Validation of New Epistemological Scales Related to Inquiry Learning

Project READI Technical Report #6

Carlos Salas, Thomas Griffin,
Jennifer Wiley, M. Anne Britt,
Dylan Blaum
and Patricia Wallace

Citation for this Report: Salas, C., Griffin, T., Wiley, J., Britt, M. A., Blaum D., & Wallace, P. (2016). *Validation of new epistemological scales related to inquiry learning.* READI Technical Report #6. 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.

Validation of new epistemological scales related to inquiry learning

Salas, C., Griffin, T.D., Wiley, J., Britt, M.A., **Blaum**, D., & Wallace, P. (July, 2015). *Validation of New Epistemological Scales Related To Inquiry Learning.* Paper presented at the 25th Annual Meeting of the Society for Text and Discourse, Minneapolis, Mn

(75 word abstract)
Data is presented showing the validity of 2 subscales that can be used to assess epistemological beliefs of particular importance for multiple-document-inquiry tasks. The subscales show similar psychometric properties across grade-levels and for parallel versions worded for science and history. The subscales provided unique predictive utility for the quality of explanations that students constructed for two science topics that differed in difficulty. In particular, the integration subscale reflects a novel dimension of epistemological beliefs.

(1200 word Summary)
Prior work demonstrates that epistemological beliefs affect the *Task Model* that students adopt when engaging in multiple-document inquiry tasks (MDITs, Braten, Britt, Stromso & Rouet, 2011; Griffin, Wiley, Britt & Salas, 2012). This research explores subscales designed to capture two aspects of students' epistemologies that should relate to MDITs. The first subscale corresponds to commonly measured beliefs about the simplicity and certainty of knowledge, but with greater emphasis on the need for explanations with multi-causality. An example reverse-worded item is "The best explanations are those that stick to a single cause". The second subscale was designed to capture beliefs about the importance of integrating and corroborating information across multiple documents and sources. An example item is "Getting information from multiple sources is important when trying find the causes of scientific phenomena." Students endorsed statements using a 1 (strongly disagree) to 6 (strongly agree) scale.

Following multiple iterations, a final set of 13 items (7 integration/corroboration and 6 simplicity/certainty) were developed, with parallel measures for science and history, substituting references to "science" and "phenomena" with "history" and "events". If the two sub-scales reflect the intended psychological constructs then they should have similar psychometric properties across domains and across grade levels. The scales were completed by 723 students in grades 6-12 from Illinois and California public schools during their normal science and history classes. Tests of Measurement Invariance employing MLE yielded significant Chi-Square tests for all two-factor models, but this is common when models with reasonable fits are tested with large samples (Klein, 2010). Thus, models were evaluated using several fit indices and recommend cut-off values (Hu & Bentler, 1999) shown in Table 1 collapsing across grade. The top portion of Table 1 shows that for both History and Science each index meets the cut-off values, suggesting that the two-factor structure is an overall good fit for both disciplines. The bottom portion of Table1 tests two-factor models with increasing restrictions of invariance across disciplines. The indices meet the cut-offs for all models, but the lower AIC and BIC values for Model 3 suggests that it has the best fit/complexity trade-off. Thus, the factor structure, loadings, and intercepts of the scales are similar for history and science, but measurement error differs. Though not shown, similar results were observed when collapsing across discipline and testing invariance of a two-factor model between grade levels. The two sub-scales assess similar orthogonal clusters of beliefs about information integration and simplicity/certainty of explanations, regardless whether the discipline is science or history. Also, the scales appear to measure similar constructs for both middle and high-school students.

A subset of 122 students completed the science version of the epistemology scales and then an MDIT on one of two science topics (coral bleaching or skin cancer). These units were created by the Project READi Science Design Team (Britt, Blaum, Wallace, Ko, & Goldman, 2014). The tasks presented students with 5 documents (including texts, graphs, and images) related to the topic and required them to write an essay to explain how and why "…rates of skin cancer differ around the globe" or "…coral bleaching rates vary at different times." Students had to integrate information across documents in order to construct a complete causal explanation.

Essays were coded for 4 dimensions related to the completeness and coherence of their causal explanations, including the number of key causal concepts/factors, number of causal connections among concepts, number of unique causal chains, and length of longest causal chain (IRRs .67 to .95). Relationships with epistemology subscales were highly similar for all four codes, and they were highly related to each other (*r*s = .66 to .89), so a single latent factor score was computed for *explanation quality*, which accounted for 81% of the variance in the 4 codes.

Table 2 shows the results of a Hierarchical Regression predicting explanation quality. Step 1 shows that both epistemology subscales uniquely predicted explanation quality. Also, explanation quality was higher for the topic of skin cancer than for coral bleaching. These variables accounted for 19% of the variance in explanation quality. Step 2 shows that grade level and essay length (number of sentences) were significant predictors of essay quality, and the total model accounted for 41% of variance. After accounting for these, the effect for the integration subscale was unchanged, but the effect for the simplicity/certainty subscale reduced to marginal levels. Simple correlations suggested that this was due to essay length being correlated with simplicity/certainty ($r = .17$, $p = .03$), but not with integration beliefs ($r = .06$, $p = .27$). Epistemology and topic did not interact, thus indicating similar predictive relationships for both topics.

The results support the validity of two subscales that can be used to assess epistemological beliefs of particular importance for MDITs, where the goal for the student is to construct multi-causal explanations. The subscales provided unique predictive utility for the quality of explanations that students construct for two topics that differed in difficulty (better essays for skin cancer). This suggests that scales are not merely tapping into general motivation or skills that would account for the same shared variance in learning. In addition, the differential overlap with essay length suggests different pathways by which the different beliefs relate to MDIT performance. Although tested here as predictive measures, these subscales could be examined as potential moderators of instruction, or as outcome variables capturing changes in epistemology that might result from authentic inquiry tasks.

References

Bråten, I., Britt, M.A., Strømsø, H.I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Towards an integrated model. *Educational Psychologist, 46,* 48-70.

Britt, M. A., Blaum, D., Wallace, P., Ko, M.-L., & Goldman, S. R. (2014). Evidence-based argumentation inquiry modules on coral bleaching and skin cancer. *Modules created as part of Project READi Science Design Team*.

Griffin, T.D., Wiley, J., Britt, M.A. & Salas, C. (2012).The role of CLEAR thinking in learning science from multiple-document inquiry tasks. *International Electronic Journal of Elementary Education*, 5, 63-78.

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

Table 1. *Model Fit and Measurement Invariance of Information Integration and Simplicity*

Recommended cut-off values	Restrictions	CFI/TLI (>.90)	AIC	BIC	RMSEA (< .06)	SRMR (<.08)
History		.91 (.89)			.06	.06
Science		.94 (.93)			.05	.06
Model 1 (Configural Invariance)	Equal factor structure	.93 (.91)	25648.23	25987.40	.05	.06
Model 2 (Metric Invariance)	1+ Equal loadings	.92 (.91)	25645.35	25938.69	.05	.06
Model 3 (Scalar Invariance)	2+Equal intercepts	.92 (.91)	25643.51	25891.02	.05	.06
Model 4 (Strict Invariance)	3+Equal meas. errors	.90 (.91)	25651.01	25843.5 2	.05	.07

Note. History *N* = 330; Science *N* = 393

Table 2. Unique Variance in Learning Predicted by *Beliefs across Disciplines* Epistemology Sub-Scales

Criterion variable		*F value*		*t v*
Step 1	.19	9.21*		
Integration			.28	3
Simplicity-Certainty			.20	2
Topic			.23	2
Step 2	.22	8.52*		
Integration			.31	2
Simplicity-Certainty			.18	
Topic			.18	2
Essay Length			.28	3
Grade level			.26	3

Note. Inc. R^2 = increment in explained variancer; standardized regression coefficient; ** = *p* < .05, * = *p* < .1