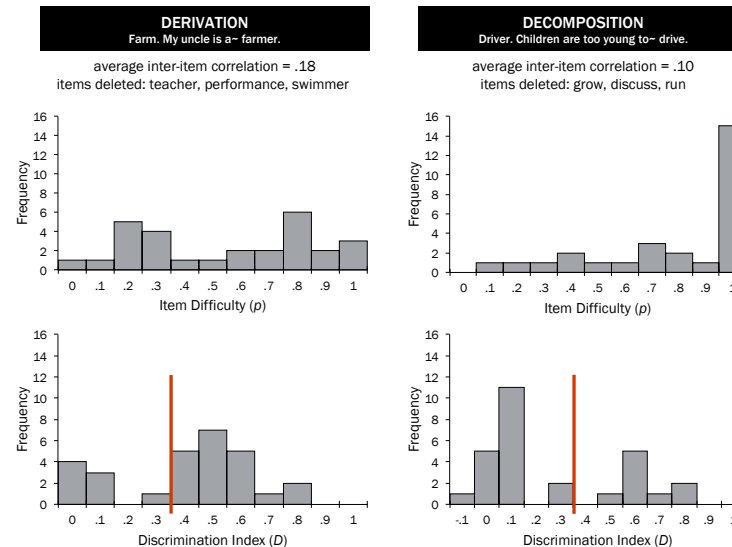
E Driver. Children are too young to~
C Drive.

Procedure: Participants individually completed the TMS as part of a larger study (Feldman et al., 2015). Responses were scored according to Carlisle (2000).

RESEARCH QUESTIONS

1. Do the TMS subtests demonstrate adequate internal consistency?
2. Do the TMS subtests demonstrate adequate item difficulty distributions?
3. Do the TMS subtest items demonstrate adequate item discrimination indices?

RESULTS



	Derivation			Decomposition		
	Mean	<i>sd</i>	range	Mean	<i>sd</i>	range
Difficulty (<i>p</i>)	0.55	4.75	0.04 – 1.00	0.78	0.28	0.08 – 1.00
Discrimination Index (<i>D</i>)	0.40	0.24	0.00 – 0.79	0.27	0.29	-0.07 – 0.79

RESULTS

1. Do the TMS subtests demonstrate adequate internal consistency?

No.
Both subtests required deletion of 3 items with no variance before alpha could be calculated. After deleting these items, $\alpha = .85$ for the Derivation subtest and $\alpha = .75$ for the Decomposition subtest. Items remained on the Decomposition subtest that were negatively correlated with the subtest score: *dry, five, elect, continue*.

2. Do the TMS subtests demonstrate adequate item difficulty distributions?

No.
Neither subtest shows a normal or uniform distribution.

3. Do the TMS subtest items demonstrate adequate item discrimination indices?

No.
The Derivation subtest contains 9 items (out of 28) with $D < .4$.
The Decomposition subtest contains 19 items (out of 28) with $D < .4$.

DISCUSSION AND IMPLICATIONS

These preliminary findings indicate opportunities for improvement in the measurement of morphological knowledge.

Ideally, item difficulty (*p*) would be uniformly or normally distributed. The bimodal distribution on the Derivation subtest indicates limited variability in item difficulty; students score similarly for most items. The skewed distribution on the Decomposition subtest indicates that most third-grade children answered most items correctly.

Discrimination indices (*D*) should exceed .4 for an item to be considered adequate. For this sample of children, several items had poor to marginal item discrimination indices (i.e., $D < .4$). These item characteristics suggest that the measure lacks the precision to capture individual differences.

It is likely that the observed lack of reliability on the TMS relates to syntactic and semantic item characteristics. Syntactically, several of the items on the TMS could be completed with multiple parts of speech despite that Carlisle designated a single correct answer for each item. For example, on the item *help. My sister is always~*, Carlisle specified *helpful* as the correct answer, but third graders in this sample often responded with *helping*. Semantically, items can be completed using vocabulary knowledge rather than morphological knowledge. For example, the item *teach. He was a very good~* can easily be completed with the familiar word *teacher* without knowledge of the agentive -er suffix. The observed weaknesses in item characteristics suggest the need for a more robust measure of morphological knowledge.

ACKNOWLEDGMENTS
This study was supported by a Preparation of Leadership Personnel grant from the US Department of Education (H325D140087; PI: Schuele) and the Vanderbilt CTSA grant UL1-RR024375-01 from NCI/NIH. Hannah Krimm's travel to SPIC2 was supported by NIH Conference Grant #P30-NS051877. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US Department of Education or National Institutes of Health.
Poster presented at the 2017 Symposium on Research in Child Language Disorders, Madison, WI
References available upon request: melanie.schuele@vanderbilt.edu Poster available at: www.mc.vanderbilt.edu/langdisorders
The authors have no financial or non-financial conflicts of interest to disclose.