
Believing in Fiction: Development and Initial Validation of the Literature Epistemology Scale (LES)

Project READI Technical Report #5

Mariya Yukhymenko-Lescroart, Stephen
Briner, Joseph Magliano, Kimberly Lawless,
Candice Burkett, Kathryn McCarthy,
Carol Lee and Susan Goldman

PROJECT **READi**

inquirium



Citation for this Report: Yukhymenko-Lescroart, M., Briner, S., Magliano, J. P., Lawless, K., Burkett, C., McCarthy, K., Lee, C.D., & Goldman, S. R. (2015). *Believing in fiction: Development and initial validation of the literature epistemology scale (LES)*. READI Technical Report #5. Retrieved from: www.projectreadi.org

Please send us comments, questions, etc.: info.projectreadi@gmail.com

Project READI was supported by the *Reading for Understanding (RFU)* initiative of the Institute for Education Sciences, U. S. Department of Education through Grant R305F100007 to the University of Illinois at Chicago from July 1, 2010 – June 30, 2016. The opinions expressed are those of the authors and do not represent views of the Institute or the U. S. Department of Education.

Project READI operated as a multi-institution collaboration among the Learning Sciences Research Institute, University of Illinois at Chicago; Northern Illinois University; Northwestern University; WestEd's Strategic Literacy Initiative; and Inquirium, LLC. Project READI developed and researched interventions in collaboration with classroom teachers that were designed to improve reading comprehension through argumentation from multiple sources in literature, history, and the sciences appropriate for adolescent learners. Curriculum materials in the READI modules were developed based on enacted instruction and are intended as case examples of the READI approach to deep and meaningful disciplinary literacy and learning.

©2016 Project READI

Believing in Fiction: Development and Initial Validation of the Literature

Epistemology Scale (LES)

Mariya A. Yukhymenko-Lescroart^a, Stephen W. Briner^b, Joseph P. Magliano^c, Kimberly Lawless^d, Candice Burkett^d, Kathryn S. McCarthy^d, Carol D. Lee^e, and Susan R. Goldman^d

^a California State University Fresno

^b Sacred Heart University

^c Northern Illinois University

^d University of Illinois at Chicago

^e Northwestern University

HIGHLIGHTS

- The Literature Epistemology Scale was developed for middle and high school students
- Factorial validity for three components was supported by confirmatory factor analysis
- Measurement invariance was established across gender and grade band
- Criterion validity was shown by associations with relevant constructs
- The final version contains 16 items to measure three subscales

Abstract

The scientific study of literature presents a unique epistemological challenge: how do people come to "believe" something about a work of fiction? Two studies were conducted to develop the Literature Epistemology Scale (LES) for middle and high school students and to establish its content validity, factorial validity, criterion validity and reliability. In Study 1, we established a multidimensional framework for understanding students' epistemological approaches to literature by identifying potential factors for the literary epistemology scale. Drawing upon a review of personal epistemology research, we pilot-tested a pool of 25-items with 199 school students. Confirmatory factor analyses (CFA) suggested a three-factor structure that may be important to measuring epistemological beliefs toward literature: multiple meanings, social functioning, and multiple reading. In Study 2, we used a split-half approach to develop the LES based on three emergent subscales and patterns of findings for the individual items in Study 1. Particularly, following content validation, a set of 29 items was pilot tested with 798 middle and high school students. A CFA with Sample 1 ($n = 399$) suggested several modifications, resulting in a refined scale with reasonable factor structure and internal consistency that led to a scale of 16 items. A CFA with Sample 2 ($n = 399$) confirmed a 3-factor structure and reliability. Factorial validity was evidenced by establishing invariance across grade band and gender; and criterion validity was shown by associations with relevant constructs. The 16-item LES is suggested as a reasonable measure of epistemological beliefs about literary texts of middle and high school male and female students. The implications of the results and directions for potential future research are discussed.

Keywords: epistemological beliefs; literary texts; self-report questionnaire; measurement invariance

Believing in Fiction: Development and Initial Validation of the Literature Epistemology Scale (LES)

When we read literature, we process more than just the nature of the characters and the structure of the plot. Often, literary works differ from other narratives in that literature often contains carefully crafted features that can invite the reader to form an interpretation of underlying themes supported by the text (Bortolussi & Dixon, 2003; Vipond & Hunt, 1984). Consider Frank Stockton's short story "The Lady, or the Tiger?" culminates with a young man opening a door, behind which is either a woman who wants to marry him, or a tiger who will devour him (Stockton, 1882). The story ends before we learn what was really behind the door, leaving the reader to form their own conclusions about the protagonist's fate. In leaving out the outcome of opening the door in "The Lady, or the Tiger?", Stockton asks the reader to interpret the story ending given content that he chose to include or not include, thereby coming to their own interpretation of the meaning of the story. But will readers choose to interpret how this story ends and how do readers come to these conclusions? Whether or not readers choose to infer how the story concludes may be contingent on their beliefs about literature and a reader's role in constructing meaning from it (Wilner, 2002). For example, some readers may believe that Stockton had a specific ending in mind, and may therefore attempt to search the text for clues about the author's intent. Other readers may be less concerned what ending the author intended (or whether he intended one at all), and instead focus on interpreting the text based on what their personal opinions or their affective response to the text. Readers who see literature as open to interpretation may therefore take the opportunity to interpret the fate of the protagonists in "The Lady, or the Tiger?", whereas readers who believe that meaning resides in the texts might find the open-ended conclusion to be irritating.

We contend that readers have epistemic beliefs about the nature of literature that affect the extent that they engage in literary interpretation. There is currently a wealth of research on how different epistemological beliefs may relate to learning and instruction (e.g., Braten & Stromso, 2008; Hofer & Pintrich, 1997). However, there is currently very little research on the epistemology of literature, and to date no instrument for measuring epistemological beliefs about literature exists. Thus, the goal of this study was to develop a scale for measuring epistemological beliefs about reading literature. Studies of literary comprehension would therefore benefit from a scale for measuring students' beliefs about literary reading and interpretation.

To create an epistemology of literary beliefs, we first must identify the ways in which readers differ in their beliefs about how to approach literary interpretation, and how to apply what one learns from literature in the everyday world. These beliefs would not be limited to specific literary works, as the demands made by the storytelling techniques of each story are likely to differ depending on the story, genre, author, or surrounding culture (Bogger-Mehall, 1996; Oska et al., 2010; Spiro et al., 1991). Second, this endeavor must draw on existing knowledge about literature and literary interpretation, including research conducted in psychology, literary studies, and education. Third, one must take into consideration existing epistemological theories and scales developed for domains in history and science, as there is much more research in these domains than in literature. Although it is unlikely that the constructs described in these previous epistemological theories will have one-to-one correspondence with an epistemological theory of literature, the existing theories of epistemology in other domains provide an important starting point for developing a theory of literary reading and interpretation.

Prior Research on Epistemic Beliefs

To date, a number of studies examining epistemology have focused on general epistemological beliefs as well as beliefs about epistemology in specific domains and topics (e.g., Braten & Stromso, 2009; Conley, Pintrich, Vekiri, & Harrison, 2004; Heft & Nasar, 2000; Hofer, 2000; Perry 1970; 1971; Schommer, 1990). There has been a wide variety of approaches to the study of epistemology in education, and a thorough survey of these studies provided important frameworks for building our epistemology of literature.

Early theories of epistemology assumed that epistemological belief and reasoning was domain-general, and that individuals used essentially the same skills to reason about science, history, literature, and other domains (e.g., Perry 1970; 1971). According to the Perry framework of epistemology, individuals experience changes to their epistemological beliefs in fixed stages, moving from believing that facts are immutable, to a more relativistic belief system in all beliefs hold equal value, to a more evidence-based epistemology in which truth may be discerned by examining the evidence for competing claims. Although Perry's work was highly influential on later theories of epistemology, it has met with several challenges, for example the lack of evidence for students moving through the epistemological stages in a fixed pattern (Schommer, 1990). Schommer's work reconceptualized epistemological beliefs as a more complex system of independent beliefs and suggested that more than one belief exist and that individuals may be sophisticated in some of them, but not necessarily in others.

More recent research and theories of epistemic beliefs and reasoning have suggested that epistemology is domain specific, rather than general (Hofer & Pintrich, 1997). This view, referred to as *personal epistemology*, suggests that individuals hold discrete theories about how knowledge works in different domains, such that a person could think that knowledge was more likely to be certain and concrete in one discipline (e.g., science) than in other disciplines (e.g.,

poetry) (Hofer, 2000). According to personal epistemology theory, there are two broad dimensions to consider when investigating epistemological attitudes to specific disciplines. The first is *the nature of knowledge* or beliefs about what knowledge is. The second dimension is the process by which a person believes they come to know something. This approach on domain-specific beliefs has been adopted in the subsequent studies, especially in the domain of science (e.g., Greene, Torney-Purta, & Azevedo, 2010; Stathopoulou & Vosniadou 2007).

There are several approaches of assessing epistemic beliefs. In some studies, researchers have employed qualitative methods such as interviews to explore epistemic beliefs of individuals (e.g., Perry 1970; 1971). Although qualitative research can provide insights related to epistemic beliefs, the limitations of such research include researcher bias, labor intensity, and challenges in terms of generalizable results and wider implications. Later domain-general and domain-specific scales of epistemology have taken a quantitative approach to assess epistemic beliefs (e.g., Braten & Stromso, 2009; Schommer, 2000). In these studies, each of these scales featured a series of multiple-choice items designed to measure theoretically identified constructs associated with epistemological beliefs. Some of these multiple-choice scales have adopted Hofer's (2000) framework and revealed several dimensions related to specific domains (e.g., psychology, earth science) and even topics (e.g., climate change). First, the dimension of *certainty* addresses whether people believe truth is certain and unchanging or complex and subject to change. Next, *simplicity* addresses whether people believe that truth tends to be simple and easy to understand, or complex and difficult to understand. Sometimes, the constructs of simplicity and certainty are collapsed into a single factor (e.g., Qian & Alverman, 1995). The dimension of *justification for knowing* focuses on how people evaluate claims, evidence, and experts. Lastly, *source of knowledge* dimension includes beliefs about whether a person can construct knowledge on

one's own or whether that knowledge must be transmitted from an outside source.

Of course measuring epistemology beliefs about a domain or topic has value to the extent that these beliefs affect learning outcomes, which has been shown to be the case (Braten, Britt, Stromso, & Rouet, 2011; Hofer & Pintrich, 1997; May & Etkina, 2002). For example, students who think that knowledge is simple or is acquired quickly learn less information during reading than those who believe that knowledge construction is complex and takes time to acquire (Peischl, Stahl, & Bromme, 2008; Schommer, Calvert, Christy, Garigiletti, & Bajaj, 1997). The reason for these differences may be based on how students approach a reading task. For example readers who assume that knowledge takes time to acquire tend to look for multiple sources of information (e.g., conduct more thorough searches of the internet) when learning about a topic than do readers who assume that knowledge can be acquired quickly (Braten & Stromso, 2006; Peischl et al., 2008; Schommer et al., 1997). As another example, students who hold more tentative beliefs about the certainty of knowledge within a topic have been shown to benefit more from an intervention that was designed to promote argument evaluation skills than those that hold beliefs that knowledge on the topic was fixed (Braten & Stromso, 2009; Braten, Stromso, & Samuelstuen, 2008; Stromso, Braten, & Samuelstuen, 2008).

Currently, there is very little research on the epistemology of literary reasoning, but several studies have focused on the epistemology of reading in general. For example, Schraw and Bruning (1996) measured students "beliefs about texts," especially about how readers believe they learn information from texts. They found that readers tend to hold "transmission" beliefs about reading (i.e., the author tells you what is true) or "translation" beliefs (i.e., the reader forms beliefs by challenging the author's claims). Similarly, Simsek (2010) constructed a scale about the beliefs about the function of language. The scale yielded three hypothesized factors:

epistemic function (whether reality can be conveyed through words), referential function (whether language corresponds to real life), and communicative function (whether words are sufficient for expressing thoughts, feelings, or experiences). These two studies have provided valuable insights into the epistemology of language and reading, however, neither scale specifically explored beliefs about literature. Given that literary reading may involve different psychological processes than when individuals read non-literary texts (Laszlo, 1999), it is likely that individuals will hold specific beliefs about how one reads literature as opposed to other texts. Therefore, to investigate the epistemology of literature, we must first develop a scale which specifically measures epistemological beliefs about the nature of knowledge in literature as well as beliefs about how one comes to understand a literary work.

Epistemology of Literary Reading and Reasoning

A scale to measure literary epistemology would be useful in helping researchers and educators understand why individuals can form such vastly different interpretations from the same literary works. Consider how a reader's beliefs about reading may affect how they interpret literature. For example, the poetry of Emily Dickinson often violates the traditional rhythm of the ballad stanza, leading to unusual meters in her poems (Denman, 1993). A reader who believes that Dickinson's poems are deliberately off-meter may form a very different interpretation from a reader who believes that the lack of traditional rhythm is a technical flaw in Dickinson's writing. Nevertheless, such differences in beliefs may not be enunciated by the readers, especially in educational settings where readers may feel intimidated questioning an author or the instructor. Finding a reliable way to measure these and other beliefs about literature would help lend transparency to this interpretive process.

Our approach for developing an epistemology of literary reasoning was informed by the

development of domain and topic specific scales in science and history (Braten & Stromso, 2006; Maggioni, VanSledright, & Alexander, 2009). Similar to domain and topic specific theories of epistemology, we assume that an epistemology of literary reasoning is comprised of multiple dimensions. However, as will become evident in the next section, those dimensions are unique to the domain of literature

Based on the literature review and consultation with experts in literary critique we identified five potential factors for our literary epistemology scale (Booth, 1961; Foucault, 1963; Lee, 2011; Mar & Oatley, 2008). Regarding beliefs about the *nature of knowledge*, we have identified potential factors regarding beliefs about an author's use of *rhetorical devices* to communicate a message and beliefs about whether a work of literature has *multiple meanings*, or several different possible interpretations. We identified three dimensions regarding the *nature of knowing*. The first involves beliefs about whether literature promotes *social functioning*, particularly, whether readers believe that events portrayed in fiction provide sufficient evidence for how to behave in everyday life, and whether readers believe literary authors are experts on understanding human behavior. A second concept related to the nature of knowing are beliefs about whether *multiple readings* of a literary work are necessary in order to fully understand it. The third concept refers to beliefs about the extent that *reflection* on a piece of literature influences one's understanding of that literary work.

Rhetorical devices. This factor relates to readers' knowledge that authors use rhetorical devices to convey their messages in literary works, such as symbolism or irony (Booth, 1961; Lee, 2011; Rabinowitz, 1987). Some readers, for example, may believe that most of the wording and phrasing in a literary work reflects deliberate choices on the part of the author, and are meant to convey specific meanings. Other readers, however, may view these word choices as

idiosyncratic or random, and thus may not believe that they reflect anything beyond a random choice on the author's part. Conceptually, this factor is similar to Hofer's (2000) factor of the simplicity or certainty of knowledge.

Multiple meanings. Although an author may have a particular intent or rhetorical goal when creating a work of literature, it is possible that readers may take away meanings from the text that the author did not intend, and it is further possible that these alternative interpretations are no less valid than the author's interpretation (Foucault, 1963; Ricoeur, 1973). Further, some stories are meant to be deliberately ambiguous. Thus, this dimension focuses on readers' beliefs about multiple interpretations to the same piece of literature. Some readers may view the story as having only one possible interpretation, whereas other readers may believe that many, even conflicting, interpretations are possible. As with rhetorical devices, multiple meanings is conceptually similar to Hofer's factor of beliefs about whether knowledge is simple and certain or complex and uncertain.

Social functioning. Works of literature are often fictitious, and frequently feature fantastical environments with no clear correspondence to the real world. However, these works of fiction may nonetheless provide readers with important cues for how to function in everyday life. Reading literature can serve as a simulation of real-life events, giving readers an opportunity to explore and understand cultures and scenarios they may not have otherwise been exposed to (Mar & Oatley, 2008; Mar, 2011; Mar, Oatley, & Peterson, 2009). Despite this, some readers may not believe that fiction can tell us about the real world, or that applying lessons from fiction to everyday life may in fact be counter-productive. This is similar to Hofer's concept of beliefs about the justification for knowing, as our social functioning factor measures whether knowledge derived from literature is useful for understanding the real world.

Multiple reading. Some readers will re-visit works of literature they have read, while other readers will avoid re-reading works they have previously encountered, instead preferring to read works they have not read before. Teaching the effectiveness of multiple readings of a work of literature is often on a focus in English Language Arts classes, and yet students often resist instructions to re-read (Hsieh & Dwyer, 2009). This dimension therefore focuses on the extent to which readers believe that multiples readings will improve their comprehension of a literary work. Conceptually, our multiple reading factor relates to Hofer's (2000) concept of the source of knowledge, as readers may hold differing beliefs about whether multiple reading conveys more information than a single reading.

Reflection. This factor explores whether students believe that thinking about a piece of literature that you have read in the past (i.e., *reflection*) increases one's comprehension of that literary work. This factor is related to Hofer's (2000) factor of the source of knowledge, except that in our factor, the source of knowledge is one's own reflection rather than an authority per se. Some readers may understand reflection to be an essential part of the process of understanding literature, while others may see the task as unnecessary. Such attitudes have been the focus on several studies of literature education (e.g., Kabuto, 2009), hence its inclusion in our initial list of epistemological factors.

The Present Study

The goal of this study was to develop a Literature Epistemology Scale (LES) for middle and high school students and to establish its content validity, factorial validity, criterion validity and reliability. Study 1 constituted a pilot study that involved creating the items for the LES and providing initial empirical data of psychometric properties of the LES. The purposes of Study 2 were to test factorial structure of the scale and to assess the factorial and criterion validity and

reliability of the scale.

Study 1

Phase 1: Items Development and Content Validation

Following literature review and establishment of a multidimensional framework for understanding students' epistemological approaches to literature, we generated a list of items to measure the five potential factors that we identified based on the review of literature and consultation with experts in literacy critique: multiple meaning, rhetorical devices, multiple reading, reflection, and social functioning. Existing scales of domain-general and domain-specific epistemological beliefs were considered. Once items were developed, they were submitted for content validation to ensure that they accurately measure the constructs being investigated, as recommended by DeVellis (2003), Polit, Beck, and Owen (2007), and Rubio, Berg-Weger, Tebb, Lee, and Rauch (2003). Establishing content validity is an important step in evaluating items prior to the pilot study conducted in Phase 2.

First, we created an initial pool of 35 items to measure five subscales: (1) five items for the multiple meaning; (2) five items for the rhetorical devices; (3) five items for the multiple reading; (4) five items for the reflection; and (5) eight items for the social functioning. These items were evaluated by a panel of six experts in the field of literacy critique. The experts were asked to identify construct category for each item by checking one of the six response options (including category "Other"), indicate their certainty of placing each item under the specified category on a 4-point scale ranging from 1 (*very unsure*) to 4 (*very sure*), and evaluate the relevance of the item to the corresponding construct category on a 3-point response scale ranging from L (*low*) to H (*high*). Additionally, the experts provided comments on the on wording of the items and, if needed, generated additional items.

Once the experts returned the evaluations, a factorial-validity index (FVI, Rubio et al., 2003) for construct category and item-level content validity index (I-CVI, Polit et al., 2007) for certainty and relevance were examined for each item to decide whether each individual item would be retained in the scale. The cut-off values for the FVI and I-CVI were set at .83 to reflect one disagreement among the experts and select items that were placed under the correct category by at least four experts. Based on the FVI and I-CVI and comments provided by the experts, 11 items were removed (i.e., one item in multiple meaning, two items in rhetoric devices, five items in reflection, and three items in social functioning), two items were reworded, and one item was added, yielding a 25-item scale designed to measure the five hypothesized latent variables.

Lastly, scale-level content validity indices (S-CVI, Polit et al., 2007) were computed for each construct with the final set of items using the averaging approach. S-CVIs for relevance were above the recommended minimum of .90, particularly, .96 for multiple meaning, 1.0 for rhetoric devices, 1.0 for multiple meaning, .93 for reflection, and .97 for social functioning, indicating excellent content validity of the final set of items. Thus, the 25-item scale was retained and pilot tested with a sample of middle and high school students to provide empirical evidence of the multidimensional framework for literature epistemology.

We hypothesized that a five-factor structure will emerge and the 25 items will measure the five hypothesized factors: multiple meaning, rhetorical devices, multiple reading, reflection, and social functioning.

Methods

Participants

Participants included a convenience sample of 199 conveniently secondary school students (57.8% females) from 11 classrooms studying in 7th ($n = 62$, 31.2%), 8th ($n = 126$,

63.3%), and 9th grades ($n = 11$, 5.5%). Only students for whom consents and assents were available were included in this study.

Measure

The 25-item scale developed for this study was designed to measure five hypothesized factors, and consisted of four-item multiple meaning, five-item rhetorical devices, five-item multiple reading, five-item reflection, and five-item social functioning subscales. Each subscale consisted of both positive, and negative, items. The response rate ranged from 1 (*strongly disagree*) to 5 (*strongly agree*) with all response options labeled. Followed by the items, participants were presented with a short set of instructions specifying what we meant by “literature” in the items, particularly, “You will read a series of statements about literature. By literature, we mean short stories, novels, poems, plays, or other things like that. Next to each question you should circle a number that indicates your agreement or disagreement with the statement. If you really, really agree with the statement you should circle the 5. If you really, really do not agree with the statement you should circle the 1. If you kind of do and kind of don’t agree, you should circle the 3.”

Procedure

After obtaining ethical approval from the institutional review board, parental consents, and student assents, participants were invited to complete a confidential survey on paper. The surveys were administered by the research staff during a class period.

Data Analysis

MPlus 6.12 (Muthen & Muthen, 1998-2010) was used to perform confirmatory factor analysis (CFA) and evaluate psychometric properties of the latent constructs. Before performing the CFA, we assessed multivariate normality with Mardia’s (1970) measure for multivariate

kurtosis, and Maximum Likelihood estimation was used with robust standard errors and scaled test statistics if data were not normally distributed. Robust Maximum Likelihood estimation allows for obtaining more accurate test results when data are not normally distributed and include missing values (Muthen & Muthen, 1998-2010; Yuan & Bentler, 2000). Then, items were submitted for CFA. Several fit indices were used to evaluate the model fit, including chi-square index, root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI), and Tucker–Lewis index (TLI). A non-significant chi-square indicates a good model fit; the chi-square test, however, is oversensitive to sample size, minor misspecifications, and multivariate normality. Alternatively, we examined chi-square to degrees of freedom ratio, which is considered a good fit if it is less than 2.0. Values of SRMR and RMSEA less than .08, and values of CFI and TLI greater than .90 indicate a reasonable fit, while a value of RMSEA less than .06 and values of CFI and TLI greater than .95 are required for a relatively good fit (Hu & Bentler, 1999). We also examined item loadings, modification indices, latent correlations, and omega reliability (McDonald, 1970). Item loadings should be significant and be at least .40; high modification indices may be a result of high correlations between measurement errors indicating cross-loadings or wording redundancy; high correlations between the factor scores indicate multicollinearity and may potentially distort parameter estimates or suggest that the model has too many factors; and omega reliability reflects the ratio of true score variance on total variance and is interpreted similarly to Cronbach’s alpha.

Results

Missing values ranged from 0 to 2.5% and were below the guideline of 5.0% (Fichman & Cummings, 2003). Before submitting items for analyses, negative items were recoded. Mardia’s (1970) measure for multivariate kurtosis indicated that data were not normally distributed

(737.5, $p < .001$), therefore a Maximum Likelihood estimation with robust standard errors was used.

Next, the 25 items were specified as indicators of their intended factors of multiple meaning, rhetorical devices, multiple reading, reflection, and social functioning. The results showed a misfit of the model to the data: $\chi^2(265, N = 199) = 441.9, p < .001$ ($\chi^2/df = 1.67$), RMSEA = .058, 90% CI [.048, .067], CFI = .78, TLI = .76, SRMR = .074. In attempt to pinpoint a model misfit, we examined factor loadings, modification indices, and latent correlations. Examination of factor loadings revealed eight items with low loadings (less than .40) for eight items. Modification indices suggested that two multiple reading items were potentially redundant. Additionally, high correlations emerged between Multiple Meaning and Rhetorical Devices, $r(197) = .90, p < .001$, and between Social Functioning and Reflection, $r(197) = .82, p < .001$, indicating multicollinearity. Highly correlated latent variables are problematic because one variable can account for all of the shared variance, which potentially can distort parameter estimates (Tabachnik & Fidell, 2001), and suggests that the model has too many factors (Kline, 2011). Multicollinearity can be resolved, where considered needed, by deleting one of the redundant variables or creating a composite variable if it makes conceptual sense to do so (Tabachnik & Fidell, 2001). Based on conceptual considerations, we created a composite variable by collapsing multiple meaning and rhetorical devices for one pair of correlated factors and deleted one subscale (i.e., reflection) and its corresponding items from the other pair of correlated factors.

Our decisions to collapse and delete factors were driven by theoretical and statistical considerations. Multiple meaning and rhetorical devices were collapsed because, conceptually, rhetorical devices is one way in which an author can convey multiple meanings to readers.

Furthermore, the two subscales measure beliefs about the nature of knowledge. Reflection was deleted because reflecting on a piece of literature to understand a literary work is one way in which beliefs about social functioning might be developed. Creating a composite variable by collapsing reflection and social functioning was not justified conceptually: social functioning is related to beliefs about justification for knowing, whereas reflection is related to the source of knowing. Statistically, reflection and its association items showed worse psychometric properties compared with social functioning, particularly, higher skewness ($M = -3.28$ vs. $M = -1.87$) and lower mean squared multiple correlations ($M = .15$ vs. $M = .21$). Finally, deleting the subscale can be beneficial for applied researchers, because a shorter scale will take students less time to complete, which is especially important when the scale has to be administered along with several other scales.

The second run of the CFA tested a 3-factor model with 20 items: four multiple meaning and five rhetoric devices items were specified as indicators of multiple meaning construct, and six multiple reading items and five social functioning items were specified as indicators of their intended constructs. Results were similar to that of the first model: the model had an inadequate fit, $\chi^2(167, N = 199) = 293.2, p < .001$ ($\chi^2/df = 1.76$), RMSEA = .062, 90% CI [.050, .073], CFI = .79, TLI = .76, SRMR = .074. Notably, the exact same eight items had low factor loadings (less than .40); and the two multiple reading items had high modification index. Nevertheless, the multicollinearity problem was resolved.

Next, eight items with low factor loadings and one of the pair of items with high modification index were removed in a series of CFAs, resulting in a 3-factor model with 11 items. In this model, three multiple meaning and two rhetoric devices items were specified as indicators of the multiple meaning construct, and three multiple reading items and three social

functioning items were specified as indicators of their indented constructs. This model showed a much better fit, $\chi^2(41, N = 199) = 62.4, p = .017$ ($\chi^2/df = 1.52$), RMSEA = .051, 90% CI [.022, .076], CFI = .94, TLI = .92, SRMR = .048. Modification indices did not suggest any further changes to the model. All factor loadings were significant (see Table 1 for standardized factor loadings). Omega reliabilities were .66 for multiple meaning, .70 for multiple reading, and .70 for social functioning. The correlations among the factors representing the subscales of the 11-item model were: $r(197) = .31, p = .019$ between multiple meaning and multiple reading, $r(197) = .66, p < .001$ between multiple meaning and social functioning, and $r(197) = .31, p = .012$ between multiple reading and social functioning.

[Table 1]

Discussion

The findings in this study represent the first step toward conceptualizing epistemological beliefs in literary comprehension, which is essentially multidimensional in nature. We hypothesized two factors related to beliefs about the nature of knowledge (multiple meaning and rhetorical devices) and three factors related to beliefs about the nature of knowing (multiple reading, social functioning, and reflection). However, our findings indicated an alternative three-factor latent structure. One latent factor involved multiple meanings, which was aligned with the nature of knowledge. Multiple meaning relates to readers' openness to new interpretations of a piece of literature. The degree to which readers interpret a piece of literature appears as a continuum moving from a novice to an expert view. For a novice reader, there is only one, direct and concrete interpretation of the literature. For an expert reader, interpretation of literature is tentative and evolving, thus, warranting multiple interpretations. The other two latent factors were multiple reading and social functioning, which were aligned with the nature of knowing.

Multiple reading concerns the extent to which readers believe that re-readings improve comprehension of a literary work. In the domain of literature, a piece of literature itself is a source of knowledge: knowledge about and interpretation of a literary work originates in an author and resides in the piece of literature, from which it may be transmitted. Multiple reading allows for constructing knowledge about the meaning of literature in interaction with the piece of literature and author. Thus, multiple reading allows for a shift in the action of knowing, with the reader moving from a passive reader to an active meaning constructor and interpreter. Finally, beliefs about social functioning relate to how readers evaluate a piece of literature and use it to understand the everyday world. As readers learn to evaluate their interpretation of and justify their beliefs about a piece of literature, they move through a continuum of beliefs that each piece of literature is nothing else but fictional, created by an author work to the realization that literature is connected to their lives. By relating fiction to everyday life, readers are able to justify their interpretation and meanings of literature. The three subscales were retained for Study 2, and modifications to items were made.

Study 2

Using three emergent subscales and patterns of findings for the individual items in Study 1, new items were generated and added to the final items from Study 1, resulting in a new pool of 37 items to measure three subscales: (1) 14 items for the multiple meaning; (2) 13 items for the social functioning; and (3) 10 items for the multiple reading. Then, a panel of seven experts evaluated the items. Based on the FVI and I-CVI, eight items were deleted and two items were reworded resulting in a 29-item scale. The values of S-CVI for relevance were 1.0 for multiple meaning, .93 for social functioning, and .98 for multiple reading, indicating excellent content validity of the final set of items. The 29 items were tested with a sample of middle and high

school students in Phase 2.

Study 2 was conducted to obtain psychometric data for the new items retained in the LES. We used a split-half approach to data analyses, which provided an opportunity to identify any further changes, optimize the length of the scale, and show initial evidence of reliability, factorial and criterion validity. Factorial validity was examined by exploring the factor structure of the LES. We hypothesized that the following three factors will emerge: multiple meaning, social functioning, and multiple reading, and that the subscales will have an adequate reliability.

Criterion validity was examined by exploring relationships between the scores on the LES and the Speed of Knowledge Acquisition subscale from the Wood and Kardash (2002) epistemology scale. Because the Speed of Knowledge Acquisition measures beliefs about quick learning (ranging from a naïve view that learning is a quick and straightforward process to a more sophisticated view that learning is a complex and gradual process), we hypothesized that it will be correlated with the multiple meaning and multiple reading subscales of the LES (Hypothesis 1).

Additionally, we examined relationships between reading habits and the scores on the LES. Students who like reading outside of schoolwork were expected to have higher scores on the multiple meaning and social functioning subscales and lower scores on the multiple reading subscale (the multiple reading subscale exclusively consists of negative items) (Hypothesis 2); and students who spend more time reading outside of school were expected to have lower scores on the multiple reading subscale (Hypothesis 3).

Finally, criterion validity was also assessed by determining whether gender moderates the relationship between reading habits and epistemological beliefs in literature. Recent research has suggested that gender differences do not emerge when epistemological thinking is defined in

terms of separate dimensions of epistemological beliefs (Conley et al., 2004; Pintrich, 2002). That is, there may not be gender differences when individuals are asked to focus on specific dimensions of epistemological beliefs (i.e., literature). However, boys in general read or enjoy reading less than girls (e.g., Cole, 1997; OECD, 2010; Pirie, 2002; Smith and Wilhelm, 2009), therefore, we wanted to separate out the effects of reading habits in the comparisons of epistemological beliefs among boys and girls. We expected that liking of reading and self-reported frequency of reading would be associated with students' epistemological beliefs in literature, meaning students who in general read or like reading more express more sophisticated beliefs. This allowed us to test whether boys and girls would differ in their epistemological beliefs in literature if they were equivalent on liking and frequency of reading. We hypothesized that the relationship between reading habits and epistemological beliefs were similar for boys and girls after controlling for their reading habits (Hypothesis 4).

Methods

Participants

Participants were 798 (53.5% females) middle and high school students from 47 classrooms in four middle and four high schools. Race-Ethnicity self-identification was 33.4% Hispanic or Latino, 24.1% White, 21.4% Asian, and 21.1% "Other" or multiracial. Middle school students ($n = 455$, 57.1% of the total sample) were in grades 6, 7, and 8 and between the ages of 11 – 15 years ($M = 13.2$, $SD = 0.95$). High school students ($n = 343$, 42.9%) were in grades 9, 10, 11, and 12 and between the ages of 13 – 18 years ($M = 16.1$, $SD = 1.27$). Gender was evenly distributed across the middle and high school grades, $\chi^2(1, N = 796) = 1.18, p = .283$. Race was not distributed evenly across the grade bands, $\chi^2(4, N = 794) = 47.13, p < .001$. Specifically, self-identification of middle school students was 39.8% Hispanic or Latino, 18.1%

Asian, 17.0% White, and 25.1% “Other” or multiracial; whereas high school students identified themselves as 24.9% Hispanic or Latino, 25.7% Asian, 33.3% White, and 16.1% “Other” or multiracial.

Measure

The 29-item scale, developed for this study, consisted of a 10-item multiple meaning, 10-item social functioning, and 9-item multiple reading subscales. The multiple meaning and social functioning subscales consisted of all positively worded items so that higher scores reflected more sophisticated beliefs. The multiple reading subscale consisted of all negatively worded items with higher scores reflecting less sophisticated beliefs. The response scale ranged from 1 (*strongly disagree*) to 5 (*strongly agree*) with all response options labeled. In this study, we also changed the instructions to “Many people enjoy reading literature. There are many genres to choose from: traditional literature, realistic fiction, science fiction, mystery, fantasy, informational non-fiction, and poetry, to just name a few. The following questions ask about your beliefs about the literature pieces you read. There are no right or wrong answers. Please read each of the statements listed below and indicate how much you personally agree with each statement. Circle the response that best represents your opinion.”

To assess criterion validity of the LES, students were also asked to complete a Speed of Knowledge Acquisition subscale from the Wood and Kardash (2002) epistemology scale. The scale consisted of seven items (e.g., “Almost all the information you can understand from a textbook you will get during the first reading”) and used a 5-point response scale asking students to indicate their agreement from 1 (*strongly disagree*) to 5 (*strongly agree*). Low scores on this factor represent a naïve view that learning is a quick and straightforward process, while high scores represent a more sophisticated view that learning is a complex and gradual process (Wood

& Kardash, 2002). Additionally, participants were asked to answer two questions evaluating their liking of reading (i.e., “How much do you like reading outside of schoolwork”) with three response options, particularly, 1 (*not at all*), 2 (*somewhat*), and 3 (*a lot*), as well as how often they read (i.e., “Compared to your peers, how much time do you spend reading outside of schoolwork”) with three response options, particularly, 1 (*less time*), 2 (*as much time*), and 3 (*more time*).

Procedures

After obtaining parent consents and student assents, a survey was administered on paper during a class period by research staff. The survey was confidential.

Data Analysis

To analyze the data, we randomly divided the 798 surveys into split-half samples. The analyses were performed using MPlus 6.12 (Muthen & Muthen, 1998-2010). Sample 1 ($n = 399$) was used to perform a confirmatory factor analysis (CFA) to test factorial structure of the scale, optimize its length by reducing the number of items that are redundant or may threaten the dimensionality of the scale, and assess its factorial validity and reliability. Sample 2 ($n = 399$) was used to perform CFA and structural equation modeling (SEM) to confirm the factorial structure of a refined scale identified with Sample 1, and to further establish its factorial validity and criterion validity. Factorial validity was examined through tests of measurement invariance for gender and grade band (middle and high school). Testing for equivalence of measures allows checking whether members of different groups ascribe the same meaning to items of the scale, thus, providing evidence of factorial validity. Measurement invariance evaluates whether structural relationships across the subgroups are equivalent and whether the scale is a valid instrument for the subsamples (Estabrook, 2012; Muthen & Muthen, 1998-2010; Netemeyer et

al., 2003). Specifically, we tested whether factor patterns, factor loadings, and factor variances were the same (i.e., invariant) across middle and high school students as well as males and females. Before testing for measurement invariance, baseline models were established for each subgroup, specifically, (a) across grade band for middle and high school students, and (b) across gender for males and females. Then, a sequence of progressively restrictive models was tested for equality of factor patterns, loadings, and intercepts. The models were compared based on changes in fit indices, using cut-off scores of .15 on changes in RMSEA and .01 on changes in CFI and TLI (Cheung & Rensvold, 2002; Chen, 2007). Factorial invariance is demonstrated when progressively restrictive models do not significantly change the fit of the model compared to the previous model as indicated by changes in RMSEA, CFI, and TLI.

Criterion-validity was examined based on latent correlations and path coefficients of the SEM models. The first model included three latent variables of literature epistemology and their corresponding observed items, a latent variable of speed of knowledge acquisition and its corresponding observed items, and two single indicator latent variables, specifically, liking of reading and reading frequency. Then, correlations among the variables of the first model were examined. The second model included three latent variables of literature epistemology and observed items, two single indicator latent variables, specifically, liking of reading and reading frequency, and gender. In this model, three latent variables of literature epistemology were specified as outcomes, and liking of reading and frequency of reading, and gender as indicators. Then, parameter estimates were examined for significant paths between reading habits and literature epistemological outcomes. In these analyses, multiple reading items were recoded so that higher scores on the multiple reading subscale reflected more sophisticated beliefs.

Results

Sample 1

Missing values ranged from 0 to 2.3% and were within the recommended norms (Fichman & Cummings, 2003). Mardia's value for multivariate kurtosis (518.5, $p < .001$) indicated non-normal distribution of the data, therefore, a maximum likelihood estimation with robust standard errors was used again. The 29 items were specified as indicators of their intended factors of multiple meaning, social functioning, and multiple reading. The results showed inadequate model fit: $\chi^2(374, N = 399) = 822.2, p < .001$ ($\chi^2/df = 2.20$), CFI = .892, TLI = .882, RMSEA = .055, 90% CI [.050, .060], SRMSR = .068. Items that could threaten the dimensionality of the scale or were redundant were revealed based on high values of modification indices, and these items were removed in a series of CFAs. A three-factor 16-item model reflecting five items for multiple meaning, five items for relevance to life, and six items for multiple reading showed a good fit: $\chi^2(101, N = 399) = 128.4, p = .034$ ($\chi^2/df = 1.27$), CFI = .984, TLI = .981, RMSEA = .026, 90% CI [0, .039], SRMSR = .041. Modification indices did not suggest any further changes to the model. All item loadings were significant (see Table 2 for item loadings). Omega reliabilities were reasonable: .79 for multiple meaning, .85 for social functioning, and .89 for multiple reading. The latent correlations among the factors were $r(397) = .30, p < .001$ for multiple meaning and relevance to life; $r(397) = -.43, p < .001$ for multiple meaning and multiple reading; and $r(397) = -.34, p < .001$ for relevance to life and multiple reading.

[Table 2]

Sample 2

Confirmatory factor analysis. The 16 items were specified as indicators of their intended factors of multiple meaning, social functioning, and multiple reading. Results indicated

a good model fit: $\chi^2(101, N = 399) = 124.3, p = .058$ ($\chi^2/df = 1.23$), CFI = .987, TLI = .985, RMSEA = .024, 90% CI [0, .037], SRMSR = .035. Table 2 presents item loadings, which all were significant; Table 3 presents latent correlations among the factors. Omega reliabilities were .79 for multiple meaning, .85 for relevance to life, and .89 for multiple reading, indicating reasonable reliabilities of the subscales.

Factorial validity. Table 3 shows the fit indices for the models testing for measurement invariance across grade level and gender. The scale was invariant across grade levels and gender as indicated by changes of less than .15 in RMSEA and less than .01 in CFI and TLI. Thus, factorial validity was established by demonstrating reasonable fit indices that did not change substantially across the models measuring invariance of factor pattern, loadings, and variances, suggesting an enhanced generalizability of the LES for middle and high school students as well as for both genders (Netemeyer et al., 2003).

[Table 3]

Descriptive statistics and scale reliabilities. Scale scores were computed for each subscale by averaging corresponding items (i.e., mean scores were computed for multiple meaning, social functioning, and multiple reading). All multiple reading items were recoded prior to computing the means so that higher scores on the multiple reading subscale indicated more sophisticated beliefs. Likewise, a scale score was computed for *speed of knowledge acquisition* subscale. Table 4 shows descriptive statistics, Cronbach's alphas, and correlations for literature epistemology subscales (i.e., multiple meaning, social functioning, and multiple reading) and speed of knowledge acquisition, as well as single items that measured students' liking of reading and frequency of reading. Mean values on the LES showed that, on average, students "agreed" that a literary work has several different possible meanings, had more neutral

beliefs about whether literature promotes social functioning by providing example scenarios which readers can apply to their lives; and reported that multiple readings of a literary work were effective to fully understand it.

[Table 4]

Criterion validity. To estimate measurement and structural components of the first model, the 16 items were specified as indicators of their intended factors of multiple meaning, social functioning, and multiple reading; seven items were specified as indicators of speed of knowledge acquisition; and 2 items were specified as single indicator latent variables of liking of reading and reading frequency. The model showed an adequate fit to the data: $\chi^2(262, N = 399) = 388.81, p < .001$ ($\chi^2/df = 1.48$), CFI = .954, TLI = .947, RMSEA = .033, 90% CI [.027, .042], SRMSR = .047. Factor loadings for each item were significant. Table 4 shows latent correlations among the variables. As expected, speed of knowledge acquisition was negatively correlated with multiple reading, $r(397) = -.49, p < .001$ and multiple meaning, $r(397) = -.50, p < .001$, providing support for the hypothesis 1. Liking of reading was positively correlated with multiple reading, $r(397) = .52, p < .001$, relevance to life, $r(397) = .44, p < .001$, and multiple meaning, $r(397) = .32, p < .001$, supporting hypothesis 2. Self-reported frequency of reading was positively correlated with multiple reading, $r(397) = .36, p < .001$, relevance to life, $r(397) = .21, p < .001$, and multiple meaning, $r(397) = .17, p = .006$, thus supporting hypothesis 3.

[Table 5]

Finally, we tested whether there were differences in literary epistemological beliefs across gender. Given that boys generally read or enjoy reading less than girls do, it is important to control for reading habits in order to obtain reliable estimates of gender effects. To do so, we specified a model with relationships between gender, liking of reading, reading frequency, and

three latent factors of literature epistemology. The outcomes were latent factors of multiple meaning, social functioning, and multiple reading their corresponding items. This model statistically adjusts for differences in students' reading habits, allowing testing differences in the epistemological beliefs across gender if boys and girls were equivalent on reading habits. The regression coefficients are shown in Table 6. Effects were found for only liking of reading, not for gender (supporting hypothesis 4), or frequency of reading. Students who did not like reading had lower scores on multiple meaning, social functioning, and higher scores on multiple reading than students who liked reading, indicating that students who did not like reading tended to hold less sophisticated beliefs about literary knowledge and knowing.

It is important to note that when analyses were run without including liking of reading and reading frequency, the results suggested gender differences in epistemological beliefs, particularly, for multiple meaning and multiple reading constructs (with boys showing less sophisticated beliefs). Nevertheless, the differences in multiple meaning and multiple reading were a functioning of liking of reading, not gender. Had we not included reading habits, we would have erroneously concluded that there were gender differences in epistemological beliefs in literature.

[Table 6]

Discussion

Using multidimensional conceptualization of epistemological beliefs in literature identified in Study 1, we submitted an initial pool of items for content validation. This resulted in a set of 29 items designed to measure multiple meaning, social functioning, and multiple reading. The scale was pilot tested with a sample of middle and high school students. We then used a split-sample approach in this study to test the factorial structure of the scale and establish its

initial validity and reliability. Results with Sample 1 confirmed the hypothesized three-factor structure and suggested several survey modifications. A refined version of the LES that was comprised of 16 items showed a reasonable three-factor structure based on such model diagnostics as fit indices, item loadings, and omega reliabilities. Results with Sample 2 confirmed the factorial structure of the 16-item scale and provided evidence of factorial validity and criterion validity. Factorial validity of the LES was established across gender and grade band by demonstrating reasonable fit indices that did not change substantially across the models measuring invariance of factor patterns, loadings, and intercepts. Criterion validity was demonstrated by associations between the scores on the LES and the speed of knowledge acquisition subscale as well as reading habits in the predicted directions. Additionally, as expected, no gender effects on epistemological beliefs in literature were found when we controlled for students' reading habits.

General Discussion

The objective of this multistudy was to develop and provide content, factorial, and criterion validity for a measure of the epistemological beliefs in literature of middle and high school students that capture beliefs about the nature of knowledge and knowing in literary comprehension. The concepts of epistemological beliefs in literary texts were introduced in Study 1 and the results provided an operational definition of epistemological beliefs in literature, suggesting that three factors may be important to measuring epistemological attitudes and beliefs toward literature: multiple meanings, social functioning, and multiple reading. The concept of epistemological beliefs in literature was scrutinized in Study 2 by making modifications to items of the scale. Results in Study 2 confirmed the hypothesized three-factor structure of the scale and provided evidence of reliability and validity of the LES.

An important consideration in developing the LES was to identify the dimensions of epistemological beliefs of readers, which then served as subscales for the scale. Factor analyses revealed three stable factors: one representing the beliefs about the nature of knowledge (i.e., multiple meaning) and two representing the nature of knowing (i.e., social functioning and multiple reading). This distinction between the nature of knowledge and the nature of knowing is consistent with the core structure of individuals' epistemological theories (Hofer, 2006a; Hofer & Pintrich, 1997; Muis, Bendixen, & Haerle, 2006; Schommer-Aikins, 2004). Our findings showed that the two aspects of epistemological beliefs are clearly evident in the domain of literary comprehension.

Our findings showed that the multiple meaning and multiple reading subscales of the LES were associated with the speed of knowledge acquisition subscale, providing support to Hypothesis 1. Associations between the subscales of the LES, students' liking of reading and frequency of reading supported Hypotheses 2 and 3. Particularly, the three subscales were correlated with a measure of how much students like reading; and the multiple reading subscale was correlated with the measure of how much time students spend reading outside of school. Additionally, our findings indicated that the relationship between reading habits and epistemological beliefs were similar for boys and girls when their liking and frequency of reading was accounted for, providing support to the Hypothesis 4. This finding is consistent with recent research (e.g., Conley et al., 2004; Pintrich, 2002) suggesting that gender differences do not emerge when epistemological thinking is defined in terms of separate dimensions of epistemological beliefs.

This is the first study that focused on epistemological beliefs in the domain of literature. Recent studies employing survey methodology have focused on domain-general (e.g., EQ,

Schommer, 1990; Epistemic Scales, Heft & Nasar, 2000), discipline-specific (e.g., DFEBQ, Hofer, 2000; 26-item Epistemological Beliefs Measure, Conley et al., 2004), and topic-specific epistemological beliefs (e.g., TSEBQ, Braten & Stromso, 2009). The vast majority of these measures were developed based on samples of college students and designed for young adults aged 18 and older. As pointed by Hofer and Pintrich (1997), young adolescents' epistemological beliefs have received very little attention. Furthermore, existing domain- and topic-specific questionnaires were hard to apply to the discipline of literature. The challenge that we faced when developing the scale was identifying the dimensions of epistemological beliefs in literature. Muis et al. (2006) noted that, "knowledge in the natural sciences is more structured than knowledge in the social sciences" (p. 37), emphasizing that the ill-structured nature of knowledge in mathematics, science, or history is not typically challenged in teaching practices. In contrast to these domains, in the domain of literature there is a maintained lack of consensus among readers regarding the nature of knowledge and nature of knowing in literary comprehension. Dimensions of epistemological beliefs that have been typically used in the other domains (e.g., mathematics, science, history), particularly, certainty and simplicity of knowledge as well as sources of knowing and justification for knowing in particular, were not directly transferable to the domain of literature. Nevertheless, we were able to conceptualized epistemological beliefs in literary comprehension through three comparable epistemic components (i.e., multiple meanings, multiple reading, and social functioning).

Although the present study expanded knowledge about epistemological beliefs in the domain of literature, a few limitations and directions for future research should be noted. First, the instrument was developed based on the data from middle and high school students; thus it should be used only with young adolescents. Before using the scale with younger or older

participants, evidence of reliability and validity of the scale in those age groups is needed. Second, this study relied on a Likert scale, which brings several pertinent concerns. Although we identified three dimensions of epistemological beliefs in literature, other dimensions that were not identified in this study may exist. Furthermore, a Likert scale may not be able to capture a multidimensional set of interrelated beliefs about knowledge and knowing (Muis et al., 2006). Future research should further develop the present scale by describing other dimensions of beliefs about knowledge and ways of knowing in literary comprehension and exploring the complexity and multifaceted nature of epistemological dimensions. Third, development of a scale is a cumulative and continuous process (DeVilles, 2003). Although we provided evidence of validity and reliability of the scale, future studies are needed to verify the results of this study and further explore its validity. Finally, this study used cross-sectional design, which does not allow for making inferences regarding developmental trends of epistemological beliefs in literary comprehension. Led by the initial work of Perry (1970), questions of change and development of the epistemological beliefs have been central to the research in this field. Recent studies showed that young children's epistemological beliefs about science changed over time (e.g., Conley et al., 2004). Therefore, future studies should focus on cognitive developmental process and examine whether participants change their responses to the scale as a result of an intervention aimed to promote close reading and reasoning in literature.

Conclusions

This research represents an important contribution to the issue of domain specificity within the epistemological beliefs that focus on literary comprehension. Drawing upon a review of personal epistemology research, the multidimensional conceptualization was used to establish a multidimensional framework for understanding students' epistemological approaches to

literature. Results established content validity, factorial validity, criterion validity, and reliability for a quantitative measure of epistemological beliefs in literature that captures readers' epistemological beliefs about literary texts. The LES can be used to measure middle and high school male and female students' epistemological beliefs about literary texts. Ultimately, this instrument can give teachers of English language an idea as to whether their students hold naïve or more expert views of the texts they read to maximize their positive impact on students through teaching practices. However, additional studies are required to establish further validity of the LES.

Acknowledgements

Research funded by the Institute of Education Sciences

Project READI (Reading for Evidence Based Argumentation in Disciplinary Instruction)

Award: R305F100007

References

- Boger-Mehall, S. R. (1996). Cognitive Flexibility Theory: Implications for Teaching and Teacher Education. In B. Robin et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 1996* (pp. 991-993). Chesapeake, VA: AACE.
- Booth, W. C. (1961). *The Rhetoric of Fiction*. Chicago, IL: University of Chicago Press.
- Braten, I., & Stromso, H. I. (2009). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes, 47*, 1-31.
- Braten, I., & Stromso, H.I. (2006). Epistemological beliefs, interests, and gender as predictors of internet-based learning activities. *Computers in Human Behavior, 22*, 1027-1042.
- Chen, F. F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indices for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233-255.
- Cole, N. (1997). *The ETS Gender Study: How Females and Males Perform in Educational Settings*. Princeton, NJ: Educational Testing Service.
- Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology, 29*(2), 186-204.
- Denman, K. (1993). Emily Dickinson's volcanic punctuation. *The Emily Dickinson Journal, 2*, 22-46.
- DeVellies, R. F. (2003). *Scale development: Theory and application*. Thousand Oaks, CA: Sage.
- Estabrook, R. (2012). Factorial invariance: Tools and concepts for strengthening research. In G.

- Tenenbaum, R. C. Eklund, & A. Kamata (Eds.), *Measurement in sport and exercise psychology* (pp. 53-63). Champaign, IL: Human Kinetics.
- Fichman, M., & Cummings, J. N. (2003). Multiple Imputation for Missing Data: Making the Most of What you Know. *Organizational Research Methods, 6*(3), 282-308.
- Foucault, M. (1963). *Death and the labyrinth: The world of Raymond Roussel*. Paris: Gallimard.
- Greene, J. A., Torney-Purta, J., & Azavedo, R. (2010). Empirical evidence regarding relations among a model of epistemic and ontological cognition, academic performance, and educational level. *Journal of Educational Psychology, 102*, 234-255.
- Heft, H., & Nasar, J. L. (2000). Evaluating environmental scenes using dynamic versus static displays. *Environment and Behavior, 32*(3), 301-322.
- Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology, 25*(4), 378-405.
- Hofer, B. K. (2006a). Domain specificity of personal epistemology: Resolved questions, persistent issues, new models. *International Journal of Educational research, 45*, 85-95.
- Hofer, B. K. (2006b). Beliefs about knowledge and knowing: Integrating domain specificity and domain generality: A response to Muis, Bendixen, and Haerle (2006). *Educational Psychology Review, 18*, 67-76.
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*, 88-140.
- Hsieh, P., & Dwyer, F. (2009). The instructional effect of online reading strategies and learning styles on student academic achievement. *Educational Technology and Society, 12*, 36-50.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis:

- Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Kabuto, B. (2009). Parents and children reading and reflecting together: The possibilities of family retrospective miscue analysis. *The Reading Teacher*, 63, 212-221.
- Kline, R. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Laszlo, J. (1999). *Cognition and representation in literature: The psychology of literary narratives*. Budapest, Hungary: Akademiai Kiado.
- Lee, C. D. (2011). Education and the study of literature. *Scientific Study of Literature*, 1, 49-58.
- Maggioni, L., VanSledright, B., & Alexander, P.A. (2010). Walking on the borders: A measure of epistemic cognition in history. *The Journal of Experimental Education*, 77, 187-214.
- Mar, R. A. (2011). The neural basis of social cognition and story comprehension. *Annual Review of Psychology*, 62, 103-134.
- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, 3, 173-192.
- Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications*, 34, 407-428.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.
- May, D.B., & Etkina, E. (2002). College physics students' epistemological self-reflection and its relationship to conceptual learning. *American Journal of Physics*, 70, 1249-1258.
- McDonald, R. P. (1970). Theoretical foundations of principal factor analysis and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, 23, 1-21.

- Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-general and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review, 18*, 3-54.
- Muthen, B., & Muthen, L. (1998-2010). *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthen & Muthen.
- Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks, CA: Sage Publications.
- OECD (2010). *PISA 2009 Results: Learning to Learn – Student Engagement, Strategies and Practices* (Volume III). Retrieved December 14, 2013 from <http://www.oecd.org/pisa/pisaproducts/48852630.pdf>
- Oksa, A., Kalyuga, S., & Chandler, P. (2010). Expertise reversal effect in using explanatory notes for readers of Shakespearean text. *Instructional Science, 38*, 217-236.
- Peischl, S., Stahl, E., & Bromme, R. (2008). Epistemological beliefs and self-regulated learning with hypertext. *Metacognition and Learning, 3*, 17-37.
- Perry, W. G. (1970). *Forms of intellectual and ethical development during the college years*. New York, NY: Holt, Rinehart & Winston.
- Perry, W. G. (1981). Cognitive and ethical growth: The making of meaning. In A. Chickering (Ed.), *The modern American college* (pp. 76–116). San Francisco, CA: Jossey–Bass.
- Pintrich, P. R. (2002). Future challenges and directions for theory and research on personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 389-414). Mahwah, NJ: Erlbaum.
- Pirie, B. (2002). *Teenage boys and high school English*. Portsmouth, NH: Heinemann

- Polit, D. E., Beck, C. T., & Owen, A. V. (2007). Is the CVI an acceptable indicator of content Validity? Appraisal and recommendations. *Research in Nursing & Health, 30*, 459-467.
- Qian, G., & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology, 87*, 282-292.
- Ricoeur, P. (1978). *Interpretation theory: Discourse and the surplus of meaning*. Fort Worth, TX: The Texas Christian University Press.
- Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: conducting a content validity study in social work research. *Social Work Research, 27*(2), 94-104.
- Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist, 39*(1), 19-29.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498-504.
- Schommer, M., Calvert, C., Garigletti, G., & Bajaj, A. (1997). The development of epistemological beliefs among secondary students: A longitudinal study. *Journal of Educational Psychology, 89*, 37-40.
- Schraw, G., & Bruning, R. (1996). Readers' implicit models of reading. *Reading Research Quarterly, 3*, 290-305.
- Simsek, O. F. (2010). Language and the inner: Development of the beliefs about functions of language scale. *European Psychologist, 15*, 68-79.
- Smith, M. W., & Wihelm, J. (2009). Boys and literacy: Complexity and Multiplicity. In L.

- Christenbury, R. Bomer, & P. Smagorinsky (Eds.), *Handbook of Adolescent Literacy Research* (pp. 360-371). New York, NY: Guilford Press.
- Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Knowledge representation, content specification, and the development of skill in situation-specific knowledge assembly: Some constructivist issues as they relate to cognitive flexibility theory and hypertext. *Educational Technology, 31*(9), 22-25.
- Stathopoulou, C., & Vosniadou, S. (2007). Exploring the relationship between physics-related epistemological beliefs and physics understanding. *Contemporary Educational Psychology, 32*, 255-281.
- Stockton, F. R. (1882). The lady, or the tiger? *The Century, 25*, 83–86.
- Stromso, H. I., Braten, I., & Samuelstuen, M. S. (2008). Dimensions of topic-specific epistemological beliefs and predictors of multiple text understanding. *Learning and Instruction, 18*, 513-527.
- Tabachnik, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn and Bacon.
- Trautwein, U., & Ludtke, O. (2007). Predicting global and topic-specific certainty beliefs: Domain-specificity and the role of the academic environment. *British Journal of Educational Psychology, 4*, 907-934.
- Ullman, J. B. (2007). Structural equation modeling. In B. G. Tabachnik & L. S. Fidell (Eds.), *Using multivariate statistics* (5th ed., pp. 676-780). Boston, MA: Allyn and Bacon.
- Vipond, D., & Hunt, R. A. (1984). Point-driven understanding: Pragmatic and cognitive dimensions of literary reading. *Poetics, 13*, 261-277.
- Wilner, A. (2002). Confronting resistance: Sonny's Blues—and mine. *Pedagogy, 2*, 173-196.

Wood, P., & Kardash, C. A. (2002). Critical elements in the design and analysis of studies of epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 231-260). Mahwah, NJ: Erlbaum.

Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. In M. E. Sobel & M. P. Becker (Eds.), *Sociological Methodology* (pp. 165-200). Washington, D.C.: ASA.

Tables

Table 1.

Standardized Factor Loadings and Standard Errors for the Resulting 11-Item 3-Factor Model in Study 1, N = 199

Subscale	Items	Estimate	S.E.
MM	When two people disagree about the meaning of a piece of literature, they could both be right (MM).	.45	.09
MM	Most literature we read has several possible "right" meanings (MM).	.43	.08
MM	Two people can read the same piece of literature and take away different meanings (MM).	.70	.05
MM	Authors can use characters and objects to represent something other than what they seem (RD).	.61	.07
MM	Unexpected events in a piece of literature can be clues to its meaning (RD).	.47	.08
MR	I get nothing out of re-reading a piece of literature again (MR).	.41	.10
MR	I notice new things every time I re-read the same piece of literature (MR).	.79	.10
MR	Every time I re-read a piece of literature, I learn something new (MR).	.75	.11
SF	What I read in literature can help me understand my own life (SF).	.42	.07
SF	Reading literature helps me see the world from other people's point of view (SF).	.75	.06
SF	Reading literature helps me appreciate the way other people see the world (SF).	.79	.06

Note. MM = Multiple Meaning; RD = Rhetorical Devices; MR = Multiple Reading; SF = Social Functioning. Subscales enclosed in parenthesis represent initially hypothesized subscale.

Table 2
Standardized Factor Loadings and Standard Errors for the Final 16-Item Model in Study 2

Subscale	Items	Sample 1 (<i>n</i> = 399)		Sample 2 (<i>n</i> = 399)	
		Estimate	<i>SE</i>	Estimate	<i>SE</i>
MM	A piece of literature can have several possible meanings.	.77	.04	.71	.05
MM	Different people can see different messages in a piece of literature.	.56	.06	.57	.06
MM	One piece of literature can be interpreted in more than one way.	.66	.06	.69	.05
MM	The meaning of a piece of literature is different for each person who reads it.	.59	.05	.60	.05
MM	The same piece of literature can have different meanings to different readers.	.70	.05	.70	.04
RL	The more I read, the more I understand other people.	.76	.03	.79	.03
RL	Reading literature helps me understand why people act the way they do.	.73	.04	.77	.03
RL	Reading literature helps me appreciate the way other people see the world.	.72	.04	.75	.04
RL	Literature helps me value another person's point of view.	.76	.03	.72	.04
RL	I can better understand the behaviors of my friends because I read.	.67	.04	.64	.04
MR	Reading the same piece of literature again is useless.	.85	.03	.90	.02
MR	Re-reading the same piece of literature is meaningless.	.82	.03	.84	.03
MR	Re-reading a piece of literature is a waste of time because I already know the ending.	.85	.03	.83	.03
MR	I do not learn anything new when I read the same piece of literature again.	.68	.04	.73	.04
MR	There is no point in re-reading something you have already read before.	.75	.04	.69	.05
MR	I do not understand why people read their favorite books multiple times.	.52	.05	.59	.05

Note. MM = Multiple Meanings; RL = Relevance to Life; MR = Multiple Readings

Table 3

Model Fit Indices for Tests of Measurement Invariance by Grade Band and Gender

Model	χ^2	<i>df</i>	χ^2/df	RMSEA	CFI	TLI
Grade Band						
Middle School, <i>n</i> = 227	115.4	101	1.14	.025	.987	.984
High School, <i>n</i> = 172	139.5**	101	1.38	.047	.957	.949
Model 1 (Factor Pattern)	255.8**	202	1.27	.037	.972	.967
Model 2 (Factor Loadings)	265.3*	215	1.23	.034	.974	.971
Model 3 (Factor Variances)	319.5***	244	1.31	.039	.961	.962
Gender						
Males, <i>n</i> = 187	116.1	101	1.15	.028	.985	.982
Females, <i>n</i> = 211	102.5	101	1.01	.008	.998	.998
Model 1 (Factor Pattern)	218.3	202	1.08	.020	.991	.990
Model 2 (Factor Loadings)	235.0	215	1.09	.022	.989	.988
Model 3 (Factor Variances)	255.0	244	1.05	.015	.994	.994

Note. RMSEA = root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index.

p* < .05. *p* < .01. ****p* < .001.

Table 4.

Descriptive Statistics for the Measured Variables for Sample 2 by Grade Band and Gender, n = 399

Variables	Range	Total	Middle School	High School	Males	Females
		<i>M (SD)</i>				
Multiple Meanings	1-5	4.24 (0.62)	4.19 (0.63)	4.31 (0.60)	4.15 (0.64)	4.32 (0.58)
Relevance to Life	1-5	3.27 (0.86)	3.38 (0.87)	3.16 (0.84)	3.22 (0.88)	3.34 (0.84)
Multiple Readings*	1-5	3.92 (0.94)	3.95 (0.95)	3.89 (0.93)	3.71 (1.02)	4.12 (0.82)
Speed of Knowledge Acquisition	1-5	2.39 (0.63)	2.49 (0.65)	2.27 (0.58)	2.47 (0.62)	2.33 (0.63)
Like Reading	1-3	2.04 (0.62)	2.10 (0.59)	1.97 (0.64)	1.87 (0.59)	2.19 (0.60)
Frequency of Reading	1-3	1.90 (0.74)	1.98 (0.68)	1.80 (0.81)	1.71 (0.69)	2.07 (0.75)

Note. *All negatively worded; means are based on re-coded values.

Table 5.
Latent Correlations for the Measured Variables for Sample 2, n = 399

Variables	1	2	3	4	5
1. Multiple Meanings	-				
2. Relevance to Life	.39 ^{***}	-			
3. Multiple Readings	.46 ^{***}	.42 ^{***}	-		
4. Speed of Knowledge Acquisition	-.50 ^{***}	-.12	-.49 ^{***}	-	
5. Like Reading	.32 ^{***}	.44 ^{***}	.52 ^{***}	-.28 ^{***}	-
6. Frequency of Reading	.17 ^{**}	.21 ^{***}	.36 ^{***}	-.24 ^{***}	.56 ^{***}

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 6.

Parameter Estimates for Reading Habits and Literature Epistemology Outcomes, Controlling for Gender, n = 399

Variables	Multiple Meanings		Relevance to Life		Multiple Readings	
	Estimate (SE)	t-value	Estimate (SE)	t-value	Estimate (SE)	t-value
Gender	.05 (.05)	0.85	-.03 (.05)	-0.54	.06 (.04)	1.45
Liking of Reading	.29*** (.06)	4.81	.40*** (.06)	7.13	.41*** (.05)	7.70
Reading Frequency	-.01 (.06)	-0.14	.01 (.06)	0.23	.13 (.05)	2.38

Note. Standard errors are in parenthesis. * $p < .05$. ** $p < .01$. *** $p < .001$.