




Assessment tools for reflection in healthcare learners: A scoping review

Cassandra Préfontaine, Isabelle Gaboury, Hélène Corriveau, Jacinthe Beauchamp, Chantal Lemire & Marie-Josée April


To cite this article: Cassandra Préfontaine, Isabelle Gaboury, Hélène Corriveau, Jacinthe Beauchamp, Chantal Lemire & Marie-Josée April (2021): Assessment tools for reflection in healthcare learners: A scoping review, Medical Teacher, DOI: [10.1080/0142159X.2021.1998400](https://doi.org/10.1080/0142159X.2021.1998400)


To link to this article: <https://doi.org/10.1080/0142159X.2021.1998400>

 [View supplementary material](#) 

 Published online: 13 Nov 2021.

 [Submit your article to this journal](#) 

 Article views: 371

 [View related articles](#) 

 [View Crossmark data](#) 



Assessment tools for reflection in healthcare learners: A scoping review

Cassandra Préfontaine^a, Isabelle Gaboury^b, Hélène Corriveau^c, Jacinthe Beauchamp^d, Chantal Lemire^e and Marie-Josée April^f

^aDepartment of Anatomy, Université du Québec à Trois-Rivières, Trois-Rivières, Canada; ^bDepartment of Family Medicine and Emergency Medicine, Université de Sherbrooke, Longueuil, Canada; ^cSchool of rehabilitation, Université de Sherbrooke, Sherbrooke, Canada; ^dCentre de formation médicale du Nouveau-Brunswick, Université de Moncton, Moncton, Canada; ^eDepartment of Pediatrics, Université de Sherbrooke, Sherbrooke, Canada; ^fVice-décanat au Développement professionnel et à la responsabilité sociale, Université de Sherbrooke, Sherbrooke, Canada

ABSTRACT

Reflection has been integrated in many healthcare education programs to achieve deeper learning and improve professional practice. A variety of evaluation tools are used to assess reflection, but few guides are available to inform educators in their choice of a relevant evaluation tool. The aim of this paper is to identify all existing evaluation tools published along with their strengths and weaknesses. A review strategy retrieved tools available in Medline, PsychInfo, CINAHL and Eric databases. The procedures outlined by Munn and colleagues were used to synthesize the information. Additionally, the reflection dimensions assessed in each tool (when sufficient information was available) were analyzed deductively, using thematic analysis according to the Killian and Todnem framework. Subthemes were identified inductively. Forty-five papers were identified, reporting on 34 different tools. The tools were based on a variety of theoretical models. Some had evidence of adequate validity and fidelity. Eleven components of reflection were identified across tools. No tool encompassed all components, but most tools included between three and five components. Context surrounding evaluation should be carefully considered when choosing an evaluation tool for reflection. There is a need for further research to validate the psychometric properties of reflection evaluation tools.

KEYWORDS

Reflection; evaluation; healthcare learner; assessment; higher education

Main text introduction

Reflection in medical education is defined as 'a metacognitive process that occurs before, during and after situations with the purpose of developing greater understanding of both the self and the situation so that future encounters with the situation are informed from previous encounters' (Sandars 2009). The concept of reflection is now an integral part of best practices in medical education (Sandars 2009; Hargreaves 2016; Pretorius and Ford 2016; Uygur et al. 2019) and role definitions of a healthcare professional (Sandars 2009). Reflection is part of the attributes of competent health care professionals (Mann et al. 2009) allowing for deeper learning (Sobral 2000; Chinniah and Nalliah 2012; Wain 2017), better understanding of the concepts being learned (Hargreaves 2016; Carter et al. 2017; Wain 2017; Farahhana et al. 2018), and improved professional practice (Glaze 2001; Sandars 2009; Pretorius and Ford 2016; Wain 2017). These include social skills (Chen and Forbes 2014), therapeutic relationships (Sandars 2009) and motivation for continuous learning (Hargreaves 2016). Assessment of reflection is also required for professional licensing (Mann et al. 2009; Pretorius and Ford 2016; Uygur et al. 2019), for example in Canada, the Netherlands, and Australia through the CanMEDS (Frank et al. 2015), as well as requalification (Mann et al. 2009; Wain 2017). As such, activities promoting reflection are now included in

Practice points

- Several tools are available for educators wishing to assess reflection in healthcare students.
- Considering the variety of tools and significant limitations in their psychometric properties, context surrounding evaluation should guide the selection of the most appropriate assessment tool.
- Educators should also consider the definition and theoretical model of reflection taught in their program to select the most appropriate tool.

undergraduate, graduate, and continuing medical and health science education (Mann et al. 2009).

Some scholars claim that reflection assessment tools could limit reflection by being too rigidly structured (Driessen et al. 2005; Ross 2014) or by fear of the judgement of others (McMullan et al. 2003; Macfarlane and Gourlay 2009; Ross 2014). Others have reported on the tools' lack of validity and faithfulness (Stewart and Richardson 2000; Bourner 2003); that they are very demanding in terms of resources (Pee et al. 2002; McMullan 2006) and that they generate anxiety among learners because of the uncertainty surrounding the criteria

(Stewart and Richardson 2000; McMullan 2006). Many argue, however, that the assessment of reflection emphasizes the importance of the process and increases both learners' and educators' engagement in the task (Bourner 2003; Driessen et al. 2005); ensures an adequate depth of reflection among learners (Mitchell 1994; Driessen et al. 2005) and provides feedback to scaffold the skills of learner and to enable learning (Kennison and Misselwitz 2002; Pee et al. 2002). Moreover others, despite the tool methodological limitations, suggest that learners' reflection skills could be improved through teaching, that mentoring and support are two important factors for such learning to occur (Mann et al. 2009), and thus that assessing of reflection remains relevant. While ontological and epistemological stances of competency assessment tools have been the subject of inquiring (Tavares et al. 2019), it is beyond the scope of this review.

In the context of initial professional training (training program leading to a first diploma allowing the practice of the profession), educators have used a variety of activities such as written reflection, interviews, simulation, or video watching to initiate and spark the reflection process among learners. A variety of evaluation tools, such as grids and checklists have been developed to assist those educators in the assessment of the process. The large number of tools available and the large variability between them makes it difficult to choose the most appropriate tool for a specific context. Therefore, it appears necessary to conduct a review of the literature to identify what tools are available and to summarize their strengths and weaknesses for medical and health science educators. The goal of this scoping review, as defined by Munn et al. (2018), was to help these initial training educators to make an informed choice in their evaluation methods for the process of reflection.

Material and methods

Four databases that include healthcare related content published between 1966 and May 2021 were searched: MedLine, PsychINFO, CINHALL and ERIC. The search strategy was based on the following terms: (reflexive* or introspect* or autoregulation or reflect*) AND (student* OR trainee* OR learner* OR resident*) AND (evaluat* or assess* or measur* or summati* OR apprais*) AND (diar* or journal* or vide* or report* or task* or writ* or tool* or activit* OR scale* OR score* OR questionnaire*). The search strategy was adapted to each database. A snowballing technique, searching the reference list of each included paper, was used to identify additional tools. The searches were performed in May 2021.

The titles and abstracts were screened (CP and IG). The primary author screened all abstracts, and the second author independently reviewed a random 15% of the generated list. Inter-rater agreement was 99.9% (kappa = 0.76), which was considered satisfying to stop dual review (Stoll et al. 2019). Full texts of the selected papers were read by two reviewers (CP and IG). Discrepancies between the two reviewers were resolved by consensus for all steps. Articles were included if authors: (1) evaluated the process of reflection; (2) presented a tool being used on healthcare learners; and (3) published in French or English. Data were

extracted by the first author (CP) and validated by the second author (IG). There were no exclusion criteria.

Analysis: A data extraction chart was compiled, categorizing the tools according to their target population, the evaluation tasks assigned to the trainees, the evaluation criteria, and the psychometric properties of the assessment tool. The data extraction form was developed by consensus among all co-authors. This group includes medical educators, interprofessional education experts, medical education experts, individuals involved in medical and health science trainee evaluation, as well as curricula developers. The first author then conducted a thematic analysis (Clarke et al. 2015) of the assessment criteria for each tool, to identify the dimensions of reflection assessed by each tool. The themes were determined deductively using the Killion and Todnem (1991) classification for reflection in action, on action, and for action. Subthemes related to the framework and beyond were identified inductively. Codification was reviewed and discussed with the second author (IG).

The Killion and Todnem framework describes engaging students in taking actions to learn and to enhance their practice, using reflection as a continuous process. In this framework, reflection *in* action is the reflection of the learners while the action was occurring, reflection *on* action is the reflection that learners make retrospectively following a situation, and reflection *for* action refers to reflection learners do to map out their behavior in the future how they will act similarly or differently in the future. The framework includes recognition by learners of the new knowledge developed from the experiences and planning of objectives or strategies for change.

Results

Thirty-eight papers were identified by applying the inclusion criteria to literature searches. Seven papers were added by searching the names of specific tools identified through the initial search, for a total of 45 papers (Figure 1). From these 45 papers, 34 different tools were identified.

This review shows a high diversity in the activities used to evaluate reflection (Supplementary Table 1). Numerous studies focused on situations experienced by the students, as in portfolio learning activities (Rees and Sheard 2004; Dalal et al. 2012; Pitts and Ruggirello 2012; Carter et al. 2017; Devi et al. 2017; Costa et al. 2018), journaling (multiple entries over time about a longitudinal experience) (Kember et al. 1999; Williams et al. 2000; Kennison and Misselwitz 2002; Plack et al. 2005; 2007; Findlay et al. 2010; Padden 2013), written reflection (a single entry about a one-time experience) (Wong et al. 1995; Jensen and Joy 2005; Plack et al. 2005, 2007; Learman et al. 2008; Wallman et al. 2008; Dempsey et al. 2009; Findlay et al. 2010; Wittich et al. 2010; Findlay et al. 2011; Root and Waterfield 2015; Liao and Wang 2016; McEvoy et al. 2016; Miller-Kuhlmann et al. 2016; Devi et al. 2017; Lucas et al. 2017; Tutticci et al. 2017; Tully et al. 2018) verbal reflection (Lewis et al. 2011), objective structured clinical examination (Bogo et al. 2011), observation of the student in practice (Carter et al. 2016) and a reflective worksheet (i.e. multiple specific questions about a situation) (Pee et al. 2002). In contrast, other studies turned to fictitious situations or situations that had not been experienced by the learner himself to organize the

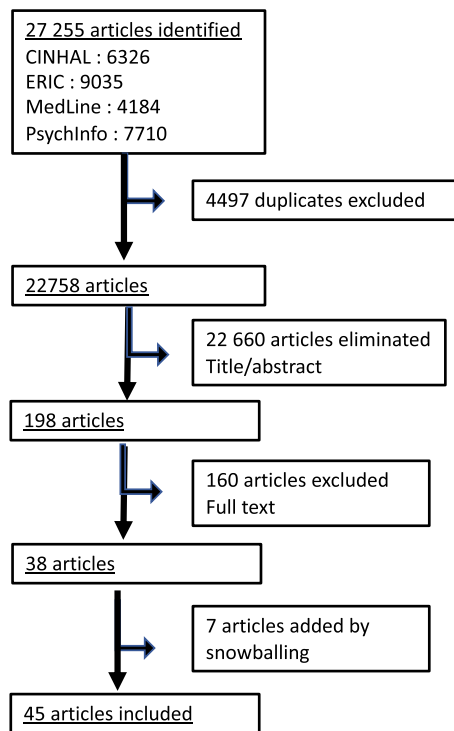


Figure 1. Paper selection flow chart.

reflexive activities. These include vignettes (Boenink et al. 2004), video-cases (Koole et al. 2012; Tsingos-Lucas et al. 2016) and reflection about a situation that happened to somebody else (Tully et al. 2018). Reflection activities were carried out throughout the learners' curriculum, in various disciplines including medicine, nursing, midwifery, social work, health policy, radiation therapy, psychology, health science, pharmacy, dental therapy, physical therapy. The amount of time necessary to train the evaluators, complete the activity, and assess the learners vary from few minutes to more than an hour.

Seven theoretical frameworks were used to develop the tools included in this inventory. Four theoretical frameworks specifically focus on reflection (Greenwood 1993; Boud et al. 1996; Moon 2013; Bass et al. 2017); while others consider reflection as part of a learning process (Kitchenham 2008; Armstrong 2016; Costa et al. 2018), either linear or cyclical. Furthermore, frameworks differ in their ability to identify different stages of the reflection process (Greenwood 1993; Boud et al. 1996; Bass et al. 2017) or to conceptualize the quality of depth of reflection (Kitchenham 2008; Moon 2013; Bass et al. 2017) All framework but one (Bass et al. 2017) shared a transdisciplinary approach. These characteristics influence the nature of the tools and may be considered when the time comes to select the appropriate tool.

Most tools were designed to be used by an external evaluator, less than a quarter of the studies used learner self-assessment tools, in which learners assess their own reflexive abilities (Kember et al. 2000; Sobral 2000, 2001; Grant et al. 2002; Leung and Kember 2003; Sobral 2005; Aukes et al. 2007; Akeroyd 2012; Lethbridge et al. 2013).

Few studies assessed psychometric properties such as validity and fidelity. Using the Messick Unitary Framework for validity, evidence of good validity based on the content was present for nine of the tools (Kember et al. 2000; Aukes et al. 2007; Dempsey et al. 2009; Wittich et al. 2010;

Lewis et al. 2011; Padden 2013; Root and Waterfield 2015; Carter et al. 2016, 2017; Devi et al. 2017); evidence of good validity based on process response was present for two of the tools (Pee et al. 2002; Root and Waterfield 2015); evidence of good validity based on internal structure was present for sixteen of the tools (Kember et al. 2000; Sobral 2000, 2001; Williams et al. 2000; Grant et al. 2002; Boenink et al. 2004; Sobral 2005; Aukes et al. 2007; Learman et al. 2008; Wittich et al. 2010; Bogoy et al. 2011; Koole et al. 2012; Lethbridge et al. 2013; Andersen et al. 2014; Carter et al. 2016; Miller-Kuhlmann et al. 2016; Carter et al. 2017; Tutticci et al. 2017; So et al. 2018); and evidence of good validity based on relations to other variables was present for nineteen of the tools (Wong et al. 1995; Kember et al. 2000; Sobral 2000, 2001; Grant et al. 2002; Pee et al. 2002; Leung and Kember 2003; Boenink et al. 2004; Rees et al. 2005; Sobral 2005; Learman et al. 2008; Bogoy et al. 2011; Dalal et al. 2012; Koole et al. 2012; Andersen et al. 2014; Miller-Kuhlmann et al. 2016; Tutticci et al. 2017; Costa et al. 2018; So et al. 2018; Tully et al. 2018). Validity based on consequence was never assessed. Although nine of the tools reported evidence of validity based on the content, for many of them the level of evidence was superficial. Eight tools did not have evidence of validity for their tool. Concerning fidelity, good inter-rater reliability was reported for twenty-five tools (Wong et al. 1995; Kember et al. 1999; Williams et al. 2000; Kennison and Misselwitz 2002; Pee et al. 2002; Boenink et al. 2004; Rees and Sheard 2004; Jensen and Joy 2005; Plack et al. 2005, 2007; Learman et al. 2008; Findlay et al. 2010; Wittich et al. 2010; Findlay et al. 2011; Dalal et al. 2012; Koole et al. 2012; Pitts and Ruggirello 2012; Padden 2013; Miller-Kuhlmann et al. 2016; Carter et al. 2017; Devi et al. 2017; Lucas et al. 2017; Costa et al. 2018; Tully et al. 2018) and good test-retest reliability for four tools (Sobral 2000; Grant et al. 2002; Rees et al. 2005; Sobral 2005; Andersen et al. 2014). Six tools did not evaluate fidelity. None of the tools had been evaluated with more than one discipline of learners. Only one tool assessed the acceptability from the user point of view (Tsingos-Lucas et al. 2016).

Themes covered by the various tools are summarized in Figure 2. Nine subthemes emerged across the 34 tools that were identified. No subtheme emerged from Reflection *in* Action (A on Figure 2) as details on the elements evaluated were insufficient. Reflection *on* action was subdivided into six subthemes consisting of: (B) multiple perspectives on a situation, i.e. considering how other individuals might have experienced the situation; (C) the impact that the context had on the situation or on the trainee's thoughts and actions, i.e. considering the impact that the environment and the moment of the situation that was used for the reflection assessment had on subsequent events; (D) the impact of the trainee's affect and cognition on the situation, i.e. how their thought and feeling influenced the unfolding events; (E) the impact that previous experiences or life experiences in general had on their thoughts and actions, i.e. how their past experiences influenced the way that they perceived the situation and reacted to it; (F) their own limits as learners, i.e. recognition of their weaknesses and their lack of experience; and (G) alternatives, i.e. to consider how they could have reacted or thought differently, and what would have been the impacts.

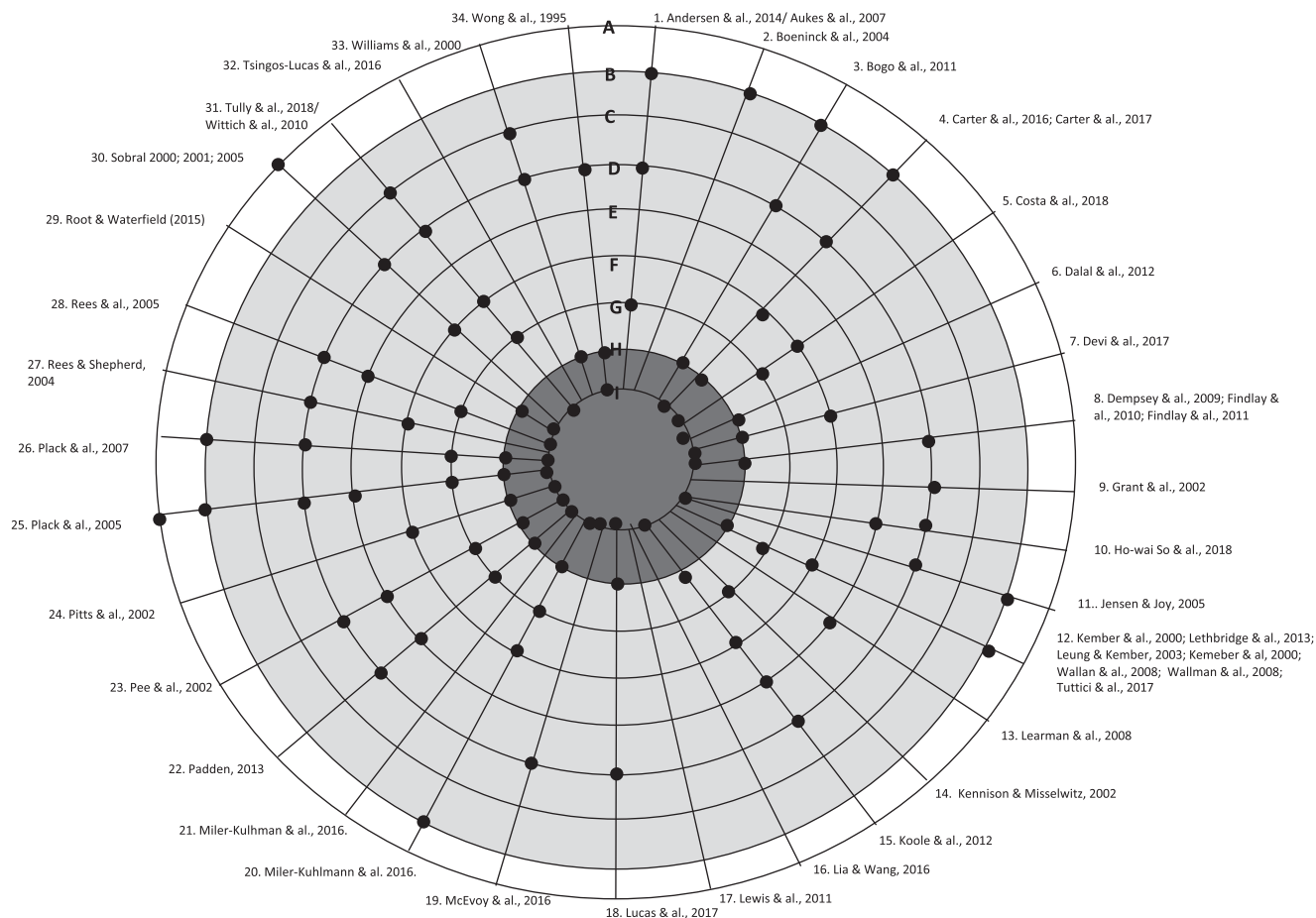


Figure 2. Components of reflection assessed in each tool.

Numbers around the circles identify each tool. Each circle represents a component of reflection as per Killion and Todnem's framework: white = reflection *in* action (A); light grey = reflection *on* action (including Consider multiple perspectives (B), Consider the impact of the context (C), Consider the impact of one's thoughts and feeling (D), Consider the impact of past situations (E), Consider one's own limits (F), and Consider alternatives (G)); dark grey = reflection *for* action (including Identify new learnings (H), and Plan for change (I)). Themes and sub-themes identified from the inductive/deductive analysis of each tool are represented by dots at the intersections.

Reflection *for* action was subdivided into two subthemes: (H) recognizing the new knowledge developed from the trainee's experiences; and (I) strategies for change. For four tools, the information on assessment criteria was insufficient to classify them. No tool encompassed all components, but most tools included between three to five subthemes.

Discussion

This study is one of the first to offer an inventory of reflection assessment tools used by medical and health science initial training educators. Thirty-four tools to evaluate reflection were identified, along with their metrological properties, as well as the contexts use. A comparison of the components of reflection that were evaluated also shows that these tools assessed different aspects of reflection.

Psychometric properties

The choice of any student assessment tool rests in part on the psychometric properties, which varied considerably among identified tools. Any tool must be valid, to reflect accurately the learning or skills that we want learners to achieve. In addition, given the correction burden associated

with reflection assessment, it is important to ensure good inter-judge fidelity since it is likely that more than one assessor will be involved with large groups of learners. There were significant limitations to the psychometric properties of the tools identified. Indeed, the assessment of the validity of tools was generally limited to apparent validity and did not include thorough evaluation of validity. No tools have used evidence concerning consequences to support the validity of a tool, despite the fact that assessment in higher education can have significant impact on the length and success of a student throughout the academic curriculum. In addition, despite the relevance of measuring progress in a learning context, only one study assessed the tools' sensitivity to change, and information regarding the feasibility and acceptability of the tools was limited.

Considering the context

In order to make the best possible choice of assessment tools, it is important that the relative importance given to psychometric properties be specifically contextualized to the situations in which the tool will be applied. Indeed, the time required of students, the influence of students' motivation to produce the best possible work, and costs related to use of the tool in terms of human resources for the training of evaluators and scoring – all must be considered

(Banta and Palomba 2014). Even with robust psychometric properties, it is important to consider the resources, culture, politics and structures in place in the department to ensure optimal implementation of the tool (Pawson and Tilley 1997). Although a set of theoretical frameworks has been used by these authors, some more frequently than others, it is important that the framework be consistent with the teaching of reflection that is done. In the absence of a common, consensus definition of the concept of reflection, the validity of any tool in real-life application is uncertain. Thus, classifying the dimensions of reflection captured by each of the tools studied can help to ensure that the tool responds well to the way in which the concept of reflection is operationalized in the educational setting. In this sense, it is not possible to recommend a universal assessment tool. Rather, this review is intended to provide a picture of the available tools and their characteristics so that initial training educators can make an informed choice by considering their assessment needs and context.

Limitations of this paper relate to the quality of reporting of the tools identified. The lack of detail in some papers as to the criteria of their tools did not permit classification of the components of reflection being assessed. In addition, no tools have been studied in more than one discipline. It is therefore not possible to discuss the applicability or utility of these tools to students from all health disciplines nor in the context of interprofessional education. Finally, our search strategy did not include any sources of grey literature, unless it was referenced by a tool selected for this review.

Conclusion

This article analyzed 34 tools available to healthcare educators, identified through a complex and rigorous review process. Considering the variety of tools included in this inventory and the generalized limitations in the tools' psychometric properties, the authors recommend that users consider primarily the components of reflection they wish to assess (reflection in action, on action or for action) as well as the context in which the tool will be used, including the resources available to the user. There is a need for research on the metrological properties of reflection assessment tools, and their generalizability to a variety of contexts and disciplines.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Notes on contributors

Cassandra Préfontaine, MSc, Department of Anatomy, Université du Québec à Trois-Rivières, Trois-Rivières, Canada.

Isabelle Gaboury, PhD, Department of Family Medicine and Emergency Medicine, Université de Sherbrooke, Longueuil, Canada.

Hélène Corriveau, PhD, School of rehabilitation, Université de Sherbrooke, Sherbrooke, Canada.

Jacinte Beauchamp, PhD, Centre de formation médicale du Nouveau-Brunswick, Université de Moncton, Moncton, Canada and Department of family medicine and emergency medicine, Université de Sherbrooke, Sherbrooke, Canada.

Chantal Lemire, MD, Department of Pediatrics, Université de Sherbrooke, Sherbrooke, Canada.

Marie-Josée April, PhD, Vice-décanat au développement professionnel, à la pédagogie et à la responsabilité sociale.

References

- Akeroyd J. 2012. Evaluating students' reflective thinking: critique of measurement tools. *Commun Nurs Res*. 45:478–478.
- Andersen NB, O'Neill L, Gormsen LK, Hvidberg L, Morcke AM. 2014. A validation study of the psychometric properties of the Groningen Reflection Ability Scale. *BMC Med Educ*. 14(1):214.
- Armstrong P. 2016. Bloom's taxonomy. Nashville: Vanderbilt University Center for Teaching.
- Aukes LC, Geertsma J, Cohen-Schotanus J, Zwierstra RP, Slaets JPJ. 2007. The development of a scale to measure personal reflection in medical practice and education. *Med Teach*. 29(2–3):177–182.
- Banta TW, Palomba CA. 2014. *Assessment essentials: planning, implementing, and improving assessment in higher education*. Hoboken: John Wiley & Sons.
- Bass J, Fenwick J, Sidebotham M. 2017. Development of a model of holistic reflection to facilitate transformative learning in student midwives. *Women Birth*. 30(3):227–235.
- Boenink AD, Oderwald AK, de Jonge P, van Tilburg W, Smal JA. 2004. Approaches to learning. Assessing student reflection in medical practice. The development of an observer-rated instrument: reliability, validity and initial experiences. *Med Educ*. 38(4):368–377.
- Bogo M, Regehr C, Katz E, Logie C, Mylopoulos M. 2011. Developing a tool for assessing students' reflections on their practice. *Soc Work Educ*. 30(2):186–194.
- Boud D, Keogh R, Walker D. 1996. Promoting reflection in learning: a model. *Boundaries of Adult Learning*. 1:32–56.
- Bourner T. 2003. Assessing reflective learning. *Educ Train*. 45(5): 267–272.
- Carter AG, Creedy DK, Sidebotham M. 2016. Development and psychometric testing of the Carter Assessment of Critical Thinking in Midwifery (Preceptor/Mentor version) [Validation Studies]. *Midwifery*. 34:141–149.
- Carter AG, Creedy DK, Sidebotham M. 2017. Critical thinking evaluation in reflective writing: development and testing of Carter Assessment of Critical Thinking in Midwifery (Reflection). *Midwifery*. 54:73–80.
- Chen I, Forbes C. 2014. Reflective writing and its impact on empathy in medical education: systematic review. *J Educ Eval Health Prof*. 11:20.
- Chinniah K, Nalliah S. 2012. Reflective writing in case summary assignments. *Int e-J Sci Med Educ*. 6(1):15–20.
- Clarke V, Braun V, Hayfield N. 2015. Thematic analysis. In: Smith JA, editor. *Qualitative psychology: a practical guide to research methods*. London: SAGE Publications; p. 222–248.
- Costa GDd, Driessen E, Silva Lsd, Campos AAdO, Costa TMTd, Donateli CP, Cotta RMM. 2018. Collective portfolio: assessment of teaching and learning in health undergraduate courses. *Cien Saude Colet*. 23(11):3779–3787.
- Dalal DK, Hakel MD, Sliter MT, Kirkendall SR. 2012. Analysis of a Rubric for assessing depth of classroom reflections. *Int J ePortfolio*. 2(1): 75–85.
- Dempsey S, Warren-Forward H, Findlay N. 2009. Development of the Newcastle reflective analysis tool. *Focus Health Prof Educ*. 11(1):32.
- Devi V, Abraham RR, Kamath U. 2017. Teaching and assessing reflecting skills among undergraduate medical students experiencing research. *J Clin Diagn Res*. 11(1):JC01–JC05.
- Driessen EW, van Tartwijk J, Overeem K, Vermunt JD, van der Vleuten CPM. 2005. Conditions for successful reflective use of portfolios in undergraduate medical education. *Med Educ*. 39(12):1230–1235.

- Farahhana RA, Mohamed Nadzri YH, Zainuddin ZI. 2018. Reflective approaches in medical education: implications on medical imaging. *Int J Allied Health Sci Pract.* 16(2):1–12.
- Findlay N, Dempsey S, Warren-Forward H. 2010. Validation and use of the Newcastle Reflective Analysis Tool: a three-year longitudinal study of RT students' reflective journals. *Reflective Pract.* 11(1): 83–94.
- Findlay N, Dempsey SE, Warren-Forward HM. 2011. Development and validation of reflective inventories: assisting radiation therapists with reflective practice. *J Radiother Pract.* 10(1):3–12.
- Frank JR, Snell L, Sherbino J, Boucher A. 2015. *CanMEDS 2015. Physician Competency Framework Series I.*
- Glaze JE. 2001. Reflection as a transforming process: student advanced nurse practitioners' experiences of developing reflective skills as part of an MSc programme. *J Adv Nurs.* 34(5):639–647.
- Grant AM, Franklin J, Langford P. 2002. The self-reflection and insight scale: a new measure of private self-consciousness. *Soc Behav Pers.* 30(8):821–835.
- Greenwood J. 1993. Reflective practice: a critique of the work of Argyris and Schön. *J Adv Nurs.* 18(8):1183–1187.
- Hargreaves K. 2016. Reflection in medical education. *J Univ Teach Learn.* 13(2):79–21.
- Jensen SK, Joy C. 2005. Exploring a model to evaluate levels of reflection in baccalaureate nursing students' journals [Evaluation Studies]. *J Nurs Educ.* 44(3):139–142.
- Kember D, Jones A, Loke A, McKay J, Sinclair K, Tse H, Webb C, Wong F, Wong M, Yeung E. 1999. Determining the level of reflective thinking from students' written journals using a coding scheme based on the work of Mezirow. *Int J Lifelong Educ.* 18(1):18–30.
- Kember D, Leung DYP, Jones A, Loke AY, McKay J, Sinclair K, Tse H, Webb C, Yuet Wong FK, Wong M, et al. 2000. Development of a questionnaire to measure the level of reflective thinking. *Assess Eval High Educ.* 25(4):381–395.
- Kennison MM, Misselwitz S. 2002. Evaluating reflective writing for appropriateness, fairness, and consistency. *Nurs Educ Perspect.* 23(5):238–242.
- Killion JP, Todnem GR. 1991. A process for personal theory building. *Educ Leadership.* 48(6):14–16.
- Kitchenham A. 2008. The evolution of John Mezirow's transformative learning theory. *J Transform Educ.* 6(2):104–123.
- Koole S, Dornan T, Aper L, De Wever B, Scherpbier A, Valcke M, Cohen-Schotanus J, Derese A. 2012. Using video-cases to assess student reflection: development and validation of an instrument. *BMC Med Educ.* 12:22.
- Learman LA, Autry AM, O'Sullivan P. 2008. Reliability and validity of reflection exercises for obstetrics and gynecology residents. *Am J Obstet Gynecol.* 198(4):461. e461–461. e410.
- Lethbridge K, Andrusyszyn M-A, Iwaswi C, Laschinger HK, Fernando R. 2013. Assessing the psychometric properties of Kember and Leung's Reflection Questionnaire. *Assess Eval High Educ.* 38(3): 303–325.
- Leung DYP, Kember D. 2003. The relationship between approaches to learning and reflection upon practice. *Educ Psychol.* 23(1):61–71.
- Lewis D, Virden T, Hutchings PS, Bhargava R. 2011. Competence assessment integrating reflective practice in a professional psychology program. *J Scholars Teach Learn.* 11(3):86–106.
- Liao H-C, Wang Y-h. 2016. The application of heterogeneous cluster grouping to reflective writing for medical humanities literature study to enhance students' empathy, critical thinking, and reflective writing. *BMC Med Educ.* 16(1):1–11.
- Lucas C, Bosnic-Anticevich S, Schneider CR, Bartimote-Aufflick K, McEntee M, Smith L. 2017. Inter-rater reliability of a reflective rubric to assess pharmacy students' reflective thinking. *Curr Pharm Teach Learn.* 9(6):989–995.
- Macfarlane B, Gourlay L. 2009. The reflection game: enacting the penitent self. *Teach High Educ.* 14(4):455–459.
- Mann K, Gordon J, Macleod A, Mann K, Gordon J, Macleod A. 2009. Reflection and reflective practice in health professions education: a systematic review. *Adv Health Sci Educ Theory Pract.* 14(4):595–621.
- McEvoy M, Pollack S, Dyché L, Burton W. 2016. Near-peer role modeling: can fourth-year medical students, recognized for their humanism, enhance reflection among second-year students in a physical diagnosis course? *Med Educ Online.* 21(1):31940.
- McMullan M, Endacott R, Gray MA, Jasper M, Miller CML, Scholes J, Webb C. 2003. Portfolios and assessment of competence: a review of the literature. *J Adv Nurs.* 41(3):283–294.
- McMullan M. 2006. Students' perceptions on the use of portfolios in pre-registration nursing education: a questionnaire survey. *Int J Nurs Stud.* 43(3):333–343.
- Miller-Kuhlmann R, O'Sullivan PS, Aronson L. 2016. Essential steps in developing best practices to assess reflective skill: a comparison of two rubrics. *Med Teach.* 38(1):75–81.
- Mitchell M. 1994. The views of students and teachers on the use of portfolios as a learning and assessment tool in midwifery education. *Nurse Educ Today.* 14(1):38–43.
- Moon JA. 2013. *A handbook of reflective and experiential learning: theory and practice.* Milton Park: Routledge.
- Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol.* 18(1):1–7.
- Padden ML. 2013. A pilot study to determine the validity and reliability of the Level of Reflection-on-Action Assessment [Validation Studies]. *J Nurs Educ.* 52(7):410–415.
- Pawson R, Tilley N. 1997. *Realistic evaluation.* Thousand Oaks (CA): Sage Publications, Inc. (Realistic evaluation).
- Pee B, Woodman T, Fry H, Davenport ES. 2002. Appraising and assessing reflection in students' writing on a structured worksheet [Research Support, Non-U.S. Gov't]. *Med Educ.* 36(6):575–585.
- Pitts W, Ruggirello R. 2012. Using the e-Portfolio to document and evaluate growth in reflective practice: the development and application of a conceptual framework. *Int J ePortfolio.* 2(1):49–74.
- Plack MM, Driscoll M, Blissett S, McKenna R, Plack TP. 2005. A method for assessing reflective journal writing. *J Allied Health.* 34(4):199–208.
- Plack MM, Driscoll M, Marquez M, Cuppernull L, Maring J, Greenberg L. 2007. Assessing reflective writing on a pediatric clerkship by using a modified Bloom's Taxonomy [Validation Studies]. *Ambul Pediatr.* 7(4):285–291.
- Pretorius L, Ford A. 2016. Reflection for learning: teaching reflective practice at the beginning of University Study. *Int J Teach Learn High Educ.* 28(2):241–253.
- Rees CE, Sheard CE. 2004. The reliability of assessment criteria for undergraduate medical students' communication skills portfolios: The Nottingham experience. *Med Educ.* 38(2):138–144.
- Rees CE, Shepherd M, Chamberlain S. 2005. The utility of reflective portfolios as a method of assessing first year medical students' personal and professional development. *Reflect Pract.* 6(1):3–14.
- Root H, Waterfield J. 2015. An evaluation of a reflective writing assessment within the MPharm programme. *Pharm Educ.* 15(1):142–144.
- Ross J. 2014. Performing the Reflective self: audience awareness in high-stakes reflection. *Stud High Educ.* 39(2):219–232.
- Sandars J. 2009. The use of reflection in medical education: AMEE Guide No. 44. *Med Teach.* 31(8):685–695.
- So SH-w, Bennett-Levy J, Perry H, Wood DH, Wong C-W. 2018. The Self-Reflective Writing Scale (SRWS): a new measure to assess self-reflection following self-experiential cognitive behaviour therapy training. *Reflect Pract.* 19(4):505–521.
- Sobral D. 2000. An appraisal of medical students' reflection-in-learning. *Med Educ.* 34(3):182–187.
- Sobral DT. 2001. Medical students' reflection in learning in relation to approaches to study and academic achievement. *Med Teach.* 23(5): 508–513.
- Sobral DT. 2005. Medical students' mindset for reflective learning: a revalidation study of the reflection-in-learning scale. *Adv Health Sci Educ Theory Pract.* 10(4):303–314.
- Stewart S, Richardson B. 2000. Reflection and its place in the curriculum on an undergraduate course: should it be assessed? *Assess Eval High Educ.* 25(4):369–380.
- Stoll CR, Izadi S, Fowler S, Green P, Suls J, Colditz GA. 2019. The value of a second reviewer for study selection in systematic reviews. *Res Synth Methods.* 10(4):539–545.
- Tavares W, Kuper A, Kulasegaram K, Whitehead C. 2019. The compatibility principle: on philosophies in the assessment of clinical competence. *Adv Health Sci Educ Theory Pract.* 2019:1–16.
- Tsingos-Lucas C, Bosnic-Anticevich S, Smith L. 2016. A retrospective study on students' and teachers' perceptions of the reflective ability clinical assessment. *Am J Pharm Educ.* 80(6):101.

- Tully V, Murphy D, Fioratou E, Chaudhuri A, Shaw J, Davey P. 2018. Learning from errors: assessing final year medical students' reflection on safety improvement, five year cohort study. *BMC Med Educ.* 18(1):57.
- Tutticci N, Coyer F, Lewis PA, Ryan M. 2017. Validation of a reflective thinking instrument for third-year undergraduate nursing students participating in high-fidelity simulation. *Reflect Pract.* 18(2):219–231.
- Uygur J, Stuart E, De Paor M, Wallace E, Duffy S, O'Shea M, Smith S, Pawlikowska T. 2019. A Best Evidence in Medical Education systematic review to determine the most effective teaching methods that develop reflection in medical students: BEME Guide No. 51. *Med Teach.* 41(1):3–16.
- Wain A. 2017. Learning through reflection. *Br J Midwifery.* 25(10):662–666.
- Wallman A, Lindblad AK, Hall S, Lundmark A, Ring L. 2008. A categorization scheme for assessing pharmacy students' levels of reflection during internships. *Am J Pharm Educ.* 72(1):05.
- Williams RM, Sundelin G, Foster-Seargeant E, Norman GR. 2000. Assessing the reliability of grading reflective journal writing. *J Phys Ther Educ.* 14(2):23–26.
- Wittich CM, Beckman TJ, Drefahl MM, Mandrekar JN, Reed DA, Krajicek BJ, Haddad RM, McDonald FS, Kolars JC, Thomas KG. 2010. Validation of a method to measure resident doctors' reflections on quality improvement [Validation Studies]. *Med Educ.* 44(3):248–255.
- Wong FK, Kember D, Chung LY, Yan L. 1995. Assessing the level of student reflection from reflective journals [Research Support, Non-U.S.]. *J Adv Nurs.* 22(1):48–57.