Published in final edited form as:

Appl Nurs Res. 2018 April ; 40: 99–105. doi:10.1016/j.apnr.2018.01.006.

A Preliminary Psychometric Evaluation of the Eight-Item Cognitive Load Scale

Grant A. Pignatiello, BSN, RN¹ [Doctoral (PhD) Candidate], Emily Tsivitse, BSN, RN [Doctoral (PhD) Student], and Ronald L. Hickman Jr., PhD, RN, ACNP-BC, FNAP, FAAN¹ [Associate Professor of Nursing and Assistant Dean for Research]

¹Case Western Reserve University, Cleveland, Ohio

Abstract

Aim—The aim of this article is to report the psychometric properties of the eight-item cognitive load scale.

Background—According to cognitive load theory, the formatting and delivery of healthcare education influences the degree to which patients and/or family members can engage their working memory systems for learning. However, despite its relevance, cognitive load has not yet been evaluated among surrogate decision makers exposed to electronic decision support for healthcare decisions. To date, no psychometric analyses of instruments evaluating cognitive load have been reported within healthcare settings.

Methods—A convenience sample of 62 surrogate decision makers for critically ill patients were exposed to one of two healthcare decision support interventions were recruited from four intensive care units at a tertiary medical center in Northeast Ohio. Participants were administered a battery of psychosocial instruments and the eight-item cognitive load scale (CLS).

Results—The CLS demonstrated a bidimensional factor structure with acceptable discriminant validity and internal consistency reliability (Cronbach's $\alpha = .75$ and $.89$).

Conclusions—The CLS is a psychometrically sound instrument that may be used in the evaluation of decision support among surrogate decision makers of the critically ill. The authors recommend application of the cognitive load scale in the evaluation and development of healthcare education and interventions.

Keywords

Cognitive Load; Decision Support; Working Memory; Psychometric Analysis; Healthcare technology

Corresponding Author: Grant Pignatiello, BSN, RN, Address: 2120 Cornell Road, Cleveland, OH 44106-4904, Telephone: 216 368-4979, Fax: 216 368-5989, gap54@case.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

As the aging population grows in the United States, the projected incidence of chronic critical illness is expected to double by the year 2020, largely affecting older adults and the families who care for them (Zilberberg, de Wit, Pirone, & Shorr, 2008). Characterized by multisystem organ dysfunction and cognitive impairment, the chronically critically ill (CCI) often require extended stays in an intensive care unit (ICU) and have higher mortality rates when compared to the general population of the critically ill. Because of the severity of illness and rates of cognitive impairment of CCI patients, their family members must often serve as surrogate decision makers (SDMs) and make complex treatment and/or end-of-life decisions on behalf of these patients. The participation of SDMs in the decision-making process for a CCI patient can present profound states of cognitive and emotional burden among SDMs who are often unprepared to serve in such a role (Pignatiello, Hickman, & Hetland, 2016). Specifically, there is a dearth of psychometrically sound instruments that capture the influential cognitive components of decision making among SDMs of ICU patients.

Working memory is an essential cognitive function for learning and making judgments that is underappreciated in the evaluation of behavioral research. Working memory involves the cognitive processes of processing, storing, and manipulating information. First described by Sweller (1988), the degree to which working memory is influenced during the learning process is contingent upon the *instructional design* used to convey the learned material. The influence of instructional design on working memory is now known as *cognitive load*. There are currently two recognized types of cognitive load: *intrinsic* and *extraneous*. *Intrinsic cognitive load* (ICL) represents the difficulty of the subject matter at hand and depends on the prior knowledge of the learner (Leppink, Gog, Paas, & Sweller, 2015). *Extraneous cognitive load* (ECL) represents the load imposed on working memory through instructional design methods that are not necessary for learning (e.g., presentation of redundant information). If an individual experiences an undesirable level of ICL or ECL during the learning process, a detriment to the learning process may occur. Within health care, such learning detriments may translate to undesirable health behaviors and/or impaired decision making.

Leppink and Heuvel (2015) proposed a psychometrically sound, self-report instrument that measures cognitive load. This eight-item instrument possesses two dimensions, with items one through four measuring intrinsic cognitive load, and items five through eight measuring extraneous cognitive load. To our knowledge, the Cognitive Load Scale (CLS) has not been directly applied to behavioral interventions within health care, and has not been used to evaluate the cognitive load imposed upon SDMs of CCI patients who are exposed to decision support. In its limited application, the CLS has been used within classroom settings to evaluate the effectiveness of educational materials and teaching styles. The evaluation of cognitive load imposed by decision support interventions is crucial in that it is hypothesized that individuals who experience undesirable states of cognitive load will demonstrate ineffective learning, potentially weakening the efficacy of the decision support intervention (Chandler & Sweller, 1991).

PURPOSE

Therefore, this psychometric study examines reliability and validity of the Cognitive Load Scale in a sample of surrogate decision makers who are exposed to one of two electronic decision support interventions.

BACKGROUND AND CONCEPTUAL FRAMEWORK

Cognitive load, defined as a “multidimensional construct representing the load that performing a particular task imposes on the learner’s cognitive system” (Paas, Tuovinen, Tabbers, Van Gerven, & Gerven, 2003, p. 64), was first described by Sweller (1988), who recognized that particular learning strategies consume a disproportionate amount of cognitive resources, hindering the learning process. Cognitive load theory attributes this process to humans’ limited working memory capacity (Paas et al., 2003). Cognitive Load Theory recognizes potential sources of ICL and ECL. Since ICL represents the cognitive load imposed by the difficulty of the learned subject matter, ICL is dependent on the knowledge of the learner. Moreover, ICL reflects the *element interactivity* of the material being presented. Element interactivity also depends on the prior knowledge of the learner, but also reflects the ontological organization of the subject matter and the relationships of the interacting elements. Low element interactivity, resulting in low ICL, reflects simple elements to the learned material that can be learned in isolation. Conversely, element interactivity is increased when learned materials are presented in a way such that they can only be understood when in relation to other elements. Sources of ECL stem from the learner being exposed to information that is not necessary for learning. Jong (2009) discusses several potential sources of ECL.

Measuring Cognitive Load

Initially, cognitive load was evaluated indirectly by observing problem-solving strategies, learning time, and error rates (Ayres, 2001; Ayres & Sweller, 1990; Sweller, Chandler, Tierney, & Cooper, 1990). Furthermore, subjective measurements of mental effort, mental workload, and learning efficiency were used as proxy measures of the cognitive load experience (Gerjets, Scheiter, & Catrambone, 2006; Van Gog & Paas, 2008; Hart & Staveland, 1988; Kester, Lehnert, & Gerven, 2006). However, due to questioning of validity and conceptual ambiguity, these measurement techniques fell out of favor to distinct measures of cognitive load. The first measures of cognitive load were commonly single-item instruments evaluating one or more types of cognitive load (Ayres, 2006; DeLeeuw & Mayer, 2008). Eventually, a 10-item, psychometrically sound instrument was developed by Leppink, Paas, Van der Vleuten, Van Gog, and Van Merriënboer (2013). This instrument was tested, refined, and re-introduced by Leppink and Heuvel (2015).

The refined Cognitive Load Scale (CLS) introduced by Leppink and Heuvel (2015) is eight items and possesses 2 four-item subscales measuring ICL and ECL, respectively. It is administered after the completion of a learning activity, as all the questions on the CLS relate to the perceived learning experience of the learner. Individuals rate the extent they agree with each question on a scale from 0 (not at all the case) to 10 (completely the case). A total score is not calculated; instead, subscale scores are calculated by summing the

individual responses from each subscale item, with higher scores indicating a greater degree of cognitive load. Psychometric evaluation of the refined eight-item CLS has not been reported; thus, this will be the first known study to report the psychometric properties of the Leppink and Heuvel (2015) CLS. However, the initial version of the 10-item CLS demonstrates adequate goodness of fit indices ($\chi^2 = 36.89$, $p = .25$; RMSEA = 0.04) and internal consistency reliability (Cronbach's $\alpha > .80$) (Hadie & Yusoff, 2016; Leppink et al., 2013). Moreover, another cognitive load scale developed by Sewell et al. (2016) demonstrated acceptable psychometric properties.

METHODS

Design

This psychometric evaluation used data generated from a randomized, controlled trial of SDMs of decisionally impaired CCI patients who were receiving two types of decision support. One decision support intervention, Information Support (IS), produced a passive experience, consisting of videos related to communicating with healthcare providers. The alternative intervention, Interactive Virtual Decision Support for End of Life and Palliative Care (INVOLVE), produced an avatar-based, experiential-based learning experience which taught a communication strategy to the user and provided an opportunity for the user to practice the communication strategy in a simulated experience one might encounter within the ICU. Upon completion of informed consent, participants were allocated to a study group (control, IS, or INVOLVE) through a minimization allocation procedure (Scott, McPherson, Ramsay, & Campbell, 2002). To ensure balance among the three study conditions, participants were allocated to a study condition according to three factors: sex (male/female), relationship to patient (spouse, non-spouse), and race (White, non-White). Participants were administered a battery of psychosocial instruments, which included the Decision Fatigue Scale, the Emotion Regulation Questionnaire, the Family Decision Making Self-Efficacy Scale, the Preparation for Decision Making Scale, and the National Institutes of Health (NIH) Toolbox Flanker Inhibitory Control and Attention Test (Flanker).

Sample

A convenience sample of 62 SDMs were recruited from four different ICUs (cardiac, medical, neuroscience, and surgical) at an academic medical center in Northeast Ohio. All participants were: (1) aged 18 or older, (2) able to understand English, (3) recognized by the ICU team as the next of kin or legal representative for healthcare decision making for a decisionally impaired patient requiring at least three consecutive days of acute mechanical ventilation. Surrogate decision makers were excluded if: (1) the critically ill patient was not expected to be in the ICU for two days past study enrollment or (2) the surrogate decision maker could not hear audio using a standard set of headphones and/or unable to view the decision support material on a 10-inch computer screen.

Instruments

An objective measure of executive function (i.e., attention) and subjective measures of decision fatigue, emotion regulation, decision-making self-efficacy, and preparation for decision making were used in part to assess construct validity of the CLS with

psychometrically established instruments. Demographic characteristics of SDMs and patients were captured through investigator-developed forms.

Decision fatigue scale—Decision fatigue describes the impairments experienced while making decisions that result from performing acts of self-regulation and/or making decisions (Hickman, Pignatiello, & Tahir, 2017). The decision fatigue scale, developed by Hickman, Pignatiello, and Tahir (2017) is a 10-item instrument that captures the subjective experience of decision fatigue over a 24-hour period. Each item consists of a four-item Likert scale ranging from 0 (*strongly disagree*) to 3 (*strongly agree*). The individual scores from each item are totaled, providing a total decision fatigue score, with higher scores indicating greater levels of decision fatigue. The decision fatigue scale demonstrates adequate construct validity, internal consistency reliability ($\alpha = .87$), and stability reliability ($r = .56, p = < .001$) across eight weeks among SDMs of critically ill patients (Hickman et al., 2017). In this sample, the DFS demonstrated acceptable internal consistency reliability ($\alpha = .82$).

Family decision making self-efficacy (FDMSE) scale—The FDMSE scale is a 13-item self-report instrument that has individuals rate their confidence in making certain types of decisions for loved ones who cannot make healthcare decisions on their own. Items are scored using a five-point Likert scale from 1 (*cannot do at all*) to 5 (*certain I can do*). Item totals are summed for a total score, with higher scores indicating a greater degree of decision-making self-efficacy. The FDMSE scale demonstrates adequate validity and reliability among surrogate decision makers of amyotrophic lateral sclerosis (ALS) patients (Nolan et al., 2009). In this sample, the FDMSE demonstrated good internal consistency reliability ($\alpha = .88$).

Preparation for decision making (PrepDM) scale—The PrepDM evaluates perceived utility of decision support interventions. This scale consists of 10 items, using a five-point Likert scale ranging from 1 (not at all) to 5 (a great deal). Furthermore, the PrepDM demonstrates acceptable validity and reliability among adults receiving conversational, audio, print, or video-based decision aids for conditions like hormone-replacement therapy, prostate cancer, and orthopedic procedures (Bennett et al., 2010). In this sample, the PrepDM demonstrated excellent internal consistency ($\alpha = .91$).

Flanker inhibitory control and attention test (Flanker)—The Flanker task tests individuals' ability to regulate their attention by challenging their ability to ignore irrelevant task dimensions. Each trial displays a center arrow facing either left or right, with two arrows on each side that all face either the same or opposite direction as the central arrow. The goal of the task is to correctly indicate which direction the middle arrow is pointing. A scoring algorithm generates a total score based on the subject's accuracy and reaction time. Accuracy is scored on a five-point scale, with 0 indicating no correct responses and 5 indicating perfect accuracy, and reaction time is measured in milliseconds. The Flanker has established validity and reliability in a population of healthy adults. It demonstrates adequate convergent and discriminant validity when compared to other neurocognitive tasks such as the Delis-Kaplan Executive Function Scales ($r(229) = .52$) and the Peabody Picture Vocabulary Test, 4th edition ($r(234) = .06$; in addition, the Flanker demonstrates excellent

stability reliability ($r(237) = .85$ (Dunn, L. & Dunn, L., 1981; Swanson, 2005; Zelazo et al., 2014).

Procedures

Prior to study recruitment, institutional review board (IRB) approval was obtained. Participant screening consisted of two phases that occurred 5 to 7 days a week. The first phase involved the identification of all patients within the study ICUs who had received at least three days of mechanical ventilation. The second phase involved confirming the patient's lack of decisional capacity and the identification of the SDM, which was achieved by communicating with members of the ICU team (e.g., registered nurses, medical residents, attending physicians). Once lack of capacity was confirmed and the SDM identified, the research assistant approached the SDM for written informed consent. Upon attainment of written informed consent, participants were assigned to one of three study treatment arms and the research assistant began the interview.

All subjective data were collected through structured face-to-face interviews. Each interview was approximately 30–40 minutes. Following the interview, the subject was exposed to either the passive or experiential decision support intervention. Allocation of decision support was determined through minimization techniques. Following the receipt of the decision support intervention, subjects completed the CLS. Data was directly entered by the research assistant into the Research Electronic Data Capture (REDCap) program and subsequently exported to the Statistical Package for the Social Sciences (SPSS, Version 24) for data analysis (Harris, Taylor, Thielke, & Payne, 2009).

Validity and Reliability Assessment

Our statistical approach involved (1) evaluation of descriptive statistics (e.g., frequencies, means, standard deviations) and chi-square analyses to describe and compare sample characteristics and endorsement of scale items; (2) conducting an exploratory factor analysis and evaluating bivariate correlation coefficients among subscale scores of intrinsic and extraneous cognitive load with the Decision Fatigue Scale, Emotion Regulation Scale, Family Decision Making Self-Efficacy Scale, Preparation for Decision Making Scale, and the Flanker task to determine construct validity; and (3) evaluating the internal consistency reliability through interpretation of Cronbach's alpha coefficient.

Construct validity—To evaluate construct validity, we conducted statistical evaluation of factorial and discriminant validity. Factorial validity compares the theoretical and empirical organization of the instrument, and was evaluated through conduct of an exploratory factor analysis (EFA). The CLS was constructed to capture a bidimensional construct; however, because this is the first time that the CLS has been evaluated in this context, the authors conducted an EFA using principal axis factoring extraction with direct oblimin rotation. This provides clarity regarding the interrelatedness of the instrument's items within its latent factors.

Factorial validity—Factorial validity was assessed through EFA. To ensure psychometric quality of the EFA, an assessment of sampling adequacy was performed prior to factor

structure evaluation. The Kaiser-Meyer-Meyer-Olkin (KMO) test [$KMO = 0.841$] and Bartlett's test of sphericity [$\chi^2 = 287.4$, $p < .001$] were used to verify the adequacy of the 7:1 ratio of participants ($N = 62$) to the total number of instrument items ($N = 8$). Based on these statistics, a sample size of 62 was deemed adequate to proceed with the EFA.

Specification of factor structure: The authors inspected the Cattell scree plot and applied the K1 method (Kaiser, 1960), which specifies that factors with eigenvalues greater than one be maintained for interpretation.

Item retention: Items with primary factor loading coefficients ≥ 0.40 and items with secondary factor loadings ≥ 0.20 were retained. Items that did not meet the retention criteria or demonstrated a primary factor loading on more than one factor were sequentially removed. Once an item was removed, the EFA was repeated until a parsimonious factor structure was achieved that met the a priori criteria.

Discriminant validity—Discriminant validity assumes that constructs that are conceptually dissimilar will not possess a strong statistical relationship. Discriminant validity was evaluated after identification of the most parsimonious factor structure through the EFA. It was assumed that discriminant validity would be confirmed if the bivariate correlations between measures of intrinsic and extraneous cognitive load with measures of emotion regulation, decision fatigue, decision making self-efficacy, decision-making preparation, and attention were $|r| < 0.70$.

Internal consistency reliability—The internal consistency reliability of the CLS was determined by evaluating Cronbach's alpha coefficient. For this study, sufficient internal consistency reliability was designated by an alpha coefficient ≥ 0.70 .

RESULTS

Sample Characteristics

This study sample consisted of two groups of SDMs for critically patients who received different types of decision support. The two types of decision support were information support (IS), which was decision support delivered through a video providing education on how to effectively communicate with a physician, and INVOLVE, which was an avatar-based tablet application that taught a communication strategy and provided simulated interaction with an ICU team during rounds to practice the communication strategy.

The subjects were evenly assigned to one of the two study conditions (IS, $n = 31$; INVOLVE, $n = 31$). For SDMs assigned to IS condition, the mean age was 54 years ($SD = 13.8$), 72% were female, 78% identified as White, 34% were the patient's spouse, and 56% had not made healthcare decisions for the patient before the current hospitalization. For SDMs assigned to the INVOLVE condition, the mean age was 51 ($SD = 15.03$), 71% were female, 77% identified as White, 29% were the spouse/partner, 42% were adult children, and 58% had not made healthcare decisions for the patient before the current hospitalization. Of note, sample demographic characteristics (Table 1) did not significantly differ between groups except for education level ($\chi^2(2) = 7.56$, $p = .02$). When administered the CLS, IS

group ($n = 31$) scores of ICL ($M = 1.8$, $SD = 2.5$) and ECL ($M = 0.52$, $SD = .93$) were not significantly different ($t = .72$, $p = .47$; $t = 1.47$, $p = .15$) than scores of ICL ($M = 2.3$, $SD = 2.8$) and ECL ($M = 1.1$, $SD = 1.9$) reported by those in the INVOLVE group.

Descriptive Scale and Item Statistics

Descriptive scale and item statistics were calculated for the eight-item CLS (Table 2). The mean scores for ICL and ECL ($N = 62$) were 2.0 ($SD = 2.6$) and 0.80 ($SD = 1.5$), respectively. Scores for ICL and ECL ranged from 0 to 10 and 0 to 7, respectively. Mean item scores across the ICL and ECL subscales ranged from 1.7 to 2.6 and 0.40 to 1.4, respectively. Total scores of ICL and ECL possessed a relatively normal distribution (skewness = 1.4; 2.4 and kurtosis = 1.2; 5.8), respectively. While these scores are distributed within the lower range of the possible total score, our results are similar with another application of cognitive load theory within healthcare. Sewell et al. (2016) reported the psychometric properties of a Cognitive Load Inventory for Colonoscopy that reported a range of mean scores for ICL (mean range 1.21–3.14; SD range 1.27–1.78) and ECL (mean range 0.91–1.83; SD range 1.32–1.80).

Evaluation of Construct Validity and Internal Consistency Reliability

Factorial validity—To evaluate factorial validity, we conducted an EFA. Consistent with the original construction of the CLS, our EFA yielded a two-factor structure in which the items representing intrinsic and extraneous cognitive load loaded together (Table 3). Notably, item 1 of the original CLS, “The content of this activity was very complex,” demonstrated a primary loading ($-.51$) with factor 2, and a secondary loading with factor 1 (.39). The two factors of the CLS had eigenvalues of 4.5 and 1.2, and accounted for 56% and 16% of the observed variance, respectively. Consistent with the original factor structure of the CLS, factor 1 contained items that operationalized ECL and factor 2 operationalized ICL.

Discriminant validity—The CLS possesses adequate discriminant validity (Table 4). The ICL and ECL subscales correlated with established measures of decision fatigue ($r = .32$, $p = .01$; $r = .30$, $p = .02$), family decision-making self-efficacy ($r = -.21$, $p = .10$; $r = -.29$, $p = .03$), preparation for decision making ($r = -.28$, $p = .03$; $r = -.43$, $p = .001$), and the Flanker task ($r = -.15$, $p = .25$; $r = -.32$, $p = .02$), respectively.

Internal consistency reliability—Internal consistency reliability of the CLS was confirmed through the interpretation of the Cronbach’s alpha coefficient. Calculation of Cronbach’s alpha yielded a coefficient of .89 for the ICL subscale and .75 for the ECL subscale. Per our a priori designation of an acceptable Cronbach’s alpha being .70, the CLS demonstrates adequate internal consistency.

DISCUSSION

To our knowledge, this is the first known psychometric analysis of the CLS within a sample of surrogate decision makers of critically ill patients; moreover, on a broader level, this is the second known application and comparison of cognitive load within recipients of healthcare

education. Despite the marked contrast between this study's population, SDMs of critically ill patients, and past environments in which cognitive load has been measured (i.e., students within educational settings), the CLS demonstrates adequate validity and reliability (Leppink et al., 2013; Leppink, Paas, Van Gog, van der Vleuten, & van Merriënboer, 2014; Wang, Chen, & Wu, 2016). Specifically, the CLS maintained the a priori specified factor structure (i.e., ICL and ECL), discriminated appropriately with decision fatigue, decision-making self-efficacy, decision-making preparation, and cognitive inhibition, and demonstrated adequate internal consistency reliability.

Educational level between the IS and INVOLVE groups were significantly different from one another. Consistent with cognitive load theory, one would expect education level to be a significant determinant of cognitive load when exposed to an educational intervention. Nonetheless, the underlying factor structure of the CLS should not adversely affected. Our results support this conclusion as the CLS demonstrated a bidimensional factor structure.

Consistent with the factor structure specified by Leppink & van den Heuvel (2015), item 1 of the CLS, "The content of this activity was very complex," demonstrated a primary loading ($-.53$) with factor 2, which represents ICL. However, it also demonstrated a secondary loading ($.39$) with factor 1, the representation of ECL. Similarly, item 2 of the CLS, "The problems covered in this activity were very complex," demonstrated an expected primary loading ($-.57$) with factor 2; however, it also demonstrated a secondary loading with factor 1 ($.31$). Thus, our results indicate that items 1 and 2 of the CLS could potentially serve as a representation of both ICL and ECL despite its intended representation of solely ICL. To understand this theoretical inconsistency, it is important to consider the two types of educational interventions the participants received. The first type, Information Support (IS), was an educational-based video that discussed how to effectively communicate with healthcare providers. The second type of decision support, INVOLVE, was a tablet-based application that taught a communication strategy, allowed for simulated practice of the strategy, and provided an interactive activity to elicit preferences concerning an end-of-life decision. Thus, it is possible that the differences in cognitive load elicited by the variation in the educational modality served as the impetus for the loading demonstrated within this sample. However, while probably underpowered to detect such differences, our data indicated no significant differences in ICL or ECL among those exposed to IS or INVOLVE.

Alternatively, it is possible the theoretical inconsistency is indicative of item 1's and 2's ambiguity; specifically, when the item references the "complexity" of the activity in question. For example, it is possible that when considering how to respond to items 1 and 2, participants viewed "complexity" in varying ways, such that it was appraised in terms of the difficulty of the subject matter representing ICL. However, the way in which educational material is delivered can influence the perceived complexity of the activity; therefore, if participants appraised complexity in that sense, they were describing ECL. If this theoretical redundancy is reported in subsequent populations, one might consider revising the scale such that items 1 and 2 more clearly captures the ICL experience (e.g., "The content of the activity was hard to understand" and "The problems covered in this activity were hard to understand").

In this sample, score distributions of ICL ($M = 2.0$, $SD = 2.6$) and ECL ($M = 0.80$, $SD = 1.5$) were low when considering the scoring range of the eight-item CLS (i.e., 0–10). These results are similar to another psychometric evaluation of a healthcare related cognitive load scale (Sewell et al., 2016). Because the reported levels of cognitive load were relatively low across two dissimilar healthcare populations (i.e., physicians performing colonoscopies vs. SDMs of the critically ill), it is possible that the scoring of both intrinsic and extraneous cognitive load needs to be altered to provide a greater level of cognitive load variance. However, the clinical implications associated with these “low” levels of cognitive load have not yet been described. Therefore, the authors recommend subsequent comparison of cognitive load scores to relevant healthcare behaviors and outcomes to determine the clinical significance of experiencing a particular degree of intrinsic or extraneous cognitive load. For example, within a population of SDMs for the critically ill, the varying levels of cognitive load should be compared to relevant SDM outcomes such as decision conflict and decision regret.

Furthermore, this study’s application of cognitive load within a new environmental domain (i.e., the intensive care unit) contributes to the overall development of cognitive load theory. As discussed by Choi, Van Merriënboer, and Paas (2014), the physical learning environment serves as a causal factor that interacts with both the learner and learning-task characteristics. Thus, it would not be unexpected if the same educational intervention yielded different levels of cognitive load in different physical environments. To date, empirical evidence describing cognitive load within healthcare settings is nascent; moreover, a large majority of available reports report levels of cognitive load using antiquated cognitive load scales and proxy measures (Korbach, Brünken, & Park, 2017; Schmeck, Opfermann, van Gog, Paas, & Leutner, 2015). Thus, comparison of ICL and ECL within the ICU, let alone any other learning environment, is limited. Nevertheless, this investigation serves as a preliminary comparison for subsequent application of Cognitive Load Theory within health care.

Relevance to Nursing Science and Practice

Cognitive load possesses strong implications for future healthcare self-management and behavioral science research. A major area of scientific focus for the National Institute of Nursing Research (NINR) addresses the need to personalize care among those who possess a chronic condition to ensure effective self-management (NINR, 2016). With knowledge being a key determinant of self-management outcomes (Ryan & Sawin, 2009), evidence that elucidates the best way to provide knowledge in a way that promotes wellness is a major topic of concern for many chronic conditions, including hypertension (Khatib et al., 2014), diabetes (Pal et al., 2013), and obesity (Okorodudu, Bosworth, & Corsino, 2015). Thus, cognitive load theory serves as a novel and promising construct as it specifically accounts for learner, learning-task, and environmental characteristics that may enhance or diminish the efficacy of healthcare self-management interventions. In addition to its application for patients and their family members, cognitive load theory also has promise to inform educational provision for clinicians (Leppink & Heuvel, 2015).

Limitations

This study possesses several limitations. First, while the KMO and Bartlett's test indicated that our sample ($N = 62$) was adequate to power this psychometric analysis, our analysis was most likely underpowered to determine whether the observed theoretical inconsistencies (i.e., item 1 demonstrating a strong secondary loading into the ECL subscale) were resultant of the differences in cognitive load elicited by the two educational interventions (i.e., IS and INVOLVE). However, while one could hypothesize that there could be differences in cognitive load among the two groups related to the unique attributes of the decision support interventions, the theoretical structure of the eight-item CLS should nonetheless remain consistent, regardless of study condition. Nevertheless, the authors recommend further psychometric analyses of the CLS among samples receiving a homogeneous educational experience to ensure that the factor structure of the CLS remain consistent across educational settings and modalities.

Moreover, in this study, cognitive load was measured at one point upon completion of the learning activity. However, prior evidence suggests that cognitive load values significantly differ from one another when it is calculated once at the end of task compared to the average of cognitive load values taken over the course of a learning activity (Schmeck et al., 2015; van Gog, Kirschner, Kester, & Paas, 2012). Therefore, the methods used to measure cognitive load in this study limit the generalizability of findings, yet provide direction for future evaluation of cognitive load during the provision of healthcare decision support and education. Finally, interpretation of the CLS's construct validity was limited by the lack of available triangulation measures. Notably, *pupillometry*, the measurement of pupil size and reactivity, has demonstrated promise as a physiologic indicator of cognitive load, and may serve as a useful comparison for determination of the CLS construct validity (Piquado, Isaacowitz, & Wingfield, 2010).

CONCLUSION

This study indicates that an eight-item CLS (Cognitive Load Scale) is a valid and reliable instrument for measuring cognitive load within recipients of decision support education. Although the CLS has established reliability and validity in our sample of surrogate decision makers, we recommend subsequent psychometric evaluations of the CLS in similar and varying populations with objective measures of neurocognition that are associated with aspects of human learning, such as working memory. Nevertheless, the CLS has theoretical underpinnings in Cognitive Load Theory, which provides an emerging framework to guide the future development, administration, and evaluation of educational interventions in nursing and other health disciplines.

Acknowledgments

The authors would like to thank Matthew McManus for his editorial assistance related to the completion of this work.

Financial Disclosure:

This publication was made possible by funding from the National Institute of Nursing Research (NINR; NR014213), a component of the National Institutes of Health (NIH), as well as the Robert Wood Johnson

Foundation (RWJF) Nurse Faculty Scholars Program (#72118). Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the NIH, NINR, or the RWJF.

References

- Ayres, P. Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*. 2006. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0959475206000703>
- Ayres PL. Systematic Mathematical Errors and Cognitive Load. *Contemporary Educational Psychology*. 2001; 26(2):227–248. <https://doi.org/10.1006/ceps.2000.1051>. [PubMed: 11273658]
- Ayres P, Sweller J. Locus of difficulty in multistage mathematics problems. *The American Journal of Psychology*. 1990:167–193.
- Bennett C, Graham ID, Kristjansson E, Kearing SA, Clay KF, O'Connor AM. Validation of a Preparation for Decision Making scale. *Patient Education and Counseling*. 2010; 78(1):130–133. <https://doi.org/10.1016/j.pec.2009.05.012>. [PubMed: 19560303]
- Chandler P, Sweller J. Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*. 1991; 8(4):293–332. https://doi.org/10.1207/s1532690xc0804_2.
- Choi H-H, Van Merriënboer JIG, Paas F. Effects of the Physical Environment on Cognitive Load and Learning: Towards a New Model of Cognitive Load. *Educational Psychology Review*. 2014; 26(2): 225–244. <https://doi.org/10.1007/s10648-014-9262-6>.
- DeLeeuw KE, Mayer RE. A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*. 2008; 100(1):223.
- Dunn, LM., Dunn, LM. Manual for the peabody picture vocabulary test-revised. Circle Pines, MN: American Guidance Service; 1981.
- Gerjets P, Scheiter K, Catrambone R. Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction*. 2006; 16(2):104–121.
- Hadie SN, Yusoff MS. Assessing the validity of the cognitive load scale in a problem-based learning setting. *Journal of Taibah University Medical Sciences*. 2016; 11(3):194–202.
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics*. 2009; 42(2):377–381. [PubMed: 18929686]
- Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*. 1988; 52:139–183.
- Hickman R, Pignatiello G, Tahir S. Evaluation of the decisional fatigue scale among surrogate decision makers of the critically ill. *Western Journal of Nursing Research*. 2017 In Press.
- Kester L, Lehnen C, Van Gerven PW, Kirschner PA. Just-in-time, schematic supportive information presentation during cognitive skill acquisition. *Computers in Human Behavior*. 2006; 22(1):93–112.
- Khatib R, Schwalm J-D, Yusuf S, Haynes RB, McKee M, Khan M, Nieuwlaar R. Patient and Healthcare Provider Barriers to Hypertension Awareness, Treatment and Follow Up: A Systematic Review and Meta-Analysis of Qualitative and Quantitative Studies. *PLoS ONE*. 2014; 9(1):e84238. <https://doi.org/10.1371/journal.pone.0084238>. [PubMed: 24454721]
- Korbach A, Brünken R, Park B. Measurement of cognitive load in multimedia learning: a comparison of different objective measures. *Instructional Science*. 2017; 45(4):515–536. <https://doi.org/10.1007/s11251-017-9413-5>.
- Leppink J., Gog, T., Paas, F., Sweller, J. *Researching Medical Education*. Chichester, UK: John Wiley & Sons, Ltd.; 2015. Cognitive load theory: researching and planning teaching to maximise learning; p. 207-218. <https://doi.org/10.1002/9781118838983.ch18>
- Leppink J, Paas F, Van der Vleuten CPM, Van Gog T, Van Merriënboer JIG. Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*. 2013; 45(4):1058–1072. <https://doi.org/10.3758/s13428-013-0334-1>. [PubMed: 23572251]

- Leppink J, Paas F, van Gog T, van der Vleuten CPM, van Merriënboer JG. Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*. 2014; 30:32–42. <https://doi.org/10.1016/j.learninstruc.2013.12.001>.
- Leppink J, van den Heuvel A. The evolution of cognitive load theory and its application to medical education. *Perspectives on Medical Education*. 2015; 4(3):119–127. <https://doi.org/10.1007/s40037-015-0192-x>. [PubMed: 26016429]
- National institute of Nursing Research. The NINR Strategic Plan: Advancing Science, Improving Lives. Bethesda, Maryland: 2016. Retrieved from https://www.ninr.nih.gov/sites/www.ninr.nih.gov/files/NINR_StratPlan2016_reduced.pdf
- Nolan MT, Hughes MT, Kub J, Terry PB, Astrow A, Thompson RE, Sulmasy DP. Development and validation of the family decision-making self-efficacy scale. *Palliative & supportive care*. 2009; 7(3):315–321. [PubMed: 19788773]
- Okorodudu DE, Bosworth HB, Corsino L. Innovative interventions to promote behavioural change in overweight or obese individuals: A review of the literature. *Annals of Medicine*. 2015; 47(3):179–185. <https://doi.org/10.3109/07853890.2014.931102>. [PubMed: 25011006]
- Paas F, Tuovinen JE, Tabbers H, Van Gerven PWM, Van Gerven PWM. Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*. 2003; 38(1):63–71. https://doi.org/10.1207/S15326985EP3801_8.
- Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, Murray E. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. *Cochrane Database of Systematic Reviews* (Online). 2013; 3:CD008776. <https://doi.org/10.1002/14651858.CD008776.pub2>.
- Pignatiello, G., Hickman, RL., Hetland, B. End-of-Life Decision Support in the ICU: Where Are We Now?. *Western Journal of Nursing Research*. 2016. <https://doi.org/10.1177/0193945916676542>
- Piquado T, Isaacowitz D, Wingfield A. Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*. 2010; 47(3):560–9. <https://doi.org/10.1111/j.1469-8986.2009.00947.x>. [PubMed: 20070575]
- Ryan P, Sawin KJ. The Individual and Family Self-Management Theory: background and perspectives on context, process, and outcomes. *Nursing Outlook*. 2009; 57(4):217–225.e6. <https://doi.org/10.1016/j.outlook.2008.10.004>. [PubMed: 19631064]
- Schmeck A, Opfermann M, van Gog T, Paas F, Leutner D. Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. *Instructional Science*. 2015; 43(1):93–114. <https://doi.org/10.1007/s11251-014-9328-3>.
- Scott NW, McPherson GC, Ramsay CR, Campbell MK. The method of minimization for allocation to clinical trials. a review. *Controlled Clinical Trials*. 2002; 23(6):662–74. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12505244>. [PubMed: 12505244]
- Sewell JL, Boscardin CK, Young JQ, ten Cate O, O'Sullivan PS. Measuring cognitive load during procedural skills training with colonoscopy as an exemplar. *Medical Education*. 2016; 50(6):682–692. <https://doi.org/10.1111/medu.12965>. [PubMed: 27170086]
- Swanson J. The Delis-Kaplan Executive Function System: A Review. *Canadian Journal of School Psychology*. 2005; 20(1–2):117–128. <https://doi.org/10.1177/0829573506295469>.
- Sweller J. Cognitive load during problem solving: Effects on learning. *Cognitive Science*. 1988; 12(2): 257–285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7).
- Sweller J, Chandler P, Tierney P, Cooper M. Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*. 1990; 119(2):176–192. <https://doi.org/10.1037/0096-3445.119.2.176>.
- Van Gog T, Kirschner F, Kester L, Paas F. Timing and frequency of mental effort measurement: Evidence in favour of repeated measures. *Applied Cognitive Psychology*. 2012; 26(6):833–839. <https://doi.org/10.1002/acp.2883>.
- Van Gog T, Paas F. Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*. 2008; 43(1):16–26.
- Wang, WF., Chen, CM., Wu, CH. Effects of Different Video Lecture Types on Sustained Attention, Emotion, Cognitive Load, and Learning Performance; *Proceedings - 2015 IIAI 4th International*

Congress on Advanced Applied Informatics, IIAI-AAI 2015. 2016. p. 385-390.<https://doi.org/10.1109/IIAI-AAI.2015.225>

Zelazo PD, Anderson JE, Richler J, Wallner-Allen K, Beaumont JL, Conway KP, Weintraub S. NIH Toolbox Cognition Battery (CB): Validation of Executive Function Measures in Adults. *Journal of the International Neuropsychological Society*. 2014; 20:1–10. <https://doi.org/10.1017/S1355617714000472>. [PubMed: 24365485]

Zilberberg MD, de Wit M, Pirone JR, Shorr AF. Growth in adult prolonged acute mechanical ventilation: implications for healthcare delivery. *Critical Care Medicine*. 2008; 36(5):1451–1455. <https://doi.org/10.1097/CCM.0b013e3181691a49>. [PubMed: 18434911]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- Cognitive Load Theory is a promising framework to inform healthcare intervention development
- The eight-item cognitive load scale is both valid and reliable
- Further measurement and application of cognitive load is necessary within varying healthcare settings

TABLE 1

Sample Characteristics by Study Group (N = 62)

Variables	INVOLVE (<i>n</i> = 31) <i>n</i> (%)	IS (<i>n</i> = 31) <i>n</i> (%)
Gender		
Female	22 (71)	22 (71)
Male	9 (29)	9 (29)
Race/Ethnicity		
White	24 (77)	24 (77)
Non-White	7 (23)	7 (23)
Relationship to Patient		
Spouse	9 (29)	10 (33)
Adult Child	13 (42)	9 (29)
Sibling	4 (13)	5 (16)
Power of Attorney	3 (10)	2 (7)
Other	2 (6)	5 (15)
Education		
High School or Less	5 (16)	14 (45)
1–4 Years of College	16 (52)	14 (45)
Graduate Studies	10 (32)	3 (10)
Employment Status		
Employed	22 (71)	15 (48)
Retired/Disabled/Unemployed	9 (29)	14 (45)
Other	0 (0)	2 (7)
Annual Household Income ^a		
\$20,000 and less	2 (6)	4 (13)
\$21,000 to \$49,999	7 (23)	15 (48)
\$50,000 and greater	21 (68)	12 (39)
Marital Status: Yes	18 (58)	20 (65)
Prior Decision-Making Experience: Yes	13 (42)	14 (45)

^a *n* = 30 for INVOLVE group.

TABLE 2

Item Statistics for the Eight-Item Cognitive Load Scale (N = 62)

Scale Items	<i>M</i>	<i>SD</i>
1. The content of this activity was very complex.	1.7	2.6
2. The problems covered in this activity were very complex.	2.1	2.9
3. In this activity, very complex terms were mentioned.	1.9	3.1
4. I invested a very high mental effort in the complexity of this activity.	2.6	3.4
5. The explanations and instructions in this activity were very unclear.	.53	1.5
6. The explanations and instructions in this activity were full of unclear language.	.40	1.1
7. The explanations and instructions in this activity were, in terms of learning, very ineffective.	1.4	2.7
8. I invested a very high mental effort in unclear, ineffective explanations and instructions in this activity.	.86	2.1

TABLE 3

Bidimensional Factor Structure of the Eight-Item Cognitive Load Scale (N = 62)

Items	Factor 1	Factor 2
5. The explanations and instructions in this activity were very unclear.	.96	.03
8. I invested a very high mental effort in unclear, ineffective explanations and instructions in this activity.	.76	-.14
6. The explanations and instructions in this activity were full of unclear language.	.65	.02
7. The explanations and instructions in this activity were, in terms of learning, very ineffective.	.50	.03
4. I invested a very high mental effort in the complexity of this activity.	-.10	-.95
3. In this activity, very complex terms were mentioned.	-.03	-.92
2. The problems covered in this activity were very complex.	.31	-.57
1. The content of this activity was very complex.	.39	-.51

Note. Extraction method was principal axis factoring with oblique rotation (direct oblimin using the following parameters: $\delta = 0$, $\kappa = 4$).

Table 4

Correlations Among Variables Used for Discriminant Validity Testing (N = 62)

Variable	1	2	3	4	5	6
1. Intrinsic Cognitive Load	-					
2. Extraneous Cognitive Load	.56**	-				
3. Decision Fatigue	.32*	.30*	-			
4. Family Decision Making Self-Efficacy ^a	-.21	-.29*	-.46**	-		
5. Preparation for Decision Making ^b	-.28*	-.43**	-.33**	.24*	-	
6. Flanker Task ^c	-.15	-.32*	.05	-.11	.22	-

Note.

^a
n = 60;^b
n = 61;^c
n = 58.*
p < .05**
p < .01.