ARTICLE
Use of Structural Assessment of Knowledge for Outcomes Assessment in the Neuroscience Classroom

Jennifer L. Stevenson, Samir Shah, & Joel P. Bish
Psychology Department and Neuroscience Program, Ursinus College, Collegeville, PA 19426

Outcomes assessment of undergraduate neuroscience curricula should assess the ability to think integratively about basic neuroscience concepts based on two of the core competencies established by the Faculty for Undergraduate Neuroscience. The current study investigated whether the structural assessment of knowledge (SAK) approach, which evaluates the organization of an individual's knowledge structures, is effective for demonstrating learning of basic neuroscience concepts. Students in an introductory psychology course (n = 29), an introductory neuroscience course (n = 19), or an advanced behavioral neuroscience course (n = 15) completed SAK before and after learning gross brain anatomy and neuronal physiology. All students showed improvements in their SAK after short-term dissemination for gross brain anatomy, but not for neuronal physiology, concepts. Therefore, research is needed to determine whether the effectiveness of SAK in outcomes assessment depends on the content or teaching style. Additional research using SAK should also explore effectiveness for learning over longer time frames and correlations with student performance in the course. However, the results suggest SAK is a promising technique for outcomes assessment of undergraduate neuroscience curricula.

Key words: assessment; concept mapping; knowledge representation; outcomes; structural knowledge

A workshop hosted by the Faculty for Undergraduate Neuroscience established a set of core competencies that will help with the design and assessment of neuroscience curricula (Kerchner et al., 2012). Of the six established competencies, two are focused on for the current study. First, successful neuroscience undergraduates are expected to demonstrate basic knowledge in Neuroscience, Biology, Chemistry, and Psychology. On its own, this competency could rely exclusively on rote memorization, which tends to promote short-term, low-level cognitive thinking (Momsen et al., 2010). Therefore, a second competency recommending the ability to think critically and integratively is also included among others, suggesting that high-level cognitive thinking about basic scientific information is critical at the undergraduate level (Kerchner et al., 2012). The establishment of these competencies requires designing modes of evaluation that will allow educators to assess the ability of students to apply high-level thinking to the acquisition of basic knowledge structures. While it is possible to use more typical assessment tools, such as pretest/posttest scores for quizzes or exams, these measures tend to be reliant on the types of questions being assessed and whether they truly demonstrate knowledge change may be unreliable.

Previous research has demonstrated that the number and organization of knowledge structures corresponds directly with expert-like status (Davis and Yi, 2004). The phrase structural assessment of knowledge (SAK) has been coined to describe the process of evaluating the organization of an individual's knowledge structures pertaining to a certain domain, e.g., neuroscience (Trumpower et al., 2010).

While similar to typical concept mapping approaches, SAK approaches lend themselves to more efficient and perhaps more unbiased assessment of knowledge acquisition. Instead of constructing a concept map individually using a pre-constructed blank map, SAK approaches require participants to make a judgement on the relatedness between two concepts. SAK approaches are also not prone to aesthetic biases, such as the need to minimize links or contain no un-linked terms (Trumpower et al., 2010).

Procedurally, the SAK approach can be understood as three interrelated phases: knowledge elicitation, knowledge representation, and knowledge evaluation. This paper considers the SAK approach using Pathfinder network analysis (Schvaneveldt, 1990), which is freely available and runs on any computer using Java Script (http://interlinkinc.net).

The knowledge elicitation phase consists of a participant rating the relatedness of two concepts from a list of key terms. Each possible pairing is rated. The number of concepts (n) is directly related to the number of concept pairs, $n(n-1)/2$. The predictive validity of SAK (with exam performance) increases with the number of concepts but this must be balanced with the time to rate all possible pairs (Goldsmith et al., 1991).

From these ratings, the knowledge representation is constructed. The Pathfinder network scaling algorithm transforms the ratings into both a mathematical and visual representation of the individual's knowledge organization. The representation consists of nodes, one for each concept, and links, representing the relatedness between pairs of nodes. The usefulness of the Pathfinder SAK method expands for both hierarchical and nonhierarchical relationships. Additionally, the algorithm allows for group averaging of knowledge structures.

Finally, the Pathfinder program allows for knowledge evaluation, through the comparison between networks. The program computes a number of metrics of similarity between networks (e.g., number of links in common, similarity), either at the group level or the individual level (see Schvaneveldt, 1990 for details on the Pathfinder method). Previous studies have demonstrated that these metrics predict the progression of learning and decision-making with a high degree of accuracy (Goldsmith et al., 1991; Davis and Yi, 2004).

Trumpower and Vanapalli (in press) provides a rich and extensive review of a variety of uses for the SAK approach including using SAK as an instructional tool and as an assessment technique as in the current study. However, as an example, Goldsmith et al. (1991) assessed SAK for research methods and statistics after a 16-week undergraduate-level course. Pathfinder's metric, number of links in common with the instructor, was a strong predictor of exam performance. In another study, Wilson (1994) investigated SAK for chemical equilibrium in high achieving and low achieving senior high school physics students. The high achieving group had fewer links than the low achieving group suggesting greater specificity in relationships and hierarchical organization.

Evidence of the validity of this Pathfinder SAK approach has demonstrated that the degree of similarity between a student and referent knowledge structure (e.g., expert) is correlated with common educational achievement indicators including course grades and exams (Goldsmith, et al., 1991). More importantly, the similarity scores between learners and experts show a stronger correlation with higher-order educational outcomes (e.g., essay performance and complex problem solving) compared to lower-order outcomes, such as multiple choice responses (d'Appolonia et al., 2004).

Past research has considered use of SAK in multiple domains such as accounting (Curtis and Davis, 2003; Rose et al., 2007), chemistry (Wilson, 1994), computer programming (Trumpower et al., 2010), mathematics (Gomez et al., 1996; Davis et al., 2003), nursing (Azzarello, 2007), physics (Chen and Kuljis, 2003; Trumpower and Sarwar, 2010), and research methods (Goldsmith et al., 1991). While SAK has been employed in cognitive neuroscience research, such as semantic networks in patients with amnesia (Chan et al., 1995), Alzheimer's disease (Chan et al., 2001; Aronoff et al., 2006; Razani et al., 2010) and frontal lobe lesions (Sylvester and Shimamura, 2002), to date no study has focused on SAK applied to neuroscience concepts.

Therefore, the current study investigated the acquisition of neuroscience concepts covered in introductory and advanced undergraduate courses using SAK. It was hypothesized that SAK would be effective in demonstrating knowledge acquisition for basic neuroscience information that was disseminated over a short period of time. More explicitly, it was expected that participants' conceptual maps, elicited by the SAK approach for basic neuroscience material, would look more like a prototypical expert's map after instruction than before.

MATERIALS AND METHODS

Participants

Sixty-three students participated in this study. Students were recruited based on their enrollment in either an introductory psychology course (n = 29), introductory neuroscience course (n = 19) or an advanced behavioral neuroscience course (n = 15) in Fall 2014 or Spring 2015. These courses were selected because they cover gross brain anatomy (e.g., cerebellum) and neuronal physiology (e.g., depolarization). Students received extra credit for their participation.

Materials and Procedure

JLS and JPB constructed a set of 15 concepts related to gross brain anatomy (i.e., amygdala, basal ganglia, cerebellum, cerebral cortex, cerebrospinal fluid, corpus callosum, forebrain, hindbrain, hippocampus, hypothalamus, medulla, midbrain, meninges, pons, and thalamus) and a set of 15 concepts related to neuronal physiology (i.e., action potential, axon hillock, axon terminal, dendrite, depolarization, ion channel, neuron, oligodendrocyte, receptor, resting potential, reuptake, salutatory conduction, summation, synapse, and threshold) that were covered in all three courses.

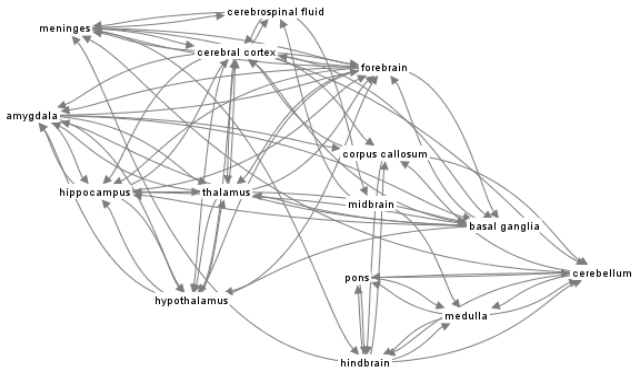
Two unique Pathfinder exercises were created (one for gross brain anatomy concepts and one for neuronal physiology). Each exercise contained 210 concept pairs (i.e., all possible pairs for a set of 15 concepts, e.g., amygdala-basal ganglia and basal ganglia-amygdala). The full set of concept pairs for each exercise is provided in a supplementary file. For each concept pair, participants were asked to assign a rating based on their relatedness on a scale from 1 *not related* to 7 *synonymous*. Participants completed each Pathfinder exercise at two separate time points (prior to learning concepts and after learning concepts for an exam). One expert (JPB) also completed the Pathfinder exercise for each set of concepts for comparison.

Participants' ratings were recoded using a 5-point Likert scale ranging from 1 *not related/slightly related* (original rating 1), 2 *somewhat related* (original rating 2 or 3), 3 *moderately related* (original rating 4 or 5), 4 *most related* (original rating 6), or 5 *synonymous* (original rating 7). The Pathfinder network scaling algorithm was used to derive network relatedness ratings using JTarget and JPathfinder (open source software available from Interlink Inc. at interlinkinc.net). JTarget was used to convert participants' raw ratings into matrices which were then imported into JPathfinder. JPathfinder compared individual participants' networks (both pre- and post-learning) to the expert's networks for gross brain anatomy and neuronal physiology. Figure 1 displays the pre- and post-learning networks of gross brain anatomy concepts in the introductory neuroscience course and the expert network for comparison.

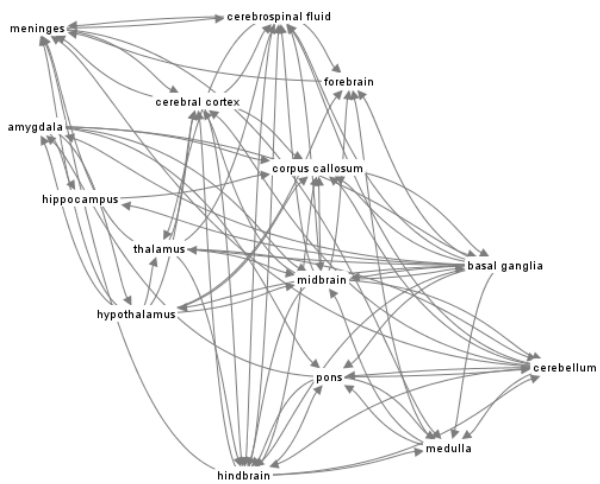
Data Analysis

The number of links (i.e., connections between nodes) in common between the participant's network and the expert comparison network was computed. The maximum score for the links in common measure is the number of links in the reference knowledge structure. In this study, the expert network comprised 83 links for gross brain anatomy and 60 links for neuronal physiology.

Introductory Neuroscience Course Pre-Learning



Introductory Neuroscience Course Post-Learning



Expert

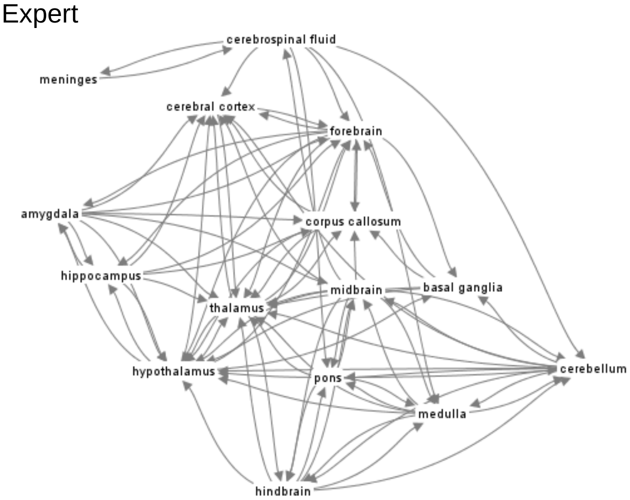


Figure 1. Pathfinder maps for the students in the introductory neuroscience course pre- (top) and post-learning (middle) and expert (bottom) for gross brain anatomy concepts.

A more complex metric of similarity between the participant's network and the expert comparison network was also computed: C / (Links1 + Links2 – C) where C refers to the number of links in common, Links1 refers to the total number of links in the participant network and Links2 refers to the total number of links in the expert network. The similarity measure ranges from 0 *no similarity in networks* to 1 *identical networks*.

Two 3 x 2 x 2 mixed-design Analysis of Variance (ANOVA) were conducted to investigate the effects of course (introductory psychology, introductory neuroscience, or advanced neuroscience), time (pre- or post-learning), and material type (gross brain anatomy or neuronal physiology) on number of links in common with the expert network and overall similarity with the expert network. Tukey HSD pairwise comparisons were used to follow-up on significant main effects of course and dependent-group *t*-tests were used to follow-up on significant interactions. All analyses used $p < 0.05$ for statistical significance.

RESULTS

For the number of links in common (see Figure 2 and Table 1), there were significant main effects of course ($F(2, 60) = 9.37$, $p < 0.001$, $\eta^2_{partial} = 0.24$), time ($F(1, 60) = 11.07$, $p = 0.002$, $\eta^2_{partial} = 0.16$), and material ($F(1, 60) = 54.23$, $p < 0.001$, $\eta^2_{partial} = 0.48$). There was also a significant interaction between time and material ($F(1, 60) = 5.54$, $p = 0.02$, $\eta^2_{partial} = 0.08$). No other interactions were significant (all *p*s > 0.05).

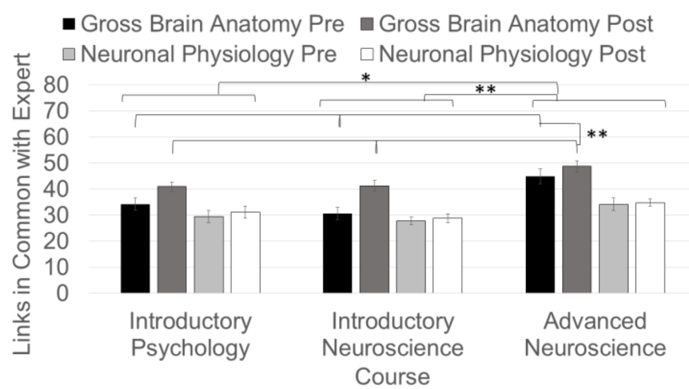


Figure 2. Number of links in common with expert by course, time, and material. Maximum number of links in common with the expert was 83 for gross brain anatomy and 60 for neuronal physiology. Error bars represent ±1 *SE*. *$p \leq 0.01$, **$p \leq 0.001$

Course	Gross Brain Anatomy		Neuronal Physiology	
	Pre	Post	Pre	Post
Intro Psychology: *M* (*SE*)	34.24 (2.33)	40.93 (1.72)	29.48 (2.28)	31.14 (2.27)
Intro Neuroscience: *M* (*SE*)	30.63 (2.48)	41.26 (2.04)	27.84 (1.50)	28.84 (1.63)
Adv Neuroscience: *M* (*SE*)	44.93 (2.86)	48.73 (2.18)	34.13 (2.39)	34.80 (1.28)

Table 1. Descriptive statistics for number of links in common with expert by course, time, and material. The expert network comprised 83 links for gross brain anatomy and 60 links for neuronal physiology.

A similar pattern of results was found for the overall similarity in networks as shown in Figure 3 and Table 2. There were significant main effects of course ($F(2, 60) = 6.43$, $p = 0.003$, $\eta^2_{partial} = 0.18$), time ($F(1, 60) = 7.35$, $p = 0.009$, $\eta^2_{partial} = 0.11$), and material ($F(1, 60) = 79.30$, $p <$

0.001, $\eta^2_{partial}$ = 0.57). There was also a significant interaction between time and material (F(1, 60) = 5.87, p = 0.02, $\eta^2_{partial}$ = 0.09). No other interactions were significant (all ps > 0.05).

Students in the advanced neuroscience course had significantly greater number of links in common and greater similarity with the expert compared with students in the introductory psychology (common links: p = 0.002; similarity: p = 0.02) or introductory neuroscience (common links: p < .001; similarity: p = 0.003) courses. Students in the two introductory courses did not differ from one another (common links: p = 0.56; similarity: p = 0.62).

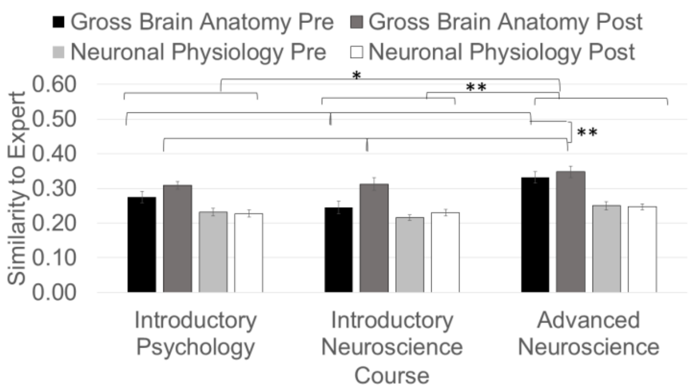


Figure 3. Overall similarity with network expert by course, time, and material. Similarity is defined as C / (Links1 + Links2 – C) where C refers to the number of links in common, Links1 refers to the total number of links in the participant network and Links2 refers to the total number of links in the expert network. Error bars represent ±1 *SE*. *$p \leq 0.05$, **$p \leq 0.01$.

Course	Gross Brain Anatomy		Neuronal Physiology	
	Pre	Post	Pre	Post
Intro Psychology: *M* (*SE*)	0.276 (0.017)	0.309 (0.011)	0.232 (0.011)	0.227 (0.010)
Intro Neuroscience: *M* (*SE*)	0.245 (0.018)	0.313 (0.018)	0.216 (0.008)	0.230 (0.009)
Adv Neuroscience: *M* (*SE*)	0.332 (0.016)	0.347 (0.017)	0.250 (0.012)	0.247 (0.009)

Table 2. Descriptive statistics for overall similarity with expert by course, time, and material.

Students made significant gains in number of links in common and similarity with the expert in knowledge related to gross brain anatomy (common links: t(62) = -3.94, p < 0.001, d = -0.50; similarity: t(62) = -3.14, p = 0.003, d = -0.40). However, no gains were made in the number of links in common or similarity with the expert for knowledge related to neuronal physiology (common links: t(62) = -0.80, p = 0.43, d = -0.10; similarity: t(62) = -0.16, p = 0.87, d = -0.02).

DISCUSSION

Structural assessment of knowledge (SAK) effectively demonstrated knowledge acquisition as measured by number of links in common with the expert network and overall similarity with the expert network. Moreover, students in both introductory and advanced neuroscience courses showed improvements in their SAK after short-term dissemination. However, significant changes were observed for gross brain anatomy, but not neuronal physiology, concepts. More specifically, students achieved 0.31 to 0.35 similarity scores with the expert network post-learning of gross anatomy. This is comparable to similarity ratings post-learning in other studies; for example, 0.30 similarity between medical students and faculty for pulmonary concepts (McGaghie et al., 1996), 0.24 similarity between nursing undergraduate students and their instructors for community nursing concepts (Azzarello, 2007), and 0.38 similarity between high and low achieving physics high school students for chemical equilibrium (Wilson, 1994). This provides evidence that SAK may be useful for both assessing knowledge of basic neuroscience concepts and the ability to think integratively as part of an outcomes assessment of a neuroscience curricula.

Given changes in SAK were not observed with neuronal physiology concepts, the effectiveness of SAK for assessment of knowledge acquisition may depend on the content and/or teaching style. While SAK has been successfully used in multiple domains, it is possible that some topics lend themselves better to concept mapping than others. Gross brain anatomy is likely taught in a hierarchical manner – the brain can be subdivided into the forebrain, midbrain, and hindbrain and the hindbrain is further subdivided into the cerebellum, pons, and medulla – which may facilitate concept mapping. In contrast, neuronal physiology is likely taught in a dynamic manner focusing on the temporal order of events – if the neuronal membrane moves from resting potential to threshold, an action potential is initiated. It is unclear whether the topic itself (neuronal physiology) or the teaching style (dynamic) hinders concept mapping. Therefore, future studies should manipulate teaching style, for example by incorporating concept mapping during learning (i.e., SAK as a pedagogical activity; see Trumpower and Vanapalli, in press for a review), to determine if teaching style affects SAK. Additionally, future research should consider other topics such as neuroscience methods and techniques to examine whether SAK depends on content.

The current study focused on basic neuroscience information that was disseminated over a short period of time. As such, it was impractical to directly relate students' performance in the class with changes in their SAK (i.e., exams covered more material). In the future, short quizzes could be administered in conjunction with the post-learning Pathfinder exercise. In addition, research could consider topics covered over a greater period of time. For example, gross brain anatomy could be expanded to incorporate concepts related to both structure and function (e.g., hippocampus, prefrontal cortex, memory, attention) that are taught as part of a unit in a course (e.g., an introductory neuroscience course with a unit on higher-order cognition) or over the course of an entire semester (e.g., an advanced course in cognitive neuroscience).

Relatedly, the scope of the current study explored the effectiveness of SAK for outcomes assessment, or summative assessment. However, as Trumpower and Vanapalli (in press) explain SAK can be used as a pedagogical activity, for formative assessment or for summative assessment. A larger-scale study could incorporate SAK during learning as either a pedagogical activity (e.g., in a manipulation of teaching style) as well as SAK for formative assessment to help gauge students' progress towards learning goals.

These learning goals may be summarized across the whole network as in the current study or expanded into much more specific network analysis of individual student networks. Additionally, this methodology allows for the specific analysis of linkages between specific nodes. If a class has missing links between nodes that are deemed critical by the faculty, the curriculum can be adjusted to focus on those specific relationships.

Finally, students' networks were compared to a single individual's network. While some individuals may serve as excellent experts in constructing the expert SAK (Acton et al., 1994) and it is even possible to reliably assess concept maps without an expert (McClure et al., 1999), in general, multiple experts in the field should be used (Acton et al., 1994). However, given the interdisciplinary nature of neuroscience, research is needed on how to select the best experts. Furthermore, it seems likely that experts should come from multiple disciplines including psychology and biology. As this was not within the scope of the current study, JPB was selected as the expert due to his extensive experience teaching neuroscience and coordinating the neuroscience program.

As pointed out by a reviewer, it is also important to consider, that depending on the assessment goal, different experts may be needed. If the goal is outcomes assessment of an individual course, either the course instructor or a set of faculty members from the department would be appropriate in constructing the expert SAK. However, if the goal is outcomes assessment of an entire program, a representative set of faculty members outside of the college (e.g., members of Faculty for Undergraduate Neuroscience) would be more desirable in constructing the expert SAK. It is likely that the inclusion of multiple experts would alter the total number of links, and potentially the location of those links within the referent network, which will alter the results. Acton et al. (1994) point out that a Pathfinder network based on the average of a number of "experts" will provide a more valid referent network compared to a single network. As mentioned, this would be particularly interesting depending upon the objectives of the assessment goal.

Despite these limitations, this is the first study to demonstrate the effectiveness of SAK for basic neuroscience concepts. Overall, these results suggest SAK is a promising method that could be used for outcomes assessment in the neuroscience classroom and merits further investigation.

REFERENCES

Acton WH, Johnson PJ, Goldsmith TE (1994) Structural knowledge assessment: Comparison of referent structures. J Educ Psychol 86:303-311.

Aronoff JM, Gonnerman LM, Almor A, Arunachalam S, Kempler D, Andersen ES (2006) Information content versus relational knowledge: semantic deficits in patients with Alzheimer's disease. Neuropsychologia 44:21-35.

Azzarello J (2007) Use of the Pathfinder scaling algorithm to measure students' structural knowledge of community health nursing. J Nurs Educ 46:313-318.

Chan AS, Butters N, Salmon DP, Johnson SA, Paulsen JS, Swenson MR (1995) Comparison of the semantic networks in patients with dementia and amnesia. Neuropsychology 9:177-186.

Chan AS, Salmon DP, De La Pena, J (2001) Abnormal semantic networks for "animals" but not "tools" in patients with Alzheimer's disease. Cortex 37:197-217.

Chen C, Kuljis J (2003) The rising landscape: a visual exploration of superstring revolutions in physics. J Am Soc Inf Sci Technol 54:435-446.

Curtis MB, Davis MA (2003) Assessing knowledge structure in accounting education: an application of Pathfinder associative networks. J Account Educ 21:185-195.

d'Appolonia ST, Charles ES, Boyd GM (2004) Acquisition of complex systemic thinking: Mental models of evolution. Educ Res Eval 10:499-521.

Davis MA, Curtis MB, Tschetter JD (2003) Evaluating cognitive outcomes: validity and utility of structural knowledge assessment. J Bus Psychol 18:191-206.

Davis FD, Yi MY (2004) Improving computer skill training: behavior modeling, symbolic mental rehearsal, and the role of knowledge structures. J Appl Psychol 89:509-523.

Goldsmith TE, Johnson PJ, Acton WH (1991) Assessing structural knowledge. J Educ Psychol 83:88-96.

Gomez RL, Hadfield OD, Housner LD (1996) Conceptual maps and simulated teaching episodes as indicators of competence in teaching elementary mathematics. J Educ Psychol 88:572-585.

Kerchner M, Hardwick JC, Thornton JE (2012) Identifying and using 'core competencies' to help design and assess undergraduate neuroscience curricula. J Undergrad Neurosci Educ 11:A27-A37.

McClure JR, Sonak B, Suen HK (1999) Concept map assessment of classroom learning: reliability, validity, and logistical practicality. J Res Sci Teach 36:475-492.

McGaghie WC, Boerger RL, McCrimmon DR, Ravitch MM (1996) Learning pulmonary physiology: comparison of student and faculty knowledge structures. Acad Med 71:S13-15.

Momsen JL, Long TM, Wyse SA, Ebert-May D (2010) Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. CBE Life Sci Educ 9:435-440.

Razani J, Chan A, Nordin S, Murphy C (2010) Semantic networks for odors and colors in Alzheimer's disease. Neuropsychology 24:291-299.

Rose JM, Rose AM, McKay, B (2007) Measurement of knowledge structures, experience and decision aid use. Int J Account Inf Syst 8:117-137.

Schaneveldt R (1990) Pathfinder associative networks: studies in knowledge organization. Norword, NJ: Ablex.

Sylvester CYC, Shimamura AP (2002) Evidence for intact semantic representations in patients with frontal lobe lesions. Neuropsychology 16:197-207.

Trumpower DL, Sarwar GS (2010) Effectiveness of structural feedback provided by Pathfinder networks. J Educ Comput Res 43:7-24.

Trumpower DL, Sharara H, Goldsmith TE (2010) Specificity of structural assessment of knowledge. J Technol Learn Assess 8(5):1-32.

Trumpower DL, Vanapalli AS (in press) Structural assessment of knowledge as, of, and for learning (Spector MJ, Lockee BB,

Childress MD, eds). Switzerland: Springer International Publishing.
Wilson JM (1994) Network representations of knowledge about chemical equilibrium: variations with achievement. J Res Sci Teach 31:1133-1147.

Received July 08, 2016; revised August 29, 2016; accepted September 02, 2016.

This work was supported by Ursinus College Student Achievement in Research and Creativity funding to SS. SS worked on this project as an undergraduate student under JLS and JPB's supervision. The authors thank Daniel Brogan, Marisa Gretz, Angelo Montenegro, and Samuel Talerico for their assistance with data collection and data entry.

Address correspondence to: Dr. Jennifer L. Stevenson, Psychology Department, 601 East Main Street, Ursinus College, Collegeville, PA 19426. Email: jstevenson@ursinus.edu