What We Do and Do Not Know about Teaching Medical Image Interpretation

Ellen M. Kok1*, Koos van Geel1, Jeroen J. G. van Merriënboer1 and Simon G. F. Robben2

1 Department of Educational Development and Research, School of Health Professions Education, Maastricht University, Maastricht, Netherlands, 2 Department of Radiology, Maastricht University Medical Centre, Maastricht, Netherlands

Educators in medical image interpretation have difficulty finding scientific evidence as to how they should design their instruction. We review and comment on 81 papers that investigated instructional design in medical image interpretation. We distinguish between studies that evaluated complete offline courses and curricula, studies that evaluated e-learning modules, and studies that evaluated specific educational interventions. Twenty-three percent of all studies evaluated the implementation of complete courses or curricula, and 44% of the studies evaluated the implementation of e-learning modules. We argue that these studies have encouraging results but provide little information for educators: too many differences exist between conditions to unambiguously attribute the learning effects to specific instructional techniques. Moreover, concepts are not uniformly defined and methodological weaknesses further limit the usefulness of evidence provided by these studies. Thirty-two percent of the studies evaluated a specific interventional technique. We discuss three theoretical frameworks that informed these studies: diagnostic reasoning, cognitive schemas and study strategies. Research on diagnostic reasoning suggests teaching students to start with non-analytic reasoning and subsequently applying analytic reasoning, but little is known on how to train non-analytic reasoning. Research on cognitive schemas investigated activities that help the development of appropriate cognitive schemas. Finally, research on study strategies supports the effectiveness of practice testing, but more study strategies could be applicable to learning medical image interpretation. Our commentary highlights the value of evaluating specific instructional techniques, but further evidence is required to optimally inform educators in medical image interpretation.

Keywords: medical image interpretation, education, instructional design, e-learning, curriculum design, diagnostic reasoning, cognitive schemas, study strategies

INTRODUCTION

How to teach medical image interpretation? For an educator in radiology, dermatology, pathology or cardiology, this might be the question in mind. Since ‘evidence-based medicine’ is held in high regard by clinicians, medical educators might aim to search the literature for evidence on how to design their instruction. Instructional design is the science and practical field of creating educational experiences (Merrill et al., 1996). This can be as broad as a curriculum or as narrow as a lesson, or even a single instructive animation. Unfortunately, evidence regarding how to teach
medical image interpretation is hard to come by. Research on instructional design in medical image interpretation suffers from being scattered all over the literature, from a lack of cross-references and a lack of theoretical background. A lot of commentaries are published (e.g., Fenderson, 2005; Gunderman and Ballenger, 2014), which might serve as an inspiration, but should not be considered ‘evidence.’ This makes it challenging for medical educators to find and apply the relevant literature to their educational practice. The aim of this paper is twofold. On the one hand we synthesize existing literature about instructional design in medical image interpretation. On the other hand we identify gaps in the literature and propose research inspired by psychological theories as a solution. While we used an extensive literature search to inform our argument, the aim of this paper is not to be systematic and exhaustive, but to provide a commentary that is informed by a literature search.

METHODS

We searched Web of Science for papers related to instructional design, using the keywords (teach* OR education OR instruct* OR curric*) AND [(Radiology OR Radiography) OR (pathology AND image) OR dermatology OR (electrocardiogra* OR ECG)]. We focused on papers less than 15 years old (2001- July 2016) to include recent papers only. This search yielded 4785 papers. Titles were scanned by EMK for relevance and subsequently the abstracts were scanned by KvG. 120 papers were selected for complete reading. EMK and KvG each checked half of these papers against the inclusion criteria. If doubt arose regarding the inclusion, both readers read the paper and discrepancy was resolved through discussion. We included only papers that implemented an educational experience and measured the effects of this intervention (against a control condition and/or against a pretest). We do not go into a discussion about what ‘effects’ of education are and should be, but consider ‘effective’ to be ‘yielding a higher score on a test,’ or ‘being evaluated more positively.’ We excluded papers where medical image interpretation was not the outcome measure (e.g., procedural knowledge), or that treated medical images as tools for teaching something else (e.g., use of radiographs as an illustration in anatomy classes). Table 1 in the Supplementary Materials provides an overview of the 81 selected papers, and they are marked with an asterisk in the reference list. We identified two broad categories of studies: (1) evaluations of curricula and courses (we discuss offline courses and e-learning courses separately) and (2) evaluations of specific instructional techniques. Three theoretical frameworks that form the basis of specific instructional techniques arose from the review: diagnostic reasoning, cognitive schemas, and study strategies. The curricula and courses in the first category commonly implement the specific instructional techniques in the second category of studies. However, studies in the first category rarely discuss specific instructional techniques, and, critically, these instructional techniques are not separately tested. We argue below that only evaluations of specific instructional techniques provide information that educators can use to design their education. In this paper, we discuss respectively evaluations of offline curricula and courses, evaluations of e-learning courses and evaluations of specific instructional techniques. For the purpose of the argument, we discuss representative papers in the rest of the manuscript and refer the reader to the Supplementary Materials for a complete overview of the reviewed papers.

REVIEW

Evaluation of Offline Curricula and Courses

Twenty-three percent of the reviewed studies evaluated a course or curriculum. These are often a combination of lectures, workshops and self-study. The outcome of these studies might seem straight-forward and reassuring: the score on the post-test is typically higher than the score on the pretest, and the ‘new’ curriculum is more effective and more positively evaluated than the ‘old’ curriculum. However, appraisal of these results is problematic, for a variety of reasons. Firstly, numerous differences exist between the ‘new’ curriculum and the ‘old’ curriculum (or other control conditions). This makes the outcome ambiguous: it is impossible to know what makes a new curriculum more effective than the old curriculum. Possibly, the instructional techniques used in the new curriculum are more effective (but, if so, which of the techniques?). But the difference might just as well be caused by other factors, such as enthusiasm of the staff or students for the new curriculum. In line with Norman (2003), we argue that the evaluation of complete courses yields trivial findings that provide no insights for educators, unless the course is carefully compared with another course that only differs on specific, well-defined aspects. If seemingly more specific techniques are compared, such as case-based learning or self-directed learning, the findings provide hardly more insights. The techniques still differ in too many aspects, and often they are very broadly and not uniformly defined. This makes it even more difficult to compare the results over studies. On top of the problem that learning effects cannot be unambiguously attributed to specific instructional techniques, in many of the studies the methodology had apparent weaknesses, e.g., no control conditions and/or the use of inappropriate randomization.

Evaluation of E-learning Modules

E-learning and blended learning are also widely investigated in medical image interpretation (44% of all papers). E-learning refers to learning activities that interactively use a computer to enhance learning (Ruiz et al., 2006). Blended learning refers to a mix of online with traditional (lecture-based) learning activities (Spanjers et al., 2015). E-learning and blended learning environments often provide participants with the opportunity to work through patient cases or provide content information in an interactive manner.

E-learning is a popular way to promote active learning: it allows for large groups of learners to engage in learning at a time and place convenient for them, has the potential to be tailored to learners’ needs, and allows for instructional designs that cannot be implemented in other formats (Cook et al., 2008).
Certainly, e-learning is widely found to be more effective or non-inferior to traditional forms of teaching, in both medical education in general (Cook et al., 2008), and medical image interpretation (Zafar et al., 2014). Indeed, the lion’s share of the studies that we reviewed conform to this. Once again, the differences between the e-learning curriculum and the control condition (often a traditional, lecture-based curriculum) are often too large to unequivocally attribute effects to instructional techniques. This issue is even more pressing when it comes to e-learning. Not only do conditions differ from each other in terms of instructional techniques, the differences in the technology used for implementation further confound the comparison.

When an online or offline curriculum is implemented, it aims to maximize learning instead of isolating the contribution of specific educational techniques, even though a curriculum applies a set of instructional techniques. This means that studies resulting from this type of implementations are not optimized for providing specific information on how instruction should be designed. Indeed, many of the methodological weaknesses in this type of studies result from practical and ethical considerations (e.g., it is often impossible to randomly assign students to conditions). However, for an educator in medical image interpretation, it is important to know what specific techniques yield more effective learning, and therefore researchers need to design studies that answer this question. In the next section, we review studies that zoom in on specific instructional techniques. We argue that these specific studies are more informative for educators, not only because they provide more detailed information about ‘what works,’ but also because they are often theory-driven.

**Evaluation of Specific Instructional Techniques**

Thirty-two percent of the studies that we reviewed evaluated a specific instructional technique. Most of these studies are (implicitly or explicitly) rooted in psychological theories. We discuss three psychological theories that together form the basis of most of these studies. We discuss respectively theories of diagnostic reasoning (Eva, 2004), cognitive schema theory (Charlin et al., 2007) and study strategies (Dunlosky et al., 2013).

**Diagnostic Reasoning**

Research in cognitive psychology proposes two modes of reasoning: analytic and non-analytic reasoning. Analytic reasoning refers to deliberate, effortful reasoning while non-analytic reasoning refers to automatic, rapid reasoning, also referred to as pattern recognition (Eva, 2004). Several studies use this framework to investigate specific educational techniques in medical image interpretation. For example, Ark et al. (2007) stimulated students to “carefully identify all features [of an ECG] while trusting guidance provided by feelings of familiarity,” i.e., balancing an analytical approach (carefully identifying features) and a non-analytical approach (trusting feelings of familiarity). This was more effective then not providing students with instructions on how to approach this task. Likewise, Baghdady et al. (2014a) found that students who were directed to diagnose a radiograph first and only then identify radiographic features outperformed participants who identified features first and then diagnosed the radiograph.

The claim that students should be instructed to diagnose a case first, based on feelings of familiarity (non-analytic reasoning) and only then collect and analyze all information (analytic reasoning) contrasts with the claim that it is crucial to systematically collect all relevant information in a medical image before making a diagnosis, which is the assumption underlying the idea of teaching a search pattern (Auffermann et al., 2015, 2016). While these studies provide evidence for a benefit of a search pattern training in radiology over no training, the benefit of systematic viewing training over a non-systematic search pattern training could not be established in radiology (Kok et al., 2016) or in ECG interpretation (Varvaroussis et al., 2014).

To sum up, these studies suggest teaching a balanced reasoning strategy, starting with non-analytic reasoning and subsequently applying analytic reasoning. Patel et al. (2015) suggest mapping and microanalysis as two tools to understand a learners’ reasoning process and provide focused feedback to train students in balancing reasoning strategies. Another option is to present students with high numbers of cases under time pressure, as a way to counteract the relatively high emphasis on analytical reasoning in medical education (Patel et al., 2015). These instructional techniques have not been investigated in medical image interpretation yet.

**Cognitive Schemas**

Diagnostic reasoning requires extensive knowledge that is structured into meaningful patterns, so-called cognitive schemas or illness scripts (Charlin et al., 2007). These contain information about pathophysiological processes underlying diseases, patients’ characteristics and signs and symptoms (Boshuizen and Schmidt, 1992; Van De Wiel et al., 2000). The acquisition of high-quality scripts is central to learning in medicine (Charlin et al., 2007) and thus several studies have explicitly or implicitly focused on helping students to develop high-quality scripts, often with a focus on either pathophysiological processes underlying diseases, patients characteristics or signs and symptoms.

Boshuizen and Schmidt (Boshuizen and Schmidt, 1992; Van De Wiel et al., 2000) argue that basic science knowledge (i.e., the understanding of pathophysiological processes underlying diseases) is fundamental to these illness scripts. However, with increased expertise, basic science knowledge is encapsulated and only used for difficult, atypical cases. Baghdady et al. (2009) argue that basic science knowledge helps diagnosis through creating a coherent mental representation of diseases and their (visual) features. Participants that were provided with causal explanations of radiological features learned more than students who were presented with feature lists or a structured algorithm. A second study, however, found that the negative effect of the structured algorithm (without basic science explanations) was mitigated by the previously discussed instruction to provide a diagnosis before summing up all features (Baghdady et al., 2014a).

The prevalence of a disease is included in illness scripts. Building solid information about actual prevalence of diseases into illness scripts can avoid the ‘prevalence bias’ in decision making (Croskerry, 2003). Pusic et al. (2012) found that the
been implemented in medical image interpretation. However, examples of active, student-centered learning have successfully surpassed teacher-centered, lecture-based education and many of those strategies are not applied to visual tasks, so the application of the study strategies in diverse situations, many of which, as Rumsfeld said, ‘we do not know that we do not know them.’ Finally, we focused on outcome measures that reflect medical image interpretation. This excludes research on other important topics such as indications for imaging and professionalism.

Research on expertise in medical image interpretation states that experts have superior perceptual and cognitive abilities (Krupinski, 2010; Manning, 2010; Nodine and Mello-Thoms, 2010; Reingold and Sheridan, 2011; van der Gijp et al., 2014). Interestingly, few studies on education explicitly relate to research about visual expertise. Another remarkable gap in the literature is the lack of studies that focus on individualized and self-regulated learning, two important topics in instructional design nowadays (Van Merrienboer and Kirschner, 2013). In other visual domains, such as air traffic control (Salden et al., 2006), it was found that adapting training to the learners’ needs makes learning more efficient, so this finding is promising for medical image interpretation. Research on self-regulated learning in non-visual diagnostic reasoning provides a possible starting point for fostering self-regulated learning (Cleary et al., 2016). However, visual metacognition is found to be rather low (e.g., Võ et al., 2016), so it is important to understand how visual metacognition can be fostered in order to optimize learning medical image interpretation. In conclusion, we discussed findings relevant for

**Limitations**

A limitation of this commentary is that we did not formally assess the quality of the reviewed studies. In general, randomized controlled trials are scarce and few studies included appropriate control conditions, although this problem was less prevalent in studies of specific interventions. A systematic assessment of study quality was beyond the scope of this literature-informed commentary but could be a relevant venue for further research. Furthermore, while this commentary discusses what conclusions cannot be drawn from the reviewed literature, there are many topics of which, as Rumsfeld said, ‘we do not know that we do not know them.’ Finally, we focused on outcome measures that reflect medical image interpretation. This excludes research on other important topics such as indications for imaging and professionalism.

Research on expertise in medical image interpretation states that experts have superior perceptual and cognitive abilities (Krupinski, 2010; Manning, 2010; Nodine and Mello-Thoms, 2010; Reingold and Sheridan, 2011; van der Gijp et al., 2014). Interestingly, few studies on education explicitly relate to research about visual expertise. Another remarkable gap in the literature is the lack of studies that focus on individualized and self-regulated learning, two important topics in instructional design nowadays (Van Merrienboer and Kirschner, 2013). In other visual domains, such as air traffic control (Salden et al., 2006), it was found that adapting training to the learners’ needs makes learning more efficient, so this finding is promising for medical image interpretation. Research on self-regulated learning in non-visual diagnostic reasoning provides a possible starting point for fostering self-regulated learning (Cleary et al., 2016). However, visual metacognition is found to be rather low (e.g., Võ et al., 2016), so it is important to understand how visual metacognition can be fostered in order to optimize learning medical image interpretation. In conclusion, we discussed findings relevant for
teaching medical image interpretation, but many open questions remain, and further evidence is required to optimally inform educators in medical image interpretation.

AUTHOR CONTRIBUTIONS

The authors conceptualized the work together. EK and KvG conducted the review, EK wrote a first draft, all authors revised the work critically for important intellectual content. All authors approved the final version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00309/full#supplementary-material
Fenderson, B. A. (2005). Strategies for teaching pathology to graduate students and allied health professionals. *Hum. Pathol.* 36, 146–153. doi: 10.1016/j.humpath.2004.09.022

Garg, A., Haley, H. L., and Hatem, D. (2010). Modern moulage evaluating the use of 3-dimensional prosthetic mimics in a dermatology teaching program for second-year medical students. *Arch. Dermatol.* 146, 143–146. doi: 10.1001/archdermatol.2009.355

Giunta, A., Di Stefani, A., and Chimenti, S. (2011). Mobile phones: a role in teaching dermatology? *Dermatology* 222, 22–23. doi: 10.1159/000317074

Gunderman, R. B., and Ballenger, Z. (2014). The golden rule of education. *Frontiers in Psychology* 6.

Krupinski, E. A. (2010). "Perceptual factors in reading medical images," in *Handbook of Medical Image Perception and Techniques*, eds E. Samei and E. Krupinski (Cambridge: Cambridge University Press), 91–106.

Lee, B. F., Chiu, N. T., and Li, C. Y. (2013). Value of case-based learning in a dental educational experiments. *BMC Med. Educ.* 13, 5–7.

Mileman, P. A., van den Hout, W. B., and Sandrink, G. C. H. (2003). Randomized controlled trial of a computer-assisted learning program to improve caries detection from bitewing radiographs. *Dentomaxillofac. Radiol.* 32, 116–123. doi: 10.1259/dmftr/58225203

Montassier, E., Hardouin, J. B., Segard, J., Batard, E., Potel, G., Planchon, B., et al. (2016). e-Learning versus lecture-based courses in ECG interpretation for undergraduate medical students: a randomized noninferiority study. *Eur. J. Emerg. Med.* 23, 108–113. doi: 10.1097/mej.0000000000000215

Nilsson, M., Bolinder, G., Held, C., Johansson, B. L., Fors, U., and Ostergren, J. (2008). Evaluation of a web-based ECG-interpretation programme for undergraduate medical students. *BMJ Med. Educ.* 8:25. doi: 10.1186/1742-6920-8-25

O'Connor, E. E., Fried, J., McNulty, N., Shah, P., Hogg, J. P., Lewis, P., et al. (2012). Effect of dermoscopy education on the ability of medical students to detect skin cancer. *Arch. Dermatol.* 148, 1016–1022. doi: 10.1001/archdermatol.2012.509

O'Connell, E. F., Fried, J., McNulty, N., Shah, P., Hogg, J. P., Lewis, P., et al. (2016). Flipping radiology education right side up. *Acad. Radiol.* 23, 810–822. doi: 10.1016/j.acra.2016.02.011

Pagnanelli, G., Soyer, H. P., Argenziano, G., Talamini, R., Barbati, R., Bianchi, L., et al. (2003). Diagnosis of pigmented skin lesions by dermoscopy: web-based...
training improves diagnostic performance of non-experts. Br. J. Dermatol. 148, 698–702. doi: 10.1016/j.bjder.2006.05.016.X

Patel, R., Sanders, J., and Carr, S. (2015). Clinical diagnostic decision-making in real life contexts: a trans-theoretical approach for teaching. AMEE Guide No. 95. Med. Teach. 37, 211–227. doi: 10.1080/0142159X.2014.975195

*Porras, L., Drezner, J., Dotson, A., Stafford, H., Berkoff, D., Agnihotri, K., et al. (2016). Novel interpretation of screening electrocardiograms and impact of online training. J. Electrocardiol. 49, 462–466. doi: 10.1016/j.jelecard.2016.02.004

*Porter, K. K., Bailey, P. D., Woods, R., Scott, W. W., and Johnson, P. T. (2015). Retrieved surgical item identification on imaging studies: a training module for radiology residents. Int. J. Comput. Assist. Radiol. Surg. 10, 1803–1809. doi: 10.1007/s11548-013-1154-9

*Pusic, M. V., Andrews, J. S., Kessler, D. O., Teng, D. C., Pecaric, M. R., Rual-Shapiro, C., et al. (2012). Prevalence of abnormal cases in an image bank affects the learning of radiograph interpretation. Med. Educ. 46, 289–298. doi: 10.1111/j.1365-2923.2011.04165.x

*Pusic, M. V., LeBlanc, V. R., and Miller, S. Z. (2007). Linear versus web-style layout in medical education. *A simple pilot study for rapidly increasing pulmonary embolism recognition among nursing students. J. Exp. Psychol. Hum. Percept. Perform. 42, 1677–1681. doi: 10.1037/xhp0000264

Võ, M. L. H., Aizenman, A. M., and Wolfe, J. M. (2016). You think you know where you looked? You better look again. J. Exper. Psychol. Hum. Percept. Perform. 42, 1477–1481. doi: 10.1037/sxp0000264

Webb, A. L., and Choi, S. (2014). Interactive radiological anatomy eLearning solution for first year medical students: development, integration, and impact on learning. *Anatomical Sciences Education 7*, 350–360. doi: 10.1002/ase.1428

Williams, J., Sato, T. S., and Policeni, B. (2013). Pulmonary embolism teaching file: a simple pilot study for rapidly increasing pulmonary embolism recognition among new residents using interactive cross-sectional imaging. *Acad. Radiol. 20*, 1048–1051. doi: 10.1016/j.acra.2012.12.020

Zafar, S., Safdar, S., and Zafar, A. N. (2014). Evaluation of use of e-Learning in undergraduate radiology education: a review. *Eur. J. Radiol. 83*, 2277–2287. doi: 10.1016/j.ejrad.2014.08.017
Zeng, R., Yue, R. Z., Tan, C. Y., Wang, Q., Kuang, P., Tian, P. W., et al. (2015). New ideas for teaching electrocardiogram interpretation and improving classroom teaching content. Adv. Med. Educ. Pract. 6, 99–104. doi: 10.2147/amep.s75316

Zhang, H. J., and Hsu, L. L. (2013). The effectiveness of an education program on nurses’ knowledge of electrocardiogram interpretation. Int. Emerg. Nurs. 21, 247–251. doi: 10.1016/j.ienj.2012.11.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer JS and the handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2017 Kok, van Geel, van Merriënboer and Robben. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.