OPEN ACCESS

ESCOLA DE MEDICINA
PUCRS

SCIENTIA MEDICA

Scientia Medica Porto Alegre, v. 30, p. 1-8, jan.-dez. 2020
e-ISSN: 1980-6108 | ISSN-L: 1806-5562

http://dx.doi.org/10.15448/1980-6108.2020.1.36918

EDUCATION IN HEALTH SCIENCE

Revisiting cognitive load theory: second thoughts and unaddressed questions

Revisitando a teoria da carga cognitiva: pensamentos secundários e perguntas não endereçadas

Jimmie Leppink[1]
orcid.org/0000-0002-8713-1374
hyjl17@hyms.ac.uk

Recebido em: 23 jan. 2020.
Aprovado em: 09 fev. 2020.
Publicado em: 15 jul. 2020.

Abstract: In cognitive load theory (CLT), learning is the development of cognitive schemas in a long-term memory with no known limits and can happen only if our limited working memory can process new information presented and the amount of information that does not contribute to learning is low. According to this theory, learning is optimal when instructional support is decreased going from worked examples via completion problem to autonomous problem solving and learners do not benefit from practicing retrieval with complex content. However, studies on productive failure and retrieval practice have provided clear evidence against these two guidelines. In this article, issues with CLT and research inspired by this theory, which remain largely ignored among cognitive load theorists but have likely contributed to these contradictory findings, are discussed. This article concludes that these issues should make us question the usefulness of CLT in health science education, medical education and other complex domains, and presents recommendations for both educational practice and future research on the matter.

Keywords: Cognitive load theory; definitions; self-reports; retrieval practice; productive failure.

Resumo: Na teoria da carga cognitiva (CLT), a aprendizagem é o desenvolvimento de esquemas cognitivos em uma memória de longo prazo sem limites conhecidos e pode acontecer apenas se nossa limitada memória de trabalho puder processar novas informações apresentadas e a quantidade de informações que não contribui para a aprendizagem é baixo. De acordo com essa teoria, o aprendizado é ideal quando diminui o suporte instrucional, passando de exemplos trabalhados, via problemas de conclusão, para uma solução autônoma de problemas, e os alunos não se beneficiam praticando a recuperação com conteúdo complexo. No entanto, estudos sobre falhas produtivas e práticas de recuperação forneceram evidências claras contra essas duas diretrizes. Neste artigo, são discutidos problemas com a CLT e com pesquisas inspiradas nessa teoria, que permanecem amplamente ignorados entre os teóricos da carga cognitiva, mas provavelmente contribuíram para essas descobertas contraditórias. Este artigo conclui que essas questões devem nos fazer questionar a utilidade da CLT na educação em ciências da saúde, educação médica e outros domínios complexos e apresenta recomendações para a prática educacional e para pesquisas futuras sobre o assunto.

Palavras-chave: Teoria da carga cognitiva; definições; autorrelatos; prática de recuperação; falha produtiva.

ABREVIATIONS: CLT, Cognitive Load Theory; PF, productive failure RP, retrieval practice.

Introduction

Cognitive load theory (CLT) postulates that learning (1) is the development of cognitive schemas in long-term memory with no known limits and (2)

[1] University of York, York, North Yorkshire (NY), United Kingdom

can happen only if (i) information to be processed is within the narrow limits of our working memory and (ii) the amount of information that does not contribute to learning is minimised (e.g., [1-4]). This theory has resulted in a series of guidelines for the design of instruction in the context of learning complex content or procedures (e.g., [5-7]), including that (1) learning is optimal when instructional support is decreased going from worked examples via completion problems to autonomous problem solving and (2) learners do not benefit from practicing retrieval with complex content. However, as demonstrated in the next two paragraphs, research on retrieval practice (RP) and productive failure (PF) has provided clear evidence against both (1) and (2). After sharing key lessons from that research, this article discusses ontological and epistemological issues with CLT that have likely contributed to cognitive load theorists' inability to explain core findings from that research. Although these issues remain largely ignored among cognitive load theorists, they should make us question the usefulness of CLT in health science education, medical education and other complex domains. This article therefore concludes with a series of recommendations for both educational practice and future research on the matter.

Key lessons (1): Retrieval practice (RP)

Research on RP has consistently demonstrated that taking a memory test (i.e., RP) not only assesses what we know but also enhances retention, an effect that is also referred to as the *testing effect* (e.g., [8-10]). Although cognitive load theorists have stated that there is no testing effect (i.e., no benefit of RP) for complex content (e.g., [11]), Karpicke and Aue [8] indicated that a key finding from RP research has been that the testing effect is alive and well for complex content, and their response to Van Gog and Sweller [11], who claimed no testing effect for complex content, neatly summarizes some of the key flaws of research inspired by CLT.

To start, a core assumption in CLT is that information to be processed imposes a load on working memory, which is also referred to as *cognitive load*, and that load depends on how many new elements of information (i.e., not yet stored in cognitive schemas to be retrieved from long-term memory) must be processed as well as how these elements are interrelated (*element interactivity*). If the total number of new elements to be processed plus their interactions exceeds the narrow limits of working memory, *cognitive overload* occurs. The problem with the concept of element interactivity is that it is not defined in any measurable way, and consequently, cognitive load and overload are not defined in such a way either. Besides, although element interactivity is recognised as a key factor in cognitive load, it is rarely clear how element interactivity is manipulated in experiments inspired by CLT. And in many experiments, it may not be an important factor after all (e.g., memorising isolated words or single sentences). Finally, a common pitfall in research inspired by CLT is that small sample sizes leave researchers (very) unlikely to detect differences of a practically relevant magnitude (e.g., half a standard deviation) and yet researchers erroneously interpret statistically non-significant outcomes as evidence in favour of "no difference". Using Bayesian methods, which – contrary to null hypothesis significance testing – can help researchers to establish evidence in favour of one hypothesis relative to one or several other hypotheses, Karpicke and Aue [8] indicated that their small-scale meta-analysis provides substantial evidence in favour of a small positive testing effect relative to the null hypothesis of no testing effect.

Key lessons (2): Productive failure (PF)

A second key statement from CLT is that learning is optimal when instructional support is decreased going from worked examples via completion problems to autonomous problem solving. CLT predicts that novices will likely face cognitive overload if asked to engage in autonomous problem solving in a complex domain without studying worked examples first, and there are studies that appear to provide some evidence in favour of that prediction (e.g., [12-13]). However, participants in these studies

worked individually, and focussed on learning rules that would not be agreed as "complex" by everyone, such as learning how to apply basic rules from probability calculus to calculate a conditional probability. Besides, studies inspired by PF have provided evidence for the notion that, at least under some conditions, initial struggle with complex content in the absence of high instructional support (i.e., worked examples, or very detailed instructions making the problem easier) can benefit learning (e.g., [14-18]).

Although many prominent cognitive load theorists have waived away this finding by arguing that these studies mainly focussed on "low element interactivity" material and therefore CLT and PF could equally well explain the findings, the absence of measures of element interactivity does not facilitate this argument, and the materials reported in for example [15-17] are not any less complex (perhaps on the contrary: somewhat more complex) than the ones used in the studies that found evidence in favour of studying worked examples before solving problems autonomously (e.g., [12-13]).

A key factor that has remained largely ignored in research inspired by CLT is learning from peers, in dyads or small groups; experiments designed from a CLT perspective have almost exclusively focussed on participants learning individually, and often so in laboratory settings in which the participants did not really have any stake in the outcome (e.g., no course in biology, programming or probability calculus coming up next). Yet, based on the literature on PF thus far, it appears that learning from peers may constitute a critical factor in PF. It is therefore surprising that most cognitive load theorists continue to dismiss the work on PF as focussing on "low element interactivity" content only, whatever that means given the lack of a clear definition and good measure of element interactivity, and that even in a recent proposal to move from CLT to collaborative CLT [19] there is no single mention of PF. Given that the apparent contradiction between findings from PF research and predictions made by CLT has been discussed at several platforms before, including by prominent cognitive load theorists (e.g., [14]), one would expect at least some consideration of future research comparing individual learning and learning from peers to see where and why CLT and PF provide different predictions and which ones are more likely under which conditions.

Hardcore cognitive load theorists state that there is more than half a century of research literature supporting direct instruction over more constructivist approaches such as PF, but most of the research indicating a preference towards direct instruction is based on laboratory studies quite isolated from everyday educational practice, involving small samples of participants studying content of questionable complexity individually without having a stake in the outcome of the experiment. However, in settings where the nature of tasks and problems, and professionals' roles and responsibilities with it, are dynamic and ever-evolving – such as security, emergency medicine, aviation, mental health, and engineering [20] – professionals have to be willing and able to learn new content and skill all the time and direct instruction may often not be an option but PF as in learning from peers may be critical. In laboratory settings where undergraduate students learn how to apply a multiplication rule to calculate a probability, cognitive overload may never occur; if a participant gets bored or thinks the problem is too difficult or not worth the investment, there may hardly be any cognitive load at all. However, in high-stakes settings like the ones just mentioned, cognitive overload will at times pose a real threat, and a lack of willingness to invest in a task or problem can have grave consequences for human lives. Even small positive effects of RP and/or PF may in such settings make a difference between life and death.

The direct instruction advocated by CLT and (most of) its followers is not without problems. To start, a lack of prior knowledge may hinder learners to understand complex problems, how they manifest, how they can be represented in a way that we can approach and try to solve them, and/or methods to solve these problems. Besides, when these problems are presented in an artificially (well-)structured manner, learners may

not come to fully understand the nature of these problems, how they manifest, how they can be represented in a way that we can approach and try to solve them, and what methods we can use to solve these problems under what conditions. PF aims to circumvent these problems by having students generate and explore the potential and limitations of different representations of a type of problem – say Type X – and methods to solve Type X (i.e., Phase 1) to then provide them with opportunities to establish useful rules for representing and solving Type X (i.e., Phase 2). When we design learning and practice tasks around Type X that are of an appropriate level of complexity, in a context that is challenging (though not frustrating), Phase 1 can help learners to activate and apply prior knowledge of concepts that are important to understand Type X, to draw attention to critical characteristics of concepts and Type X, to explain and elaborate these characteristics, and both Phase 1 and Phase 2 can create a safe space for students to explore, generate, make mistakes, and learn and practice with methods to approach and solve Type X.

Knowledge as static vs. as dynamic

The key notion in CLT that learning is the development of cognitive schemas in long-term memory is somehow based on the assumption that content to be learned is something static that can be captured in schemas which can then be retrieved from long-term memory. However, high-stakes settings like the ones in the previous paragraph have in common that the nature of knowledge, tasks, and problems is dynamic and ever-evolving. With the advancement of science and technology, many things learned once upon a time turn out to be less useful than expected or lose their usefulness because the nature of problems, roles and responsibilities has changed.

Apart from these high-stakes settings, let us take learning and maintaining a foreign language as an example. From personal experience, most of us can tell that grammar structures and proverbs in a foreign language once learned become rusty and may be retrieved with error (i.e., incorrect memories) if we do not (continue to) use that foreign language regularly. Using that foreign language regularly, with native or otherwise fluent speakers, provides a natural form of RP. Besides, language evolves; new words and proverbs are born, and the use of grammar structures may change with time as well, and that RP of using the language with others can help us to adapt to these changes. In this respect, knowledge is not necessarily exclusively about something "out there" for us to learn but is at least to some extent also cocreated in dialogue and conversation. Finally, we do not need to see a worked example or completion problem for any new grammar structure or proverb; in line with PF, much of it is learned while "struggling" in a conversation with others.

Definitions and poor methodological practice

As mentioned earlier, the concepts of element interactivity, cognitive load, and cognitive overload – key concepts in CLT – are poorly defined and good measures are lacking. In fact, the dominant measurement practice since 1992 has been to have participants self-report on a nine-point scale how much mental effort they invested in a task that just completed [21-22], depending on the study either once at the end of a learning and/or post-test stage or several times (i.e., repeatedly) during a learning and/or post-test stage, for instance after each of a series of tasks. This practice has persisted despite repeated critiques, including perfect confounding of measurement error, differences in tasks in which it is used, and a likely shift in participants' response from one task to the next [23]. A robust rule from psychometrics is that that single self-report items can be incredibly noisy (i.e., large measurement error) and are usually much noisier than measurements obtained from series of items on the same variable of interest. Task differences may make it difficult to compare ratings from different tasks not in the last place because our willingness to invest mental effort in a given task may well depend on how many tasks we have seen before and how much effort we invested in each of these. Finally, response shift is a real issue because our conceptions of

task complexity as well as our self-assessments of what we are capable of may change as we learn. Newcomers in a complex topic are often poor self-assessors in that topic (e.g., [24]); this is a skill to be improved with practice. If when seeing a counterintuitive probability problem for the first time we think it is easy and therefore invest little mental effort, then learn about the solution and steps to be taken towards the solution and realise it is more difficult than anticipated, we may invest more mental effort in a second problem of the same type not because the second problem is more complex but because we now have a better appreciation of some initially "hidden" complexities or difficulties and we have become more aware of the limitations of our probability problem solving skills.

To account for a range of empirical findings that could not be explained only in terms of a general "cognitive load" or mental effort invested, cognitive load theorists introduced different types of cognitive load, some of which linking to for learning not effective load (i.e., "bad" load) some of which potentially stimulating learning (i.e., "good" load). It is beyond the scope of this article to provide a detailed review of these different types of load and how different scholars have attempted to define and measure these types of load, but this work has been done already anyway (e.g., [1-5, 20, 23]) and can be briefly summarised as follows. On the one hand, there are cognitive load theorists who state that we need three types of cognitive load: load arising from essential aspects of the task (*intrinsic*), load due to non-essential aspects of the task (*extraneous*), and load arising from the deliberate engagement in learning (*germane*) (e.g., [7, 25]). On the other hand, there are scholars who state that *germane* load is that part of the *intrinsic* load that results in learning (i.e., not all intrinsic load results in learning); from this perspective, germane load is therefore not a third independent type of load but part of intrinsic load (e.g., [2, 5, 20, 26]). Along with this lack of consensus in definitions, we have seen the development and use of a variety of self-report questionnaires (e.g., [12, 27-30]) which all attempt to measure two or three types of load but with somewhat different wording. Each of these questionnaires suffers from question wording effects, suffers from the same task differences and response shift issues as the mental effort self-report item, and all beg the same question: if we cannot even properly define element interactivity or cognitive load let alone agree on the number and definitions of types of load, what on earth are we measuring?

Despite the disagreement on definitions and measurement, researchers continue to use their own definitions and measurement tools without mention of alternative views, as if no one ever questioned for instance the role of germane load or any of the issues with the use of self-report measurements. They usually do so in small-sample experiments with questionable manipulations of element interactivity or cognitive load that leave the reader with a variety of possible alternative explanations for the findings reported. A recent example of this comes from Lehmann and Seufert [25], who with 42 learners in a 2x2 between-subjects design (i.e., 42 learners divided into four groups) claim to have found evidence in favour of tailoring instruction to learners' preferred learning styles, despite very clear evidence against that idea (e.g., [31]). However, there are several possible alternative explanations for this finding in favour of learning styles, including the following.

To start, it is well known that in samples as small as the one at hand findings can vary wildly from one experiment to another, and while individual experiments might indicate a clear effect in one direction in an accidental sample across experiments there might be no difference at all or even a clear difference in the other direction (e.g., [31]). Statistically significant outcomes in one experiment may not be replicated in a future study, and statistically non-significant outcomes cannot be interpreted as evidence in favour of no difference (e.g., "there are no main effects" or "there is no interaction effect"). Even if the statistical power of a statistical test for a particular effect of interest in each of two independent experiments carried out under exactly the same conditions is high, provided the anticipated effect (of a given size) exists, the chance of establishing

a statistically significant outcome in both experiments is the product of the statistical power of the two experiments (e.g., [23]). For example, if in two experiments we achieve a power of 0.80, the chance of obtaining a statistically significant outcome in both experiments is 0.80 x 0.80 = 0.64, or 64%. In CLT research, sample sizes are often such that for an effect of for instance half a standard deviation the appropriate statistical test has a statistical power of around 0.50. With such a low power, the chance of a statistically significant finding for the effect of interest in two independent experiments, if the anticipated effect (of the size specified) exists, is only 25% (!). Yet, interpretations of statistically non-significant outcomes such as "there is no effect" are all over the place in CLT research.

Lehmann and Seufert asked learners to indicate their preference for either auditive or visual texts, and they found that among learners with a preference for visual text the ones given visual texts on average learned more than their peers who were given auditive texts. However, as they themselves recognise, most texts in everyday life are presented visually, so an increased closeness to real life may be a much more likely explanation for this finding than tailoring materials to learners' preferred learning styles. Furthermore, there is another potentially obvious confounder: reading skills. What if the participants who indicated a preference towards visual texts happen to have been the ones with better reading skills compared to the ones who indicated a preference towards auditive texts? When having to process information, competence and preference often go together, and if that is the case here, the ones with better reading skills may more frequently have indicated a visual preference than the ones with somewhat poorer reading skills. The finding that the presentation of visual texts on average resulted in better outcomes in the "visual preference" group than in the "auditive-ambiguous preference" group may then largely if not exclusively reflect a difference in reading skills rather than a difference in (whether or not tailoring to) style per se.

To conclude

Two key statements from CLT are that (1) learning is optimal when instructional support is decreased going from worked examples via completion problems to autonomous problem solving and (2) learners do not benefit from practicing retrieval with complex content. However, research inspired by PF has provided evidence against (1) while research on RP has provided evidence against (2). An immediate recommendation for teachers and others involved in educational practice is to not consider CLT – or any educational theory for that matter – as the holy grail providing the whole "truth" and nothing but the truth about what works and what does not work in education, but to consider the robust findings on PF and RP as well. One of the key contributors (if not the most important contributor) to PF may be having learners working in dyads or small groups to learn from each other. This is a possibility that should be investigated further in both laboratory and actual educational settings. Learning from peers may involve learning new things but may equally function as a form of RP. Future studies could experiment with this possibility to determine under what conditions RP may contribute to or diminish any potential PF effects.

The suggestion that CLT is a useless theory that can now be placed in the museum of dead theories is neither the message nor the intention of this article. However, to assess the continued relevance of CLT as a key contributor to educational research and practice, more cognitive load theorists should take note of critical arguments that have been made for quite a while now. Specifically, findings from research on PF and RP that contradict core predictions from CLT, critiques on the lack of definition and good measures and the lack of consensus on these questions in the cognitive load community, and recommendations for good methodological and statistical practice such as striving for larger samples and refraining from interpreting statistically non-significant findings as evidence of "no difference". If we take these points together, we may in the next years learn much more about conditions under which CLT, PF, and

RP converge, under which conditions they diverge, and what are the best possible recommendations for educational practice and further research based on this convergence and divergence.

Notes

Funding

This study did not receive financial support from external sources

Conflicts of interest disclosure

The authors declare no competing interests relevant to the content of this study.

Authors' contributions.

All the authors declare to have made substantial contributions to the conception, or design, or acquisition, or analysis, or interpretation of data; and drafting the work or revising it critically for important intellectual content; and to approve the version to be published.

Availability of data and responsibility for the results

All the authors declare to have had full access to the available data and they assume full responsibility for the integrity of these results.

REFERENCES

1. Sweller J. Element interactivity and intrinsic, extraneous, and germane cognitive load. Educ Psychol Rev [Internet]. 2010 [cited 2020 Mar 01];22:123-38. https://doi.org/10.1007/s10648-010-9128-5.

2. Sweller J, Ayres P, Kalyuga S. Cognitive load theory [monograph on Internet]. New York: Springer; 2011. [cited 2020 Mar 01]. https://doi.org/10.1007/978-1-4419-8126-4.

3. Sweller J, Van Merrienboer JJG, Paas F. Cognitive architecture and instructional design. Educ Psychol Rev [Internet]. 1998 [cited 2020 Mar 02];10:251-96. https://doi.org/10.1023/A:1022193728205.

4. Sweller J, Van Merrienboer JJG, Paas F. Cognitive architecture and instructional design: 20 years later. Educ Psychol Rev [Internet]. 2019 [cited 2020 Mar 02];31:261-92. https://doi.org/10.1007/s10648-019-09465-5.

5. Leppink J, Van den Heuvel A. The evolution of cognitive load theory and its application to medical education. Perspect Med Educ [Internet]. 2015 [cited 2020 Mar 02];4:119-27. https://doi.org/10.1007/s40037-015-0192-x.

6. Van Merrienboer JJG, Sweller J. Cognitive load theory in health professions education: design principles and strategies. Med Educ [Internet]. 2010 [cited 2020 Mar 02];44:85-93. https://doi.org/10.1111/j.1365-2923.2009.03498.x.

7. Young JQ, Van Merrienboer JJG, Durning S, Ten Cate O. Cognitive load theory: implications for medical education: AMEE Guide No. 86. Med Teach [Internet]. 2014 [cited 2020 Mar 02];36:371-84. https://doi.org/10.3109/0142159X.2014.889290.

8. Karpicke JD, Aue WR. The testing effect is alive and well with complex materials. Educ Psychol Rev [Internet]. 2015 [cited 2020 Mar 02];27:317-26. https://doi.org/10.1007/s10648-015-9309-3.

9. Roediger HL, Karpicke JD. The power of testing memory: Basic research and implications for educational practice. Perspect Psychol Sci [Internet]. 2006 [cited 2020 Mar 02];1:181-210. https://doi.org/10.1111/j.1745-6916.2006.00012.x.

10. Roediger HL, Karpicke JD. Test-enhanced learning: Taking memory tests improves long--term retention. Psychol Sci [Internet]. 2006 [cited 2020 Mar 02];17:249-55. https://doi.org/10.1111/j.1467-9280.2006.01693.x.

11. Van Gog T, Sweller J. Not new, but nearly forgotten: The testing effect decreases or even disappears as the complexity of learning materials increases. Educ Psychol Rev [Internet]. 2015 [cited 2020 Mar 02];27:247-64. https://doi.org/10.1007/s10648-015-9310-x.

12. Leppink J, Paas F, Van Gog T, Van der Vleuten CPM, Van Merrienboer JJG. Effects of pairs of problems and examples on task performance and different types of cognitive load. Learn Instruct [Internet]. 2014 [cited 2020 Mar 02];30:32-42. https://doi.org/10.1016/j.learninstruc.2013.12.001.

13. Van Gog T, Kester L, Paas F. Effects of worked examples, example-problem, and problem-example pairs on novices' learning. Contemp Educ Psychol [Internet]. 2011 [cited 2020 Mar 03];36:212-18. https://doi.org/10.1016/j.cedpsych.2010.10.004.

14. Kalyuga S, Singh AM. Rethinking boundaries of cognitive load theory in complex learning. Educ Psychol Rev [Internet]. 2016 [cited 2020 Mar 02];28:831-52. https://doi.org/10.1007/s10648-015-9352-0.

15. Kapur M. Productive failure. Cognit Instruct [Internet]. 2008 [cited 2020 Mar 03];26:379-424. https://doi.org/10.1080/07370000802212669.

16. Kapur M. A further study of productive failure in mathematical problem solving: Unpacking the design components. Instruct Sci [Internet]. 2011 [cited 2020 Mar 03];39:561-79. https://doi.org/10.1007/s11251-010-9144-3.

17. Kapur M. Productive failure in learning math. Cognit Sci [Internet]. 2014 [cited 2020 Mar 03];38:1008-22. https://doi.org/10.1111/cogs.12107.

18. Kapur M, Rummel N. Productive failure in learning from generation and invention activities. Instruct Sci [Internet]. 2012 [cited 2020 Mar 03];40:645-50. https://doi.org/10.1007/s11251-012-9235-4.

19. Kirschner PA, Sweller J, Kirschner F, Zambrano JR. From cognitive load theory to collaborative cognitive load theory. Inter J Comp Supp Collab Learn [Internet]. 2018 [cited 2020 Mar 03];13:213-33. https://doi.org/10.1007/s11412-018-9277-y.

20. Lee CB, Hanham J, Leppink J. Instructional design principles for high-stakes problem-solving environments [monograph on Internet]. Singapore: Springer; 2019. [cited 2020 Mar 03]. https://doi.org/10.1007/978-981-13-2808-4.

21. Paas F. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. Educ Psychol. [Internet]. 1992 [cited 2020 Mar 03];84:429-34. https://doi.org/10.1037/0022-0663.84.4.429.

22. Sweller J. Measuring cognitive load. Perspect Med Educ [Internet]. 2018 [cited 2020 Mar 03];7:1-2. https://doi.org/10.1007/s40037-017-0395-4.

23. Leppink J. Statistical methods for experimental research in education and psychology [monograph on Internet]. Cham: Springer; 2019. [cited 2020 Mar 03]. https://doi.org/10.1007/978-3-030-21241-4.

24. Bjork RA, Dunlosky J, Kornell N. Self-regulated learning: Beliefs, techniques, and illusions. Ann Rev Psychol [Internet]. 2013 [cited 2020 Mar 03];64:417-44. https://doi.org/10.1146/annurev-psych-113011-143823.

25. Lehmann J, Seufert T. The interaction between text modality and the learner's modality preference influences comprehension and cognitive load. Front Psychol [Internet]. 2020; [cited 2020 Mar 05]. https://doi.org/10.3389/fpsyg.2019.02820.

26. Kalyuga S. Cognitive load theory: How many types of load does it really need? Educ Psychol Rev [Internet]. 2011 [cited 2020 Mar 05];23:1-19. https://doi.org/10.1007/s10648-010-9150-7.

27. Klepsch M, Schmitz F, Seufert T. Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. Front Psychol [Internet]. 2017; [cited 2020 Mar 05]. https://doi.org/10.3389/fpsyg.2017.01997.

28. Leppink J, Paas F, Van der Vleuten CPM, Van Gog T, Van Merrienboer JJG. Development of an instrument for measuring different types of cognitive load. Behav Res Meth[Internet]. 2013[cited 2020 Mar 05];45:1058-72. https://doi.org/10.3758/s13428-013-0334-1.

29. Naismith LM, Cheung JJH, Ringsted C, Cavalcanti RB. Limitations of subjective cognitive load measures in simulation-based procedural training. Med Educ [Internet]. 2015 [cited 2020 Mar 05];49:805-14. https://doi.org/10.1111/medu.12732.

30. Sewell JL, Boscardin CK, Young JQ, Ten Cate O, O'Sullivan PS. Measuring cognitive load during procedural skills training with colonoscopy as an exemplar. Med Educ [Internet]. 2016 [cited 2020 Mar 05];50:682-92. https://doi.org/10.1111/medu.12965.

31. Pashler H, McDaniel M, Rohrer D, Bjork R. Learning styles: Concepts and evidence. Psychol Sci Pub Int [Internet]. 2008 [cited 2020 Mar 05];9:105-119. https://doi.org/10.1111/j.1539-6053.2009.01038.x.

Mailing address:

Jimmie Leppink

University of York

Heslington Y010 5DD

York, North Yorkshire (NY), United Kingdom