

Applications of Pathfinder Network scaling for identifying an optimal use of first language for second language science reading comprehension

Kyung Kim¹  · Roy B. Clariana²

Published online: 11 June 2018

© Association for Educational Communications and Technology 2018

Abstract Previous research has shown that the knowledge structure (KS) complexity of a first language (L1) under certain conditions can strongly influence the KS complexity established in a second language (L2), and then this more complex L2 KS reciprocally influences L2 text comprehension. This present experimental investigation seeks to identify the unique contributions of mapping and writing in L1 (Korean) tasks to support L2 (English) science text comprehension using Pathfinder Network scaling, a graph-theoretic cognitive science approach. Native Korean low proficiency English language learners ($n = 245$) read a 708-word English (L2) science lesson text, completed one of seven treatment conditions, and then completed a comprehension posttest. The seven conditions consisted of three experimental conditions that required different L1 tasks including: L1 mapping alone, L1 writing alone, or both L1 mapping and writing; and four control conditions that did not receive any L1 treatment: L2 mapping alone, L2 writing alone, both L2 mapping and writing, or reading only. All of the maps and writing artifacts were converted into Pathfinder Networks that were compared to an expert's referent network. Results show that requiring L1 lesson tasks relatively increases L2 KS complexity and concomitant comprehension posttest performance. In order of effectiveness, combined L1 mapping and writing was most effective for posttest comprehension, then L1 writing, and least effective is L1 mapping alone. These findings confirm and extend the earlier findings that the inherent L1 KS complexity can strongly influence L2 KS complexity. Educationally, requiring L1 tasks, especially in text translation, likely engenders richer L2 structure that supports higher-order understanding of the text. Also, these findings further validate this technology-based approach for measuring KS contained in bilingual learners' productions.

✉ Kyung Kim
kkim2@niu.edu

¹ Educational Technology Research and Assessment, Northern Illinois University, Dekalb, USA

² Learning, Design, and Technology, The Pennsylvania State University, University Park, USA

Keywords Pathfinder Network · Knowledge structure · Concept map and writing · Second language reading

Introduction

According to the U.S. Census Bureau's American Community Survey (Ryan 2013), one in five U.S. residents now speaks a language other than English at home. What instructional design advice is there for these students in our classrooms who must learn in English? Regarding learning from texts, the ability to read and comprehend is a critical learning skill not only in a native language, but also in a second language. Accordingly, there have been several research-based efforts to help students comprehend second language texts (see for review, Karim 2010), but the value of intentionally using first language (L1) in second language (L2) reading has been a continuous source of interest and debate (see for debate of L1 use, van Hell and Kroll 2013). This investigation considers the influence of knowledge structure (KS) of the L1 on L2 KS during and after reading in the L2 by applying a Pathfinder Network scaling, a graph-theoretic computational approach, to quantitatively evaluate comprehension as a function of changes in L1 and L2 KS.

Although this investigation considers Korean and English, this line of research is really not about languages, but rather, it is about structure. Behavioral studies (e.g., ranging from reaction time studies up to problem-solving expertise, e.g., Tse and Altarriba 2012) and neural network model descriptions (e.g., Zhao and Li 2013) of language artifacts support the premise that an individual's lexicon is structured and it is this structure that allows for "thinking". In this present investigation, describing how KS in one language influences or aligns with that of the second language could provide direct empirical information regarding the role of KS in second language reading comprehension.

The influence of first language in second language reading

The research base considering the use of L1 during L2 tasks cuts across several domains, including psycholinguistics (e.g., van Hell and Kroll 2013), neuroscience (e.g., Chee et al. 2001), and even socio-cultural spheres (e.g., Anton and DeCamilla 1998; Brooks and Donato 1994). L1-in-L2 studies are carried out for a number of different reasons and with varying research goals. Many of the studies have found that learners completing L2 tasks tend to spontaneously but idiosyncratically use their L1 and L2 interactively in order to work through a particular problem. In consonance with these research findings, other studies have reported that increased L2 proficiency may constrain the amount of L1 use, thereby exerting varied effects on learners' L2 production (Jones and Tetroe 1987; Pennington and So 1993; Qi, 1998; Roca de Larios et al. 1999; Wang and Wen 2002).

Logically, if L2 is sufficient for a specific task, then L1 is not needed; and so there is a general tendency to rely mainly on L2 as L2 proficiency increases. For example, Cumming (1990) found that highly proficient bilingual learners used their L1 sparingly and selectively—for initiating an idea, for developing a thought, for verifying the meaning of a word, and to compensate for working memory overload in L2 (see for detail, Qi 1998). In comparison, other studies have concluded the opposite, that lower L2 proficient learners used their L1 far more generally than the higher L2 proficient learners, for example, to search out and assess appropriate wording, compare cross-linguistic equivalents, but less frequently considered linguistic choices in the L2 (see for detail, Wang et al. 2002). Taken together, we can infer that the L2 proficiency level predicts L1 spontaneous use, with a

selective or no reliance on L1 for highly proficient L2 learners but a broad and general reliance on L1 for low proficient L2 learners.

Moreover, Woodall (2002) complicated this possible relation of L2 proficiency with L1 use even further by including the fundamental linguistic properties of languages as an additional independent variable in the study. He found that overall, low L2 proficient bilinguals used their L1 more than high L2 proficient learners, but this was influenced by task difficulty and language group (e.g., with non-cognate languages such as Korean/English versus cognate languages such as Spanish/English that share common etymological origins). Spontaneous L1 use when the L1 and L2 are cognate languages was related to higher quality L2 production, while spontaneous L1 use between L1 and L2 non-cognate languages was related to lower quality L2 production.

Although the recent studies cited above provide insight into the possible role that L1 plays when completing demanding L2 tasks, overall, the value of L1 use for learning in an L2 setting has remained a controversial issue. The reasons for spontaneous versus required L1 use, and which cognitive activities are carried out in the L1 also remain somewhat unclear. As mentioned above, L1 may be recruited to help solve L2 linguistic or lower-order problems (cognitive level, see Wang 2003), but L1 may also be recruited for meta-level activities such as planning or to prevent cognitive overload (metacognitive level, see Woodall 2002). This begs the question, "What characterizes a 'demanding' task in L2 that will likely elicit spontaneous L1 support?"

In summary, bilinguals with different L2 proficiency levels may use their L1 in different ways (Woodall 2002) or for different cognitive processing while completing L2 tasks (Cumming 1990). Thus, a broad conclusion is that L1 use for the completion of L2 tasks, whether spontaneously recruited or alternately required by the learning task, tends to improve L2 production, but not in all situations and not for all categories of bilingual learners (Cohen and Brooks-Carson 2001; Kobayashi and Rinnert 1992; Uzawa and Cumming 1989). The somewhat unclear and inconclusive findings reported so far in L1 use in an L2 setting make it difficult to generalize results over specific tasks and across languages, or to establish a clear and direct link between L1 use and L2 performance, which is a relevant issue for educational purposes.

With a purpose to contribute to this on-going debate, Kim and Clariana (2015) considered the role L1 plays in L2 reading by applying a Knowledge Structure (KS) analysis approach. This KS approach, explained in more detail below, visually represents L1 and L2 KS changes during L1 use in L2 reading; each student's L1–L2 relationships are explicitly observable and statistically comparable. In that investigation, low-L2 proficient Korean learners of English ($n = 50$) were randomly assigned into one of two groups, either an L2 English-only group or to a Korean-to-English L1 + L2 translation group. First, students were asked to read a 708-word English TOEFL text (Test of English as Foreign Language) and create a concept map of the text, then write a summary (L2 only or L2 + L1), then create a post concept map and take a comprehension posttest. For analysis, all of the students' maps and essays were converted into Pathfinder Network graphs (*PFnets*; Tossell et al. 2010), which are hypothesized to represent the most *salient* underlying organization, or structure, of the data. This approach allowed the researchers to compare the L1 and L2 mental representations and their relationships (see for example, Fig. 1). The results showed that the *translated* treatment's L2 KS is more similar to that of an expert than is the *English-only* treatment (58% vs. 36%) and their KS significantly correlated to the reading comprehension posttest compared to those in the *English-only* condition ($r = 0.84$, accounting for 71% of the variance vs. $r = 0.60$, for 36%). In addition, the average KS of the students in the *translated* condition converged towards a more relational KS (measured

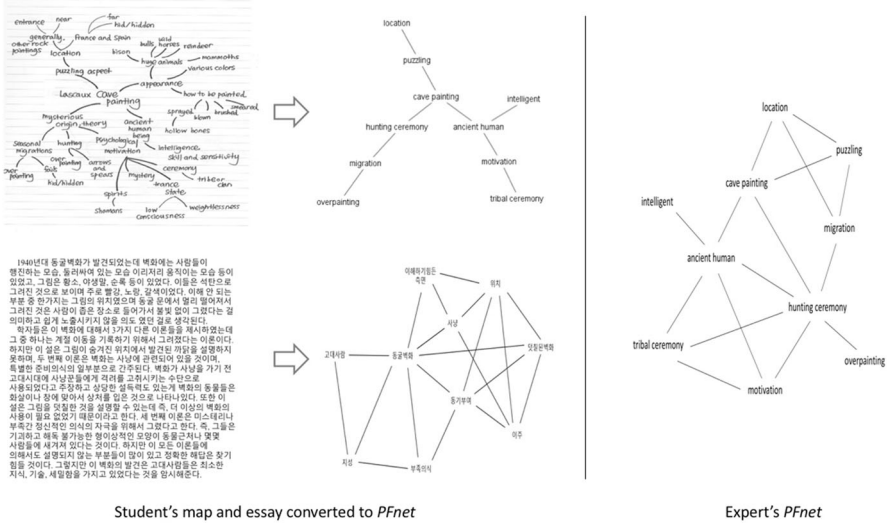


Fig. 1 Example of student’s English map and Korean essay (left) and Expert’s consensus map (right) converted to a Pathfinder Network (PFnet)

as graph centrality) like that of the content expert relative to the English-only treatment that had a more linear KS, indicating a fundamental way that L1 and L2 reading processing differs. These findings indicate that requiring L1 with L2 tasks engenders more complex and relational KS form of the L2 topic, relative to those not required to use L1 while writing, at least for these low proficient Korean learners of English across these two non-cognate languages.

This previous investigation by Kim and Clariana (2015) combined mapping and writing in L2 only, or in L1 + L2, in order to consider the influence of L1 on L2; so questions remain concerning the possible separate unique contribution of mapping alone, writing alone, and both mapping plus writing compared to a no map, no writing control. Thus, this present investigation was designed to confirm and extend this previous investigation as part of an ongoing investigation into the optimal way to use L1 during L2 text comprehension by comparing seven treatments: three experimental conditions including L1 + L2 mapping, L1 + L2 writing, L1 + L2 mapping plus writing, and four control conditions including L2-only mapping, L2-only writing, L2-only mapping plus writing, and a reading-only true control condition.

Purpose

Following Schmidt (2009), the editors of this journal Spector et al. (2015), repeated a call voiced by editors of educational journals to encourage replication and scaling-up investigations as a critical part of the scientific endeavor. This current investigation is a direct replication of Kim and Clariana (2015) that reported a positive influence of L1 use on L2 learning outcomes and KS, using the same lesson content, comprehension posttest, and KS descriptive analysis of students’ maps and essays, but extends that investigation by

isolating the separate contributions of three different lesson tasks (mapping, writing, and both) and with a far larger sample ($n = 245$).

Our working assumption is that persistent KS exists and influences cognition in a way that can be expressed as a nomological network (Cronbach and Meehl 1955, a set of constructs/concepts and their linkages) and that KS can be measured and is worth measuring. Using the KS analysis approach from Kim and Clariana (2015), the purpose of this present investigation is to estimate the separate and the combined effects of mapping and writing with or without L1 to add to the growing evidence of how L1 use influences low proficient English language learners' comprehension and KS.

Method

Participants

The participants were 245 college students in engineering departments from three different South Korean-based Universities. All the participants are native Korean speakers (aged from 20 to 22 years, with 120 females and 125 males). Students with low proficiency in English were selected for this study to better reveal the influence of an L1 in L2 text comprehension. This study used two steps to filter participants' English language proficiency level. The first filtering was to use the web-based Language History Questionnaire 2.0 (LHQ 2.0, Li et al. 2014) to estimate bilinguals' linguistic background and self-reported proficiency level. To attain a low proficient sample, only those participants were selected who had: (a) no intermediate or higher level English instruction and (b) no previous experience studying English abroad. A second round of filtering was used to check an official TOEFL score of the selected participants (from LHQ), which is considered to be a valid and reliable measure of English learners' proficiency (Jamieson et al. 2008). Their TOEFL score was below 80 on the iBT test (internet-based test), which is considered to indicate low proficiency in English (Jamieson et al. 2008). Taken together, we presumed that all students could be considered as low proficient learners of English. The students were briefed on the tasks involved and the purpose of the investigation and were asked to participate, and all agreed. All students received course credit for completing the activities, the experimental materials were not related to the course content.

Procedure

All participants ($n = 245$) read a 708-word English TOEFL text, *The Cave of Lascaux*, the same text used in Kim and Clariana (2015), this text and test were used with permission from the Educational Testing Service (ETS). Participants were randomly assigned to complete one of the seven treatment conditions (e.g., 35 per condition) including three *Experimental conditions requiring L1 tasks*: L1 mapping (L1|M), L1 writing (L1|W), both L1 mapping and writing (L1|MW), and *four Control conditions that did not require L1 tasks*: L2 mapping (L2|M), L2 writing (L2|W), both L2 mapping and writing (L2|MW), or no activity (No). Then all participants completed the ETS multiple-choice posttest that consisted of 10 comprehension-related multiple-choice items that were designed to measure *global inferences* in the text. Global inferences require a well-structured situation model (i.e., a relational-hierarchical KS representation) in contrast to local inferences that only require surface structure representations (i.e., a sequential-linear KS representation).

The posttest had a moderate level of internal consistency, as determined by a Cronbach's alpha of 0.805 in this investigation. All of the students' maps and essays were converted into *PFnets* in order to compare each to an expert's referent *PFnet* (see Fig. 2).

To establish an expert referent model for comparison analysis, the same 10 essential terms were selected from a previous study by Kim and Clariana (2015). In that study, three subject domain experts were provided with a list of all of the terms used by the participants in their maps, arranged in order of frequency of occurrence. The experts were asked to base their selection of terms on importance rather than mere frequency of occurrence, and that they should add terms on the list as needed. While considering this list and the lesson text, the experts collaborated in face-to-face to reach a consensus on the ten essential terms (these include: *cave painting*, *puzzling*, *location*, *seasonal migration*, *hunting ceremony*, *tribal ceremony*, *motivation*, *overpainting*, *ancient human*, *intelligent*). They then created a single expert referent map using the 10 essential terms that was then converted to a *PFnet*. This consensus expert *PFnet* was used as the referent map for comparing to the students' *PFnets*.

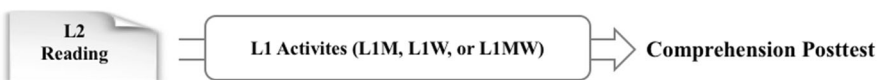
Data analysis

Concept maps to *PFnets*

According to the findings from Clariana et al. (Taricani and Clariana 2006; Poindexter and Clariana 2006; Clariana and Taricani 2010), concept map scores derived from proposition data ('links' between terms) are more related to low-dimensional *verbatim* knowledge such as facts, terminology, and definitions (i.e., a sequential-linear knowledge representation); whereas concept map scores derived from association data ('spatial distances' between terms) are more related to higher-dimensional *gist* knowledge (i.e., a relational-hierarchical knowledge representation). Because of our focus on global inferences, we provided mapping directions to the participants that intentionally deemphasized propositions (*links* between terms) and emphasized association (*distance* between terms). Thus, in this investigation, KS concept map data consists of the distances between the key terms in the maps.

For map to *PFnet* conversion, following Kim and Clariana (2015), *Jrate* software was used to calculate the raw proximity data that contained all pair-wise *distances* between the 10 terms in expert's and in students' maps (*Jrate* is available at <http://endaemon.net/jRateSuite>). All of the students' paper-based maps were recreated in *Jrate* software by the

Experimental Group (L1 task groups, n=105)



Control Group (no L1 task groups; n=140)

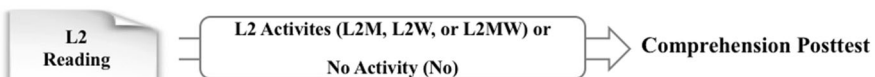


Fig. 2 Research design

researchers; this is a manual process that can introduce error, so care was taken to maintain the original spatial proportional relationships between terms in the map. Each map was converted to a proximity file with 45 elements (i.e., $(10^2 - 10)/2 = 45$, where 10 is the number of selected key concepts) by *Jrate* that are the distances in pixels between every term. Then the *Knowledge Network and Orientation Tool* software (*KNOT*, Tossell et al. 2010; available at <http://interlinkinc.net/>), was used to convert all of the students' and expert's map proximity data from *Jrate* into *PFnets* (see Sarwar 2011, for validity of *KNOT*). Note that the *PFnet* algorithm as data reduction network scaling technique has been widely used in diverse domains, for example, to describe changes in KS during team member collaboration (Clariana et al. 2013), to measure individual and group knowledge convergence (Draper 2013), to reveal L1–L2 interaction (Kim 2017a), to elicit text structure in lesson texts (Fesel et al. 2015), to improve satellite image categorization (Barb et al. 2013), and recently to help understand neurocognitive differences between good versus poor readers by comparing their *PFnets* and neural patterns (Li and Clariana 2018).

Essays to *PFnets*

For essay to *PFnet* conversion, following Kim and Clariana (2015), we employed a text visual analytic system, *Graphical Interface of Knowledge Structure* (*GIKS*, Kim 2017b), for converting Korean and English essays into *PFnets*. The *GIKS* system was developed by integrating two standalone offline software tools, *ALA-Reader* and *KNOT*. Here, the *ALA-Reader* algorithm was used to transform the Korean and English essays into raw proximity data for the same 10 terms. The approach is too lengthy to describe here, for complete details on how *ALA-Reader* was developed and has evolved until now, refer to Clariana et al. (2009, 2014), and Kim (2017c); how it has been recently applied in online learning settings, refer to Tawfik et al. (2018) and Zimmerman et al. (2018). In brief, *ALA-Reader* is a tool designed to capture relations of preselected key terms in a text as pair-wise links in a proximity file, adding only “1 (co-occurred)” or “0” to indicate the sequential occurrence of the key terms in the text (see Kim 2012, for validity of *ALA-Reader*). Then, all of the students' essay proximity data from the *ALA-Reader* analysis were converted to *PFnets* using *KNOT*. Thus, the students' *PFnets* from their concept maps and written essay responses can be directly compared, maps-to-maps, essays-to-essays, and even maps-to-essays (refer back to Fig. 1).

PFnet analysis

Following Kim and Clariana (2015), the raw proximity data from all of the maps and essays were analyzed using two different approaches: (1) Pearson correlation of the raw proximity data as Fisher z and (2) two graph-theoretical measures of the *PFnets* derived from the raw proximity data, node centrality and graph centrality (Coronges et al. 2007).

Note that no studies have yet explicated which specific network analysis approach should be applied in specific settings. Pathfinder analysis is a severe data reduction approach where *PFnets* are believed to consist of the most salient relations in a set of proximity data (i.e., pruned proximity data, Sarwar 2011) and thus may be a better data set than comparing the raw data, although it is prudent and useful to analyze and report both data sets within the same study. Thus, the study carries out analyses using three different levels of data reduction (raw proximity data to pruned proximity data), with correlation of the raw proximity data having the greatest amount of information (both true and error), and centrality of the pruned proximity data (represented as *PFnets*) having the most salient

information; node centrality which represents networks as n -degree vectors and graph centrality where a network is represented as an integer on a continuum between zero and one.

For example, Clariana et al. (2013) used centrality to mathematically and visually describe state changes from problem space to problem solution during problem solving. Participants ($n = 140$) were randomly assigned to interdependent or non-interdependent conditions to work online in triads to create a concept map to solve a problem scenario. *Node centrality* of the interdependent group-created concept maps resembled the fully explicated problem space, while the non-interdependent group-created concept maps mainly resembled the problem solution, confirming prior studies using a ‘manual’ common relations count analysis of concept maps (Engelmann and Hesse 2010). The results demonstrated that node centrality can provide a complementary measure of the KS contained in the team concept map artifacts. *Graph centrality* was not significant in that investigation, but the interdependent triads’ average graph centrality is indicative of expertise relative to the non-interdependent triads’ centrality that was indicative of goal-oriented thought (see Fig. 4 below where centrality is explained in detail).

Correlation analysis The *KNOT* tool automatically computes the correlation (r) between the 45-element distance data in two *PFnet* files; the distances between a pair of nodes in one *PFnet* are compared to the distances between the same pair of nodes in another *PFnet*. Next, Fisher’s r to z transformation was applied to the correlation r values prior to averaging and statistical comparison of the two groups because Pearson correlation values are not interval-level data, and so correlations cannot be averaged together, while Fisher z values are additive. Then these Fisher values for each student’s network were averaged and compared to the expert’s using a Pearson correlation coefficient (r^2).

Degree centrality Centrality can provide both a local-level (node centrality; a measure of *node importance*) and a global-level measure (graph centrality; a measure of *network form*) of a network graph (see for example, Ifenthaler 2010; Clariana et al. 2013; Kim and Clariana 2015, 2016).

Node centrality indicates the relative importance of each node in the graph based on the number of links that the node has with other nodes. A node with the most links would probably be the most important term in the network graph relative to the node with fewer links. To determine the node degree vectors for each network (i.e., degree is the number of links to a node), all students’ *PFnets* and the experts’ consensus *PFnet* were manually inspected to create a table of node degree vectors using the 10 terms established by the expert panel (e.g., see Fig. 3), then each vector in the table is correlated with the expert vector. For example, the correlation of the example participant map node degree vector to the expert map is $r = 0.57$, that has a coefficient of determination $r^2 = 0.32$, indicating a 32% overlap of these two maps.

Graph centrality provides a numerical measure of map structure, or form, that ranges from 0 to 1 (see Fig. 4). These values have been categorized into a conceptual typology: where 0–0.2 indicates a linear network form that aligns with a goal-orientation approach such as a procedure, 0.2–0.4 indicates an ordered network form that indicates a simple and rigid expertise, 0.4–0.6 indicates a more cross-connected network form that indicates complex and fluid expertise, and 0.6–1.0 indicates a star-shaped network with one or a few central nodes that indicates naiveté. Kim and Clariana’s (2015) study with low proficient L2 readers showed that the KS ‘form’ of the L2 readers who used L1 tended to have expertise-like *network* structures ($C_{graph} = 0.57$) compared to the L2 only group that had a more goal-oriented *linear* structure ($C_{graph} = 0.28$). In this investigation, we assume that the shapes of KS representations from the two languages, English and Korean, would be

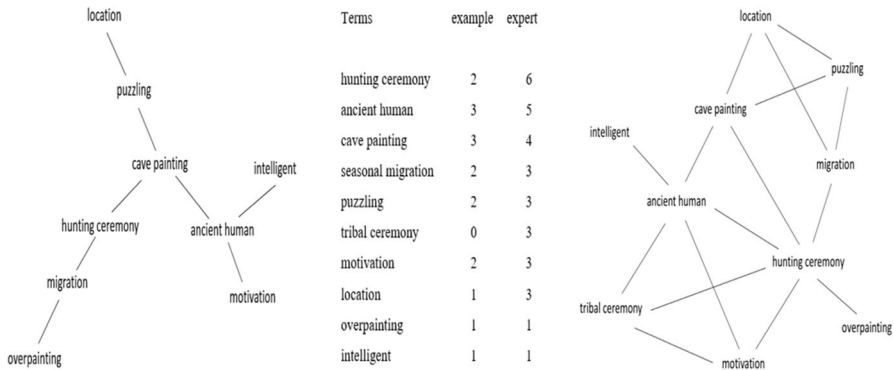


Fig. 3 An example student PFnet (left) and the expert consensus PFnet (right) along with the 10-element node degree vector for each PFnet

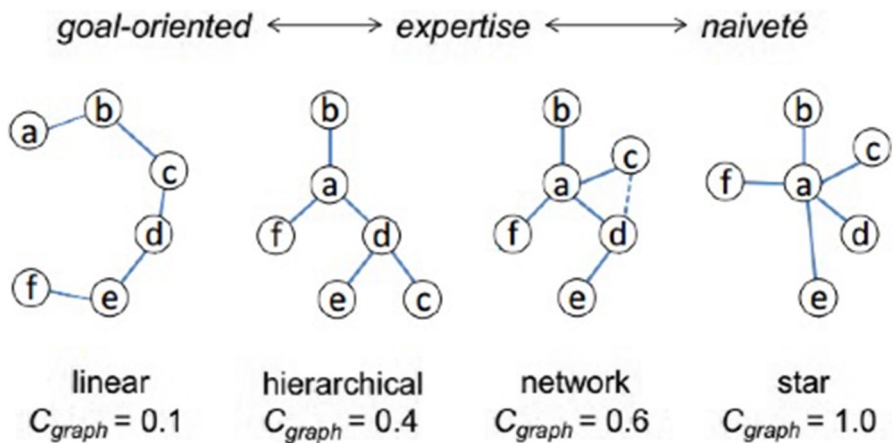


Fig. 4 Graph centrality calculated for four network graph forms

fundamentally different, as evidenced in other studies (e.g., Kim 2017c; Tang and Clariana 2017), so a graph centrality measure would be very helpful to reveal the different ‘structure’ inherent in maps and essays from the different languages.

The investigations cited above demonstrate that centrality measures provide a different vantage point and descriptive frame as a visually-based measurement approach of mental representations that is mainly descriptive.

Results

The data for analysis included the students’ comprehension posttest and individual maps and essays converted to *PFnets*. Comprehension posttest findings are presented, and then, following the analysis approach from Kim and Clariana’s (2015) study, the *PFnet* data are described and compared using Pearson correlation as (1) correlation to the expert and (2)

Table 1 Means and standard deviations for the comprehension posttest (max. 10) and for two task measures, knowledge structure similarity to the expert as Fisher Z and graph centrality as a measure of KS form (C_{graph})

	L1 task groups			No L1 task groups (control)			
	L1 M	L1 W	L1 MW	L2 M	L2 W	L2 MW	No
Posttest	5 (1.69)	6.1 (2.01)	7.5 (2.45)	3.6 (1.41)	2.7 (1.77)	4 (1.52)	3.3 (1.87)
Fisher Z	0.51	0.59	0.79	0.44	0.35	0.45	–
C_{graph}	0.43 relational	0.51 relational	0.54 relational	0.29 linear	0.24 linear	0.31 linear	–

Expert $C_{graph} = 0.47$; Cronbach’s alpha of posttest = 0.805

correlation to the posttest, and network centrality measures as (3) graph centrality and (4) node centrality (see Table 1).

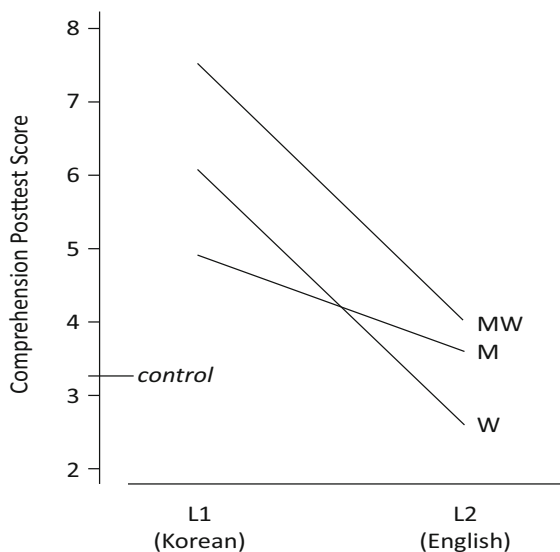
Comprehension posttest

The comprehension multiple-choice posttest data were analyzed by ANOVA (2×3 between groups) to compare the effect of two factors, L1–L2 use (L1 spontaneous or L1 required) across three tasks (mapping, writing, or both). To determine the data’s characteristics as it relates to ANOVA, three tests were performed. Skewness (-0.19) and kurtosis (-0.42) are within acceptable parameters, the sample size is adequate ($n = 120$), but Levene’s test was significant, $F(5,99) = 5.42, p < 0.001$ indicating that the error variance was not equal across groups. As ANOVA is relatively robust when only one assumption is violated, we report results.

The L1–L2 factor was significant, $F(1,99) = 127.54, MSE = 1.53, p < 0.001$; the tasks factor was significant, $F(2,99) = 16.06, MSE = 1.53, p < 0.001$; the interaction of L1–L2 and tasks, $F(2,99) = 8.87, MSE = 1.53, p < 0.001$. The significant interaction is shown in Fig. 5. Cohen’s d effect sizes for these six groups relative to the read-only control are, $d = L2|W (-0.33) < control < L2|M (0.16) < L2|MW (0.36) < L1|M (0.80) < L1|W (1.37) < L1|MW (2.09)$. Note that L2|W is worse than reading only, requiring L1 is consistently better, and combining mapping and writing is better than mapping and writing alone.

Multiple regression analysis was then used to test if the KS measures significantly predicted students’ comprehension posttest performance. The results of the regression indicated the combined measure explained 74.9% of the variance ($R^2 = 0.749, F(2,102) = 152.49, p < 0.001$). It was found that KS similarity as Fisher z significantly predicted comprehension posttest ($\beta = 0.65, p < 0.001$), as did graph centrality ($\beta = 0.36, p < 0.001$).

Fig. 5 The significant interaction of L2 use and tasks (*M* mapping, *W* writing)



Correlation of PFnets

Correlation to the expert

Participants' map scores are the Fisher z transform of the Pearson correlation between the participant maps and the expert referent map. As seen in Table 1 above, all of the map scores where L1 was required in the lesson (the experimental groups) were much higher compared to the groups' scores where L1 was not required (the control groups). The finding of the current study thus corroborates that of Kim and Clariana (2015), which reported that the process of mapping and/or writing in an L1 for the L2 text comprehension has strong learning benefits for the low L2 proficient learners using non-cognate Korean-English language L1–L2 tasks.

Further, mapping and writing in combination are additive and superior to mapping or writing alone, whether L1 is required or not. This finding is in agreement with past studies reporting that essays and concept maps are considered to be highly related and complementary forms of assessment to evaluate KS, at least in L1–L1 tasks (Clariana and Koul 2004, 2008; Gonzalvo et al. 1994), especially if the concept maps are used as a scaffold or outline for writing activity and vice versa. The present study confirms the past studies and further suggests that combining an L1 mapping and writing activity could be a more effective way to support learning in an L2 (in this case L2 text comprehension) relative to an L1 mapping alone and an L1 writing alone.

Correlation to the posttest

All individual map Fisher z scores (i.e., determined as correlation to the expert) were compared to the comprehension posttest scores using Pearson's correlation (see Table 2). This finding is in agreement with the findings explained above which showed that when L1 is required in the lesson (experimental groups), participants' maps and posttest scores were much more highly correlated compared to those in the groups where L1 is not required.

Further, the L1|MW group had the strongest correlation ($r = 0.88$) between map scores and comprehension posttest scores; this suggest that combining mapping and writing tasks leads to more an expert-like KS that is more able to perform well on a comprehension posttest that measures global inferences. Hence, in general terms, superior KS ‘form’ allows for superior ‘cognitive function’.

We can conclude from the results of the correlation analysis that requiring L1 (mapping, writing, or both) strongly influences L2 performance (L2 posttest text comprehension). Therefore, an effective way to use an L1 is to require *both* mapping and writing compared to mapping or writing alone, at least for low proficient participants working with non-cognate languages.

Centrality analysis

Graph centrality

Graph centrality as operationalized by Clariana et al. (2013) has values that range from 0 (linear network) to 1.0 (star network), with optimally relational and complex networks typically falling around 0.4–0.5. As shown in Fig. 6, the control groups’ maps on average were somewhat *linear* in form (C_{graph} , 0.29–0.24–0.31 respectively) while the three experimental groups’ maps were all quite *relational* in form (0.43–0.51–0.54 respectively) and similar to the expert (0.47). This could indicate that the average KS of participants who were required to use L1 in the lesson (e.g., experimental groups) converged toward a more *relational* structure similar to the expert referent compared to the control students that have a more *linear* structure. The result is consistent with that of Kim and Clariana (2015) who reported that the average KS of the students in the translated writing condition (i.e., L1 task group) converged towards a more relational structure ($C_{graph} = 0.51$); in contrast the direct writing condition (i.e., no L1 task group) showed a more linear structure ($C_{graph} = 0.28$), thus suggesting a fundamental KS difference when L1 tasks are required in L2 settings.

From a fuzzy-trace theory view (Reyna and Brainerd 1995), we assume that a well-developed relational and complex KS would be better for supporting high-order inferential *gist* knowledge tasks relative to a linear KS that would be better for *verbatim* knowledge tasks. To test this assumption, all individual graph centrality vectors were compared to their inference-required comprehension posttest scores. As shown in Fig. 7, the control groups’ non-linear relationships between graph centrality and posttest were not significant. But, the experimental groups’ maps all show a significant non-linear relationship between graph centrality and posttest comprehension score. That is, perhaps too little structure (that might indicate a deficient map) or too much structure (which might indicate inappropriate

Table 2 Pearson correlations of individuals’ posttest to map scores

	L1 task groups				No L1 task groups (control)				
	L1M	L1W	L1MW		L2M	L2W	L2MW		No tasks
			L1M	L1W			L1M	L1W	
Posttest	0.52*	0.61**	0.54**	0.88*	0.35	0.29	0.37	0.34	–

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

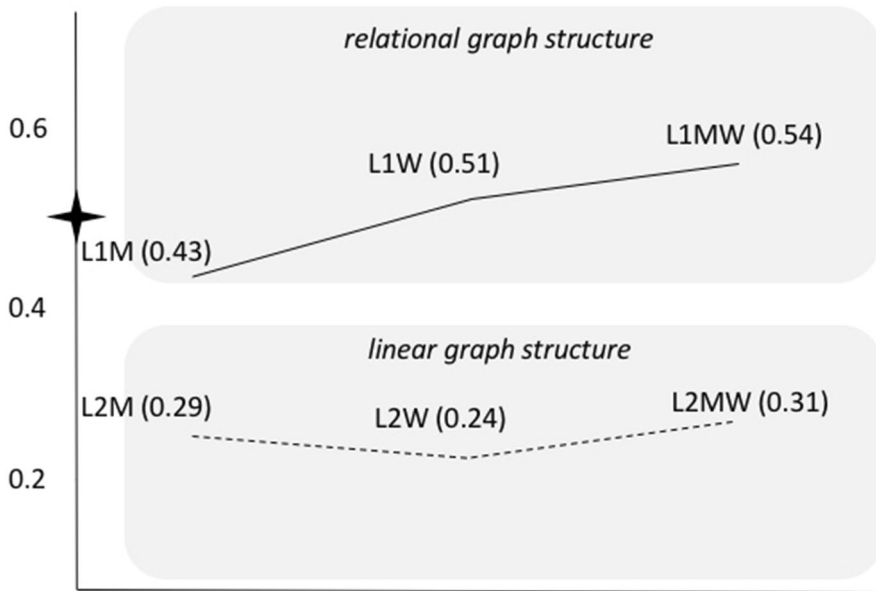


Fig. 6 Graph centrality of experimental (solid) and control groups (dash). A large star indicates the *expert* graph centrality value (at 0.47)

dominance of the structure by irrelevant terms) both negatively influence posttest performance, but optimal structure (graph centrality of about 0.45 for L1M, 0.50 for L1W, 0.55 for L1MW) relates to higher posttest scores. This result could be used to explain in part why the average graph centrality values of the experimental groups' maps ranged from 0.43 to 0.54, and why their performance on the posttest was much higher than the control groups. A nearly identical "inverted U-shaped curve" of the optimal KS centrality values was also observed in Kim and Clariana (2015), thus form relates to function.

Node centrality

Group map 'relationships' Following Kim and Clariana's (2015) approach for visually depicting group relations, Proxscal multidimensional scaling (SPSS 20.0) was used to scale the node centrality vectors of the averaged groups and the expert referent map (see Fig. 8; see for detail Clariana et al. 2013). As shown in Fig. 8, the Expert map fell toward the left of the figure. The dimensional representation indicates that the experimental maps are towards the left, so we infer from the diagram that the L1 task groups' maps moved to the left of the multidimensional scaling (MDS) representation "converging" to the Expert map. The control groups' maps were distributed across the MDS figure, far distant from the Expert map, with no clear convergence.

Group map 'convergence' Group member convergence describes how similar the maps of individuals within a condition are to each other. The rationale is that a potent treatment condition influences an individual's KS to be more like the KS of the others who received the same treatment. Group KS convergence is measured here by how much map node degree vectors are alike (i.e., as average percent overlap) with every map compared to all maps in a condition (see, Clariana, et al. 2013, for details). In the present study, the three

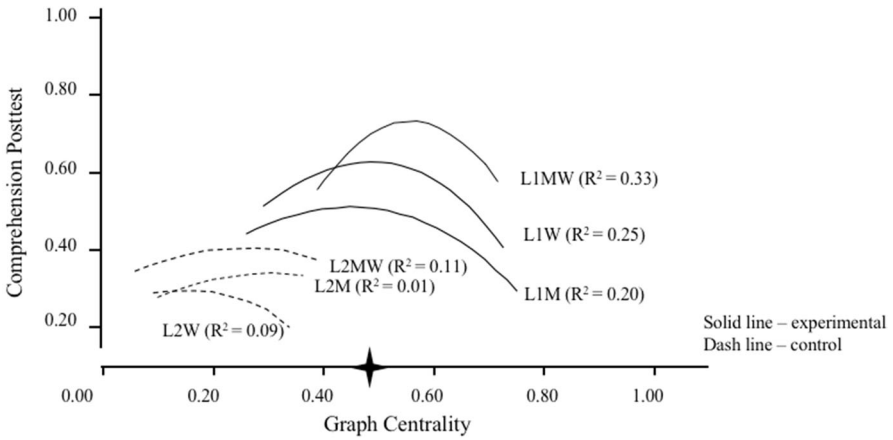


Fig. 7 The relationship between graph centrality and posttest scores. A large star indicates the *expert* graph centrality value (at 0.47)

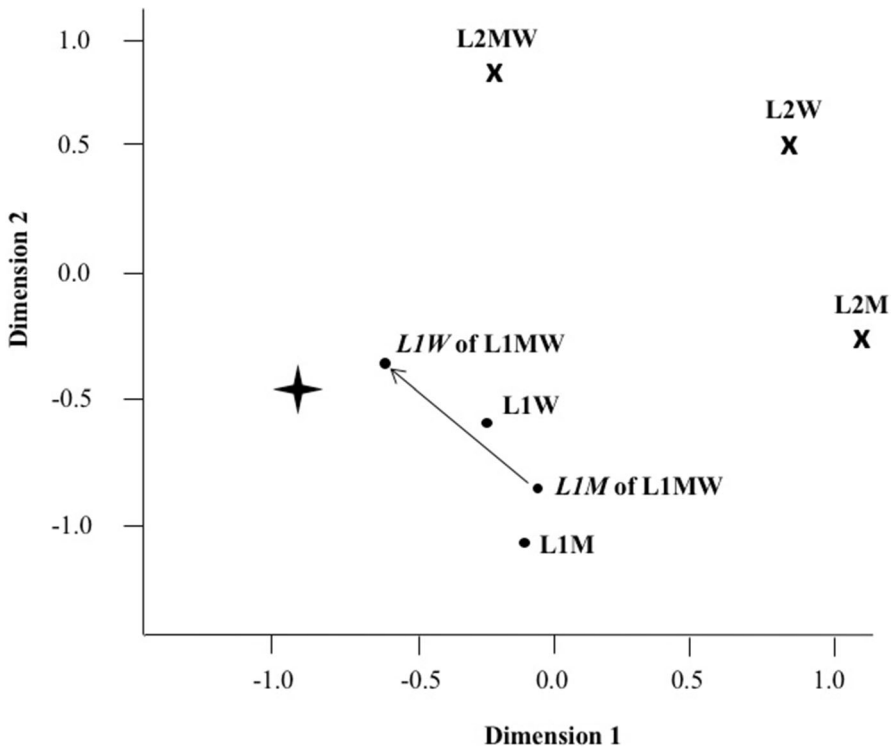


Fig. 8 The proxscal 2-dimensional representation of the group-average map node centrality. Solid dot indicates experimental group maps and "X" indicates the control group maps. A large star indicates the *expert* vector

experimental groups' average overlap percent is 72% (SD = 0.10) and the four control groups' average overlap percent is 33% (SD = 0.12), this difference has an effect size of $d = 1.25$, that is statistically significant (t test $p = 0.01$). Thus, requiring L1 tasks leads to very homogenous individual maps that are more similar to the expert referent map. The MDS in Fig. 8 above also shows this difference in convergence in each condition (convergence is represented as nearness in the MDS). The three experimental groups' average positions were closer to each other in the 2-dimensional space compared to the four control groups and also were closer to the expert referent with L1MW being the closest of all.

Discussion

This investigation confirmed and extended the findings of Kim and Clariana (2015) of the influence of L1 tasks on learning. Using KS network analysis methods, we sought to estimate the separate and the combined effects of mapping and writing with or without L1 to add to the growing evidence of how L1 use influences low proficient English language learners' L2 text comprehension and knowledge structure. Four analyses were conducted that provide clear and overlapping evidence of the influence of L1 tasks on learning as mediated by apposite knowledge structure.

Correlation results

In accordance with Kim and Clariana's (2015) study, the correlation results of this present investigation showed that requiring L1 tasks to complement L2 reading comprehension (the experimental groups) leads to knowledge structure that is more similar to that of the expert and that is significantly correlated to L2 comprehension posttest scores, relative to the control groups that were not required to compete L1 tasks. In addition, the most interesting result to emerge from the correlation data is that combining mapping and writing activities is additive compared to mapping alone or writing alone, at least for these low L2 proficient participants. Both mapping and writing contribute orthogonally to comprehension posttest performance and to KS change, although writing has a stronger effect than mapping.

Centrality results

First, the graph centrality results indicate that requiring L1 in these L2 tasks helps these low proficient participants who otherwise would obtain a fairly *linear* L2 KS for this text passage to have a more *relational* L2 KS, and this relational structure had a strong positive correlation with their L2 posttest scores. This result also corroborates the findings of the correlation analysis. This means that participants maintained a different knowledge structure in L1 and L2 about this topic (e.g., an example of L1–L2 asymmetry, see Brysbaert and Duyck (2010) and van Hell and Kroll (2013)). Requiring L1 use appears to alter the less complex linear L2 KS to be more like the complex relational L1 KS but still remain different from it.

The group-level node centrality values displayed as a 2-dimensional MDS chart clearly accounted for and distinguished between the experimental groups' and the control groups' KS. The MDS representation shows that the maps of the experimental groups 'converged' towards the expert referent map (as closeness in MDS space), while the control groups'

maps were fairly widely scattered and further away from the expert referent map. We presume that having a KS more similar to an expert results in better comprehension. The MDS results show that requiring L1 tasks leads to group-level knowledge structure convergence with the expert. Interestingly, this MDS analysis also visually showed that the L1IMW group's map is closest to the Expert map than the other experimental groups (L1IM and L1IW). This supports the use of L1 mapping and writing together as an effective way to use L1 in L2 reading. Also, group-level cognition is an important new area of theory and investigation in the Learning Sciences, and so these group-level KS measures reported here add to the empirical base for theories of group level cognition (Stahl 2010).

The findings of this investigation only generalize to low English proficient Korean students. As discussed above, high English proficient students may be quite able to read and comprehend at a high level in English-only tasks. Further, different language groups may exhibit dominant knowledge structure forms, especially Korean baseline knowledge structure which is relational complexity may be fundamentally different than that of other languages, especially English which is especially linear (Kim et al. 2016). More research is needed to determine how knowledge structures especially as manifested in different languages influences cognition.

Conclusion

In accordance with the previous investigation (Kim and Clariana 2015), this present investigation has again demonstrated that the required use of their native language can provide low proficient English language learners with additional cognitive support that leads to improved reading comprehension and better post-reading KS of the L2 topic than would be possible were they restricted to sole use of their L2. Further, this investigation clearly shows that combining L1 mapping and writing activities may be the most effective way for learning in L2 reading than just mapping alone or writing alone. And finally, these findings further validate this technology-based approach, pathfinder network scaling, for measuring KS contained in bilingual learners' productions. A fully automated version of this approach (e.g., refer to Kim 2017b, for text to PFnet conversion) would provide researchers with another measurement tool to better understand the mediating influence of lesson tasks on language processes and performances.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Anton, M., & DeCamilla, F. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Review*, 54, 314–342. <https://doi.org/10.1111/0026-7902.00018>.
- Barb, A. S., Clariana, R. B., & Chi-Ren, S. (2013). Applications of PathFinder Network Scaling for improving the ranking of satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3), 1092–1099.
- Brooks, F. B., & Donato, R. (1994). Vygotskian approaches to understanding foreign language learner discourse during communicative tasks. *Hispania*, 77, 262–274. <https://doi.org/10.2307/344508>.
- Brysbart, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition*, 13, 359–371.

- Chee, M. W. L., Hon, N., Lee, H. L., & Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *NeuroImage*, *13*, 1155–1163. <https://doi.org/10.1006/nimg.2001.0781>.
- Clariana, R. B., Engelmann, T., & Yu, W. (2013). Using centrality of concept maps as a measure of problem space states in computer-supported collaborative problem solving. *Educational Technology Research and Development*, *61*, 423–442.
- Clariana, R. B., & Koul, R. (2004). A computer-based approach for translating text into concept maplike representations. In Proceedings of the first international conference on concept mapping (pp. 14–17).
- Clariana, R. B., & Koul, P. (2008). The effects of learner prior knowledge when creating concept maps from a text passage. *International Journal of Instructional Media*, *35*, 229–236.
- Clariana, R. B., & Taricani, E. M. (2010). The consequences of increasing the number of terms used to score open-ended concept maps. *International Journal of Instructional Media*, *37*, 163–173.
- Clariana, R. B., Wallace, P. E., & Godshalk, V. M. (2009). Deriving and measuring group knowledge structure from essays: The effects of anaphoric reference. *Educational Technology Research and Development*, *57*, 725–737.
- Clariana, R. B., Wolfe, M. B., & Kim, K. (2014). The influence of narrative and expository text lesson text structures on knowledge structures: Alternate measures of knowledge structure. *Educational Technology Research and Development*, *62*, 601–616.
- Cohen, A. D., & Brooks-Carson, A. (2001). Research on direct versus translated writing: Students' strategies and their results. *The Modern Language Journal*, *85*, 169–188. <https://doi.org/10.1111/0026-7902.00103>.
- Coronges, K. A., Stacy, A. W., & Valente, T. W. (2007). Structural comparison of cognitive associative networks in two populations. *Journal of Applied Social Psychology*, *37*(9), 2097–2129.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281. <https://doi.org/10.1037/h0040957>.
- Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, *7*, 482–511. <https://doi.org/10.1177/0741088390007004003>.
- Draper, D. C. (2013). The instructional effects of knowledge-based community of practice learning environment on student achievement and knowledge convergence. *Performance Improvement Quarterly*, *25*(4), 67–89.
- Engelmann, T., & Hesse, F. W. (2010). How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning*, *5*(3), 299–319.
- Fesel, S. S., Segers, E., Clariana, R. B., & Verhoeven, L. (2015). Quality of children's knowledge representations in digital text comprehension: Evidence from pathfinder networks. *Computers in Human Behavior*, *48*, 135–146. <https://doi.org/10.1016/j.chb.2015.01.014>.
- Gonzalvo, P., Canas, J. J., & Bajo, M. (1994). Structural representations in knowledge acquisition. *Journal of Educational Psychology*, *86*, 601–616. <https://doi.org/10.1037/0022-0663.86.4.601>.
- Ifenthaler, D. (2010). Relational, structural, and semantic analysis of graphical representations and concept maps. *Educational Technology Research and Development*, *58*, 81–97. <https://doi.org/10.1007/s11423-008-9087-4>.
- Jamieson, J., Enright, M., & Chapelle, C. (2008). *Building a validity argument for the Test of English as a Foreign Language*. New York: Routledge.
- Jones, S., & Tetroe, J. (1987). Composing in a second language. In A. Matsuhashi (Ed.), *Writing in real time: Modelling production processes* (pp. 34–57). Norwood, NJ: Ablex.
- Karim, K. (2010). First language (L1) influence on second language (L2) reading: The role of transfer. *Working Papers of the Linguistics Circle*, *17*, 49–54.
- Kim, M. K. (2012). Cross-validation study of methods and technologies to assess mental models in a complex problem solving situation. *Computers in Human Behavior*, *28*(2), 703–717.
- Kim, K. (2017a). Visualizing first and second language interactions in science reading: A knowledge structure network approach. *Language Assessment Quarterly*, *14*, 328–345.
- Kim, K. (2017b). Graphical interface of knowledge structure: A web-based research tool for representing knowledge structure in text. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/s10758-017-9321-4>.
- Kim, K. (2017c). An automatic measure of cross-language text structures. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/s10758-017-9320-5>.
- Kim, K., & Clariana, R. B. (2015). Knowledge structure measures of reader's situation models across languages: Translation engenders richer structure. *Technology, Knowledge and Learning*, *20*, 249–268.

- Kim, K., & Clariana, R. B. (2016). Text signals influence second language expository text comprehension: Knowledge structure analysis. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-016-9494-x>.
- Kim, K., Clariana, R. B., Alqahtani, M. A., & Tang, H. (2016). Revealing knowledge structure in lesson texts using a computational text pattern-matching approach. Paper presented at the annual meeting of the Association for Educational Communication and Technology (AECT), Jacksonville, FL.
- Kobayashi, H., & Rinnert, C. (1992). Effects of first language on second language writing: Translation versus direct composition. *Language Learning*, 42, 183–215. <https://doi.org/10.1111/j.1467-1770.1992.tb00707.x>.
- Li, P., & Clariana, R. B. (2018). Reading comprehension in L1 and L2: An integrative approach. *Journal of Neurolinguistics*.
- Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, 17(03), 673–680. <https://doi.org/10.1017/s1366728913000606>.
- Pennington, M., & So, S. (1993). Comparing writing process and product across two languages: A study of six Singaporean university student writers. *Journal of Second Language Writing*, 2, 41–63. [https://doi.org/10.1016/1060-3743\(93\)90005-N](https://doi.org/10.1016/1060-3743(93)90005-N).
- Poindexter, M. T., & Clariana, R. B. (2006). The influence of relational and proposition-specific processing on structural knowledge and traditional learning outcomes. *International Journal of Instructional Media*, 33, 177–184.
- Qi, D. S. (1998). An inquiry into language-switching in second language composing processes. *The Canadian Modern Language Review*, 54, 413–435. <https://doi.org/10.3138/cmlr.54.3.413>.
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, 7, 1–75.
- Roca de Larios, J., Murphy, L., & Manchon, R. (1999). The use of restructuring strategies in EFL writing: A study of Spanish learners of English as a Foreign Language. *Journal of Second Language Writing*, 8, 13–44. [https://doi.org/10.1016/S1060-3743\(99\)80111-8](https://doi.org/10.1016/S1060-3743(99)80111-8).
- Ryan, C. (2013). American Community Survey Reports: Language Use in the United States: 2011. <https://www.census.gov/prod/2013pubs/acs-22.pdf>
- Sarwar, G. S. (2011). *Structural assessment of knowledge for misconceptions: Effectiveness of structural feedback provided by pathfinder networks in the domain of physics*. Kolln: LAP Lambert Academic Publishing.
- Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, 13, 90–100. <https://doi.org/10.1177/1745691613514450>.
- Spector, J. M., Johnson, T. E., & Young, P. A. (2015). An editorial on replication studies and scaling up efforts. *Educational Technology Research and Development*, 63, 1–4. <https://doi.org/10.1007/s11423-014-9364-3>.
- Stahl, G. (2010). Group cognition as a foundation for the new science of learning. In M. S. Khine & I. M. Saleh (Eds.), *New science of learning: Cognition, computers, and collaboration in education* (pp. 23–34). New York, NK: Springer.
- Tang, H., & Clariana, R. (2017). Leveraging a sorting task as a measure of knowledge structure in bilingual settings. *Technology, Knowledge and Learning*, 22(1), 23–35.
- Taricani, E. M., & Clariana, R. B. (2006). A technique for automatically scoring open-ended concept maps. *Educational Technology Research and Development*, 54, 61–78.
- Tawfik, A. A., Law, V., Ge, X., Xing, W., & Kim, K. (2018). The effect of sustained vs. faded scaffolding on students' argumentation in ill-structured problem solving. *Computers in Human Behavior*.
- Tossell, C. C., Schvaneveldt, R. W., & Branaghan, R. J. (2010). Targeting knowledge structures: A new method to elicit the relatedness of concepts. *Cognitive Technology*, 15(2), 11.
- Tse, C. S., & Altarriba, J. (2012). The effects of first-and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task. *Bilingualism: Language and Cognition*, 15(03), 663–676.
- Uzawa, K., & Cumming, A. (1989). Writing strategies in Japanese as a foreign language: lowering or keeping up the standards. *The Canadian Modern Language Review*, 46, 178–194.
- van Hell, J. G., & Kroll, J. F. (2013). *Using electrophysiological measures to track the mapping of words to concepts in the bilingual brain: a focus on translation*. Cambridge: Cambridge University Press. Retrieved from <http://repository.uhn.ru.nl/handle/2066/116759>
- Wang, L. (2003). Switching to first language among writers with differing second-language proficiency. *Journal of Second Language Writing*, 12, 347–375. <https://doi.org/10.1016/j.jslw.2003.08.003>.

- Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing, 11*, 225–246. [https://doi.org/10.1016/S1060-3743\(02\)00084-X](https://doi.org/10.1016/S1060-3743(02)00084-X).
- Woodall, B. R. (2002). Language-switching: Using the first language while writing in a second. *Journal of Second Language Writing, 11*, 7–28. [https://doi.org/10.1016/S1060-3743\(01\)00051-0](https://doi.org/10.1016/S1060-3743(01)00051-0).
- Zhao, X., & Li, P. (2013). Simulating cross-language priming with a dynamic computational model of the lexicon. *Bilingualism: Language and Cognition, 16*, 288–303.
- Zimmerman, W. A., Kang, H. B., Kim, K., Gao, M., Johnson, G., Clariana, R., et al. (2018). Computer-automated approach for scoring short essays in an introductory statistics course. *Journal of Statistics Education, 26*(1), 40–47.

Kyung Kim is an Assistant Professor of the Department of Educational Technology Research and Assessment at Northern Illinois University. He has worked on the measures and applications of knowledge structure in various learning environments.

Roy B. Clariana is a Professor of Learning, Design, and Technology program and Department Head of Learning and Performance Systems at the Pennsylvania State University. His current research focus is on knowledge structure measures.