

## **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**

### AUTHOR(S)

Jane Uygun, Ellen Stuart, Muireann De Paor, Emma Wallace, Seamus Duffy, Maire O'Shea, Susan Smith, Teresa Pawlikowska

### CITATION

Uygun, Jane; Stuart, Ellen; De Paor, Muireann; Wallace, Emma; Duffy, Seamus; O'Shea, Maire; et al. (2020): 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. Royal College of Surgeons in Ireland. Journal contribution. <https://hdl.handle.net/10779/rcsi.12886976.v1>

### HANDLE

[10779/rcsi.12886976.v1](https://hdl.handle.net/10779/rcsi.12886976.v1)

### LICENCE

**CC BY-NC-SA 4.0**

This work is made available under the above open licence by RCSI and has been printed from <https://repository.rcsi.com>. For more information please contact [repository@rcsi.com](mailto:repository@rcsi.com)

### URL

[https://repository.rcsi.com/articles/journal\\_contribution/A\\_Best\\_Evidence\\_in\\_Medical\\_Education\\_systematic\\_review\\_to\\_determine\\_the\\_most\\_effective\\_teaching\\_methods\\_that\\_develop\\_reflection\\_in\\_medical\\_students\\_BEME\\_guide\\_No\\_51/12886976/1](https://repository.rcsi.com/articles/journal_contribution/A_Best_Evidence_in_Medical_Education_systematic_review_to_determine_the_most_effective_teaching_methods_that_develop_reflection_in_medical_students_BEME_guide_No_51/12886976/1)

# **A Best Evidence in Medical Education Systematic Review to determine the most effective teaching methods that develop reflection in medical students**

## **Abstract**

**Introduction:** Reflection is thought to be an essential skill for physicians. Although much has been written about it, there is little concurrence about how to best teach reflection in medical education. The aim of this review was to determine: i) which educational interventions are being used to develop reflection, ii) how is reflection being assessed, and iii) what are the most effective interventions.

**Methods:** Inclusion criteria comprised: i) undergraduate medical students, ii) a teaching intervention to develop reflection, and iii) assessment of the intervention. A review protocol was developed and nine databases were searched. Screening, data extraction and analysis procedures were performed in duplicate. Due to the heterogeneity of studies a narrative synthesis approach was performed for the study analysis.

**Results:** Twenty-eight studies met the inclusion criteria. The interventions in these studies had at least two of the following components related to reflection: i) introduction, ii) trigger, iii) writing, iv) guidelines, v) small group discussion, vi) tutor and vii) feedback. Three validated rubrics were used to assess reflective writing in these studies.

**Conclusion:** The strongest evidence from studies in this review indicates that guidelines for, and feedback on, reflective writing improve student reflection.

## **Practice Points**

- There is a great deal of heterogeneity in the type of interventions and study designs for teaching reflection to medical students
- Interventions had at least two of the following components: i) introduction, ii) trigger, iii) writing, iv) guidelines, v) small group discussion, vi) tutor and vii) feedback
- The strongest evidence indicated guidelines for, and feedback on, reflective writing improved student reflection
- Three rubrics used to assess reflective writing in these studies were validated: i) Reflect Ability Rubric (RAR), ii) Groningen Reflection Ability Scale (GRAS) and iii) Reflection Evaluation for Enhanced Competencies Tool (REFLECT)
- A future systematic review concentrating on the validity and reliability of available tools to assess reflection is needed

**Keywords:** reflection, undergraduate medical students, medical education, learning, assessment

## Notes on Contributors

Dr Jane Uygun, MD CCFP MClSc is a Clinical Competency Tutor in the Department of General Practice, RCSI, Dublin, Ireland. She is also a PhD candidate in the Western Family Medicine program, London, Canada

Dr Ellen Stuart, MB BCH BAO BSc (Physiotherapy) MICGP DME DipWH Grad Dip (Clinical Education) Grad Cert (Palliative Care) is a Clinical Lecturer in the Department of General Practice, RCSI, Dublin Ireland

Dr Muireann De Paor, MB BCh BAO, BMedSc, MICGP, DOWH, DMSM is a Clinical Lecturer and Research Fellow in the Department of General Practice, RCSI, Dublin, Ireland

Dr. Emma Wallace MB BAO BCh (Hons), BMedSci (Hons), MICGP, PhD, HDip(Clin Ed), DCH is a Senior Lecturer in the Department of General Practice, RCSI, Dublin, Ireland

Dr Seamus Duffy, BComm, MB, MCh, BAO, LRCPSI is a Clinical Lecturer in the Department of General Practice, RCSI, Dublin, Ireland

Marié T. O'Shea, BA, M.Sc is a Primary Research Officer in the Health Professions Education Centre, RCSI, Dublin Ireland

Professor Susan Smith, MD, MSc MB BCh BAO DCH MRCPI MRCGP is a Professor in the Department of General Practice, RCSI, Dublin, Ireland

Professor Teresa Pawlikowska, BSc MB BS MSc PhD MRCP is Director of the Health Professions Education Centre, and BEME Coordinator for the BEME International Collaborating Centre, RCSI, Dublin, Ireland

## Introduction

Reflection has become a widespread topic of discussion amongst medical educators over the past decade and the literature relating to methods of teaching reflection has grown accordingly (Nelson and Purkis 2004; Walk et al. 2009; Murdoch-Eaton et al. 2014).

Reflection is thought to be an essential skill for competent healthcare professionals who are working with increasingly complex patients in correspondingly complex healthcare systems (Epstein 2008). As a result, evidence of reflective practice is becoming part of licensing and revalidation processes (College of Family Physicians of Canada. Maintenance of Proficiency Credits. Understanding Mainpro-C Credits ; General Medical Council. Supporting Information for Appraisal and Revalidation 2012 ; Stanford School of Medicine, Graduate Medical Education, Core Competencies). Despite this emphasis, however, there is little concurrence about how to best teach reflection.

Reflection is a complex construct and subsequently one of the challenges in this area has been lack of consensus around its definition. However, a recent systematic review of the literature by Nguyen et al established five core components and two extrinsic elements to reflection (Nguyen et al. 2014). In their operational definition they outline the five core components as follows:

*“Reflection is the process of engaging the self (S) in attentive, critical, exploratory and iterative (ACEI) interactions with one's thoughts and actions (TA), and their underlying conceptual frame (CF), with a view to changing them and a view on the change itself (VC).”*

In their conceptual model of reflection, they add to the five core components by describing the two extrinsic elements that impact reflection, the trigger and the context. This definition distinguishes reflection from other thinking processes and illuminates the extrinsic elements that inform and refine instances of reflection.

This process of exploring one's thoughts and actions as described in this definition has been seen by educators to have many benefits. Reflection can complement experiential learning by helping to identify learning needs, therefore leading to clinical competence (Shon 1983; Boud et al. 1985). New knowledge can be integrated with previous knowledge in this process. Professionalism can also be enhanced through reflection on personal beliefs, attitudes, values and needs and through self-regulation and monitoring (Bandura 1986; Epstein 1999).

There is increasing evidence to support the use of reflection throughout medical training. For example, higher scores on the Reflection in Learning scale and in reflection self-efficacy were linked to better academic performance in second year medical students (Sobral 2001). Similarly, reflection and re-visitation improved clinical performance with standardized patients in third year medical students in a study by Blatt et al. (2007). Internal medicine residents, studied by Mamedes et al. (2008), were more accurate when diagnosing complex,

unusual cases if they were practicing reflection. Finally Toy et al. (2009) found that residents were more likely to achieve their rotation goals when using reflective practice.

Evidence for the use of reflection is increasing and reflection is now considered by many to be an essential aspect of lifelong learning (Menard and Ratnapalan 2013). Reflection is becoming a core clinical skill in undergraduate medical curricula, as exemplified by the recommendations from the Association of American Medical Colleges (Association of American Medical Colleges, Recommendations For Clinical Skills Curricula For Undergraduate Medical Education). Internationally, the structure of continuing professional development in medicine is also placing an increased emphasis on reflection skills. For example, the College of Family Physicians in Canada and the General Medical Council in the United Kingdom (UK) now require doctors to produce evidence of reflection as part of their licence renewal process (College of Family Physicians of Canada. Maintenance of Proficiency Credits. Understanding Mainpro-C Credits; General Medical Council. Supporting Information for Appraisal and Revalidation 2012, Cutrer et al. 2017).

Despite the fact that reflection is now being used for licensing renewal there is no clear consensus in the literature about what methods are effective for teaching reflection. A systematic review investigating reflection and reflective practice in health professional education from 2009 found only ten studies in the literature which investigated the development of reflective thinking or the contextual influences which facilitated or deterred the development of reflection skills (Mann et al. 2009). A scoping literature search undertaken for this work highlighted that there has been a substantial increase in studies examining reflection since this last systematic review.

The aim of this review was to determine: i) which educational interventions are currently being used to facilitate the development of reflection, ii) how is reflection being assessed, and iii) what are the most effective interventions. Secondary research questions were also set to provide some context as to how reflection is currently being taught. These included: i) where in the curriculum are such interventions offered or recommended, ii) which faculty are generally responsible or endorsed for introducing such interventions, and iii) what are the barriers to using these interventions.

## **Review methodology**

The purpose of this review was to both understand the landscape of current practice and to consider effectiveness therefore it was based on both a constructivist and a positivist paradigm. A search to capture the full breadth of teaching practices was also expected to result in a heterogeneous group of articles. Consequently, both a systematic review and narrative synthesis of the literature was conducted to address the research questions.

The study methods followed a BEME-approved study protocol. The review is reported here in accordance with the STORIES guidelines for healthcare education evidence synthesis

(Gordon et al. 2014). These guidelines were developed through a review of existing guidance in the literature and a modified Delphi process to offer a guide for reporting evidence synthesis in health education.

### ***Types of interventions***

As this is an exploratory review, intervention types were not predefined.

### ***Types of participants***

This review was aimed at understanding how reflection is being taught to undergraduate medical students at any point in the undergraduate curriculum. The undergraduate curriculum ranges from four to six years around the world.

### ***Types of studies***

Quantitative (systematic reviews, randomised and nonrandomised control trials, cohort studies, cross-sectional studies, case control studies), case reports, descriptive studies, theses, qualitative and mixed studies were included in this review.

### ***Types of outcome measures***

Educational outcomes of the included papers were assessed using the Kirkpatrick's framework for educational outcomes adapted by Steinert et al. (2006). This version includes Barr's adaptations for medical education research and Steinert's subdivision of the original Level 3 into self-reported (3a) and observable (3b) changes in behaviour (Barr et al. 2000; Kirkpatrick and Kirkpatrick 2006; Steinert et al. 2006). See Table 1. Although, there are criticisms of the usefulness of Kirkpatrick's framework in medical education research, it has been used in many BEME reviews and has been widely used in other non-BEME medical education reviews (Steinert et al. 2006; Yardley and Dornan 2012).

Insert Table 1 Kirkpatrick Framework

### ***Search Strategies***

Several members of the review group had experience evaluating reflective writing and developing methods to teach reflection. This was helpful, during the preliminary scoping search, to determine whether the breadth and depth of teaching in reflection was being adequately captured. Appendix 1 lists a sample search string used for Pubmed.

The initial pilot search was conducted in 2015 in Pubmed and resulted in 3806 articles being retrieved. Search strings were adapted to capture the most pertinent articles. The review team agreed that the resulting search comprehensively captured relevant literature. Inclusion and exclusion criteria were further refined by discussing articles that reviewers found difficult to designate.

Databases were chosen to capture all possible interventions supporting reflection in medical education. The search strategy was designed with the assistance of two expert Medical

Education librarians. The search string was standardised and set to err on the side of maximising inclusivity without producing an unreasonable number of abstracts to review. The following databases were chosen for the search: Medline, Embase, ERIC, psychinfo, BEME published reviews, Cochrane, DARE, Web of Science and Scopus. The search period listed was not time limited in order to maximise inclusivity. The search was run on March 2, 2017.

Searches run in Google scholar were bringing up material that was too broad due to multiple meanings for reflection. Therefore, it was decided not to search further in the grey literature. However, BEME and Cochrane data bases were searched. In addition, reference lists from all papers were examined with respect to inclusion criteria. Any relevant papers that were identified were compared with the initial search results and added if they were not included in the initial search.

### ***Inclusion/exclusion criteria***

The following criteria were used to select articles for this review:

*Population:* Undergraduate medical students, however, if study subjects were inter-professional, studies were included if medical students were part of the group targeted for the intervention.

*Intervention type:* The study intervention had to involve a method of developing students' reflective skills, even if this was not the sole focus.

*Assessment:* The intervention regarding reflection needed to be assessed. Articles with limited description of the educational intervention or limited assessments were included if they provided answers to some of the secondary questions of the study.

*Study type:* Qualitative, quantitative and mixed methods studies were included.

*Exclusion criteria:* Studies focusing on reflection in professions other than medicine or on medical postgraduates, rather than undergraduates, were excluded. As the focus of the review was to ascertain which interventions have been shown to be effective for developing reflective skills in these students, viewpoint, editorial, opinion and descriptive papers were excluded. Dissertations and theses were included.

*Language and country:* Articles were not excluded due to language or country where intervention took place.

*Date of review:* Articles were not excluded due to publication date as there has not been a review specifically for undergraduate medical students in this area.

### ***Screening and Review of Abstracts***

Four team members (MdP, SD, ES and JU) independently assessed abstracts identified in the initial searches for eligibility of full text reviewing using Covidence software (Covidence 2018). To ensure that coding was consistent, team members discussed and compared a selected sample. A fifth team member (EW) reviewed articles that other team members were unsure whether to include, and made a determination around eligibility. Team members then reviewed each other's articles ensuring that two independent reviewers agreed that each article was suitable for inclusion. Abstracts were excluded at this stage if they were not relevant to the topic, and a new EndNote library was created for eligible articles only (EndNote 2013). An EndNote library of the total bibliography of abstracts was retained for reference and the Covidence file indicated which studies had been excluded.

### ***Data extraction***

Full text papers were acquired for articles that met the inclusion criteria. These papers were then again reviewed for inclusion in the final data extraction by five authors independently in order to ensure relevancy.

Using a similar method described in protocols for BEME reviews, five study authors independently reviewed included articles using a modified BEME Coding sheet which was adapted and developed following the initial pilot to ensure comprehensiveness of the tool. Reviewer's data extraction was validated for accuracy by providing 20% of coded papers to a moderator (T.P.) to assess for inter-rater reliability. A kappa of 0.8 indicates strong agreement with reliability of data falling between 64-81% (McHugh 2012). Any disagreements were discussed between two reviewers and if they were unable to agree, moderation was sought from a third member to reach consensus.

### ***Appraisal of studies***

Studies using quantitative measures were appraised using criteria developed by Buckley in their 2009 BEME review (Buckley et al. 2009) and further developed by Barrett in their 2016 BEME review (Barrett et al. 2016). Each criterion was independently marked as "met," "unmet," or "unclear." Seven of these 11 quality indicators needed to be met by the study to be deemed of high quality. The ethical aspect of these criteria were further developed by Barrett in their BEME review and this additional criterion was also evaluated in the quantitative studies (Barrett et al. 2016). Studies using qualitative measures were appraised using the Critical Appraisal Skills Program (CASP) as this has also been used in other BEME reviews (Critical Appraisal Skills Program (CASP)). Each quality indicator for the CASP score was answered as "yes" "no" or "can't tell" to appraise each study. Finally, the overall methodological strength of studies was then graded using the coding form "Strength of Findings" model as done in other BEME reviews (Buckley et al. 2009). In this model grade one indicates 'no conclusions can be drawn', whereas grade five indicates 'results are unequivocal'. Finally, the reviewer's overall impression was rated as either poor, acceptable,



good or excellent. Synthesis of studies incorporated discussion of methodological quality to address the issue of weaknesses in some of the studies with poor methodologies.

### ***Data synthesis***

The study designs for assessing educational interventions to teach reflection were very heterogeneous and as a result a meta-analysis of quantitative studies was not possible. There were a wide range of interventions and varied interpretations of the definition of reflection. Therefore, the studies identified were approached with a narrative synthesis to accommodate this diversity (Pope et al. 2006). Studies were initially grouped into qualitative and quantitative studies for analysis. As themes emerged, studies were grouped accordingly for further synthesis. Using their professional judgment reviewers assessed the contribution of a given study in answering the review research question, ‘How can reflection be taught?’ looking at studies in a holistic manner. Reviewers also noted the conceptual, theoretical, or methodological basis for any teaching method described, and the quality and appropriateness of the study. The aim of the synthesis was to provide educators with an improved understanding for selecting appropriate methods to teach reflection to undergraduate medical students, as well as the current state of the teaching of reflection in undergraduate medical education.

## **Results**

### ***Selection of papers***

Database searching identified 8047 references on reflection in medical education, of which 2943 were duplicates, leaving 5104 articles for initial screening. After initial screening 4895 articles were excluded, leaving 209 articles to review. Full text articles were obtained and of these 28 were eligible for synthesis of evidence on interventions to teach reflection in medical education. No study was excluded on the basis of quality throughout this process; methodological quality was evaluated following agreement on the final included set of articles. Figure 1 presents the flow diagram through the review process, indicating numbers of articles reviewed and retained at each stage.

Insert Figure 1 – PRISMA Diagram

Regarding data extraction, agreement was found to be within an acceptable range (kappa of 0.80).

### ***Overview***

The 28 core journal articles were published between 2005 and 2017. There were 13 qualitative studies, 11 quantitative studies and four studies using mixed methods. Two authors had two studies included in the synthesis (Aronson et al. 2011; Aronson et al. 2012; Lutz et al. 2013; Lutz et al. 2016). The majority of studies (n=13) took place in the USA. By continent, 15 studies came from North America, six from Europe, five from Asia, one from

Australia, and one from Africa. The three tables below summarize the studies by methodological approach (Table 2 Quantitative studies, Table 3 Qualitative studies and Table 4 Mixed methods).

Insert Table 2 – Quantitative studies grouped by level of evidence

Insert Table 3 – Qualitative studies

Insert Table 4 - Mixed method studies

### ***Context***

Twenty interventions took place during the clinical years of undergraduate medical school and eight took place in the preclinical years. Some of the interventions in the clinical years occurred while students covered several rotations, others were specific to a clinical attachment. Of these, four were during the medical rotation, three in obstetrics and gynaecology, two in paediatrics, one in primary care and two were in community placements. Some of the interventions were associated with specific courses or modules. Two interventions were part of humanities courses. Other courses associated with interventions included professionalism, communication skills, clinical skills, physical diagnosis, anatomy and research. In ten of the studies the intervention was either an elective or part of a pilot project and not associated with the core curriculum.

In seven studies, it was unclear which faculty members were involved in the intervention. The faculty of medicine was involved in eleven of the interventions, general practice was involved in three, paediatrics in two. Obstetrics and gynaecology, radiology, anatomy and primary care faculty were each involved in one intervention.

The length of the intervention and the number of teaching hours ranged greatly between studies. The length of the intervention ranged from four weeks to three years. The number of teaching hours was sometimes difficult to ascertain but ranged from approximately one hour to 12 hours.

### ***Interventions***

The interventions were heterogeneous and each had at least two of the following components related to reflection: i) introduction, ii) trigger, iii) writing, iv) guidelines, v) small group discussion, vi) tutor and vii) feedback. Triggers (n=23) and writing (n=22) were the most common components used in interventions. Analysis of intervention components was based solely on information provided in the studies.

#### ***Introduction (n=9)***

Introductions ranged from simply providing a definition of reflection to two hours of lecture material on the reflective process. Providing a definition or simple instruction was done in two studies. Verbal information on reflection or the course was provided in three studies.

Two studies held extensive interactive sessions. Introductions may have been present in other studies but were not explicitly described.

### *Trigger (n=23)*

The most common trigger for reflection was a clinical experience (n=14). Some reflections were initiated as part of a clinical rotation, while in others a specific clinical interaction was organized for the intervention. Other triggers included writing prompts (n=4), subject matter related to professionalism (n=3) or literature (n=2), research experiences (n=1), a virtual patient (n=1) and a test for bias (n=1).

### *Writing (n=22)*

Writing was described as a reflective assignment or essay in ten, a portfolio in three and a journal in two of the studies. For six of the interventions writing was done online. Five of these were in either a blog or discussion forum format. The sixth involved writing small reflections at points during a virtual patient interaction. In six of these interventions students only wrote one reflection. Two to 15 reflection accounts were required from students in other interventions.

### *Guidelines (n=12)*

Guidelines were described by some more broadly, as prompting questions or templates, while others gave specifics of the questions used. Four examples of specific guidelines that were used are: Learning from your Experiences as a Professional (LEaP) guidelines (Aronson et al 2011), Narrative Reflection Tool (Peterkin et al 2012), Gibbs cycle (Gibbs 1988) and Critical Incident Technique format (Flanagan 1954).

### *Small Group Discussion (n=17)*

The size of small groups ranged from three to ten students. In seven of these studies group function was guided by a specific method including Balint (n=2), Learner Centred approach (n=2), Check-in Peer support (n=1), Brookfield steps of reflection (n=1) and Clinical reflection training (n=1). The focus of the group discussions included: i) experiences with patients (n=6), ii) building skills in reflection (n=5), or iii) specific topics; professionalism (n=1) mind-body skills (n=1), managing bias (n=1), learning goals (n=1) and research (n=1). In two studies the small groups operated as online discussion forums. In one study the groups were peer led.

### *Tutor (n=20)*

The role of the tutor varied a great deal and was sometimes described as a facilitator or as a mentor. The role of the tutor ranged from overseeing to providing feedback and from facilitating to counselling. Interaction between students and tutors happened online, in groups and one-on-one. In some studies tutors engaged once in the intervention and in others engagement occurred several times during the intervention.

### *Feedback (n=17)*

In most of these studies it was not entirely clear what the nature of the feedback was and whether it related directly to the reflective process. Feedback was generally provided by the tutor but in some studies feedback also came from peers (n=7). Feedback was either face to face, written or online. In nine studies there was some indication that tutors or peers were given some training or guidelines on how to provide feedback. None of the studies described using a specific rubric but Aronson et al (2012) did describe a protocol they used for feedback.

Table 5, below, portrays the different components present in each study.

Insert Table 5 – Intervention components

### ***Barriers***

In 12 studies barriers to interventions were discussed. Both student and tutor resistance were mentioned and some of this related to the additional burden of a reflection intervention in an already full curriculum. Timelines to complete interventions and for students to submit assignments were also seen as a challenge. Students questioned the relevance of reflection in only four studies. Students attitudes were generally positive toward interventions but time commitment, workload and relevance to passing exams were the main focus of student concerns. Interventions using information technology (IT) noted that set up, arising technological problems and faculty user comfort were problematic. IT platforms could provide anonymity but were also felt to be more impersonal. A combination with an introduction or small group session prior to online activity was proposed to circumvent this. Faculty training to prevent biased feedback and develop skills to manage small group dynamics were also mentioned as ways to overcome barriers.

### ***Assessment***

Fourteen of 28 studies provided clear guidelines about how reflection was assessed. For example, one study used a validated instrument to measure student critical reflection before and after facilitated case discussions to improve professionalism in anatomy (Wittich et al. 2013; Spampinato et al. 2014) The rest of the assessment guidelines involved written reflections. There were three validated rubrics used: i) Reflective Ability Rubric (RAR) (Learman et al. 2008), ii) Groningen Reflection Ability Scale (GRAS) (Aukes L et al. 2007) and iii) Reflection Evaluation for Enhanced Competencies Tool (REFLECT) (Wald et al. 2012). RAR and REFLECT were used in three studies and GRAS in two. In Peterkin et al's study both the RAR and REFLECT rubrics were used (Peterkin et al. 2012). Three studies developed their own rubrics (Liao and Wang 2016; McEvoy et al. 2016; Devi et al. 2017). In two other studies, assessment of written reflections was described as being based on Mezirow's descriptions of reflectivity.

None of these studies indicated that they provided the students with the rubrics for assessment. However, the GRAS is a self-assessment scale and would have been completed by the students themselves.

In nineteen studies reflection was assessed formatively. Three studies had both a formative and summative assessment and one study used only summative assessment. In five studies it was not entirely clear whether the assessment was formative or summative.

### ***Quantitative studies: Risk of Bias***

Of the 16 studies with a quantitative evaluation, five were considered to be of good quality, nine acceptable and two were poor. All studies were prospective, had clear research questions, suitable study subjects and appropriate methods used in the analysis. Data collection methods were a concern in three studies and five of the studies lacked complete data, mostly due to high attrition rates. Risk of bias was the topic that was most poorly addressed with only five of the 16 studies addressing this issue. Two studies drew conclusions that were not entirely in line with the data and two studies were weak on their reproducibility due to lack of detail. Ethical issues were not clearly managed in six studies and triangulation was also not used in five of the studies.

### ***Qualitative studies: Risk of Bias***

Nineteen studies had a qualitative evaluation. Of these four were rated by coders as good quality, fourteen as acceptable and one as poor. Study design, recruitment data collection and findings were considered acceptable in all studies. Study aims and data analysis were not clearly stated in two studies. Methodology was not considered appropriate in one of the sixteen studies. In two studies, the value of the research was difficult to discern due to other methodological quality issues. There were two areas that were not commented on in many studies. Ethical issues were not clearly described in six studies and not discussed at all in five studies. In addition, the relationship between researcher and participants was not clearly articulated in eight studies and not addressed well in six studies. Although ethical permission is often in the associated material and is assumed by the publication process, delineating the steps taken to prevent coercion and ensure reflexivity of the researcher is useful in determining the methodological quality.

### ***Outcomes***

Educational outcomes were classified using the Kirkpatrick framework as adapted by Steinert et al (2006). Three studies were ranked as level 1, meaning that outcomes covered student reactions to interventions only. The majority of studies (n=17), were ranked as 2a, indicating student acquisition of knowledge and skills. Eight studies demonstrated modification of student attitudes and skills, level 2b. None reached the higher levels of three and four.

### ***Data synthesis***

Aronson et al. (2012) in their randomized trial comparing reflective writing with or without a **guideline** (LEaP) provide strong evidence for this intervention. This study follows on from the pilot study in 2011 comparing reflective writing with or without LEaP guidelines. The use of guidelines for reflective writing are further supported by findings in qualitative studies by Chretien et al. (2008) (structured blog), Makoul et al. (2010) (templated online forum posts) and Ozcarar et al. (2009) (structured portfolio).

Aronson et al.'s (2012) randomized trial also looked at the impact of **feedback** on reflective writing. The importance of feedback in enhancing reflection is also reinforced by findings in three qualitative studies (Carr and Carmody 2006; Makoul et al. 2010; Wen et al. 2015).

Devi et al. (2017) studied the impact of a two-hour **introduction** module on student reflection in a nonrandomized experimental trial. Although this proved effective, no other studies in the core articles specifically looked at this type of intervention. However, Hayton et al. (2015) provided a ninety-minute workshop with an introduction, small and large group activities and this was found to improve student reflection compared to control.

In a historically controlled study Spampinato et al. (2015) studied the effect of a **small group** intervention around professionalism in anatomy, on reflection. Reflection did not improve in the intervention group in this study. This may have been related to ethical restrictions which limited pairing pre- and post- intervention scores, therefore diminishing the ability to detect a difference with the intervention. However, Duke et al. (2015) found an improvement in reflection with a combined virtual and in person small group intervention using pre- and post-test methodology. Peterkin et al.'s (2012) pilot with combination online and in person small groups also demonstrated effectiveness in improving reflection. In addition, Liao et al. found that heterogeneous small groups produced deeper aspects of reflection than non-heterogeneous groups. Qualitative studies by Lutz et al. (2016 and 2013) and O'Neill et al. (2016) also indicate that small groups may improve reflection, decrease stress, increase self-care and lead to professional development.

Aukes et al. (2008) concluded from their pre- and post-test study that experiential learning was an effective **trigger** to enhance reflection. Use of personal experience as a trigger was also an integral part of Duke et al.'s (2015) intervention, mentioned above. Qualitative studies by Beylfield et al. (2005), Lutz et al. (2016 and 2013), and Plack et al. (2010) also support the importance of having personal experiences as triggers for reflection. These studies used clinical experiences in clerkship years or early patient contact in the preclinical years as a trigger for reflection.

Use of **online** formats to promote reflection were supported by a number of studies. In addition to Chretien et al. (2008), Makoul et al. (2010) and Ozcakar et al. (2009) mentioned above, Peterkin et al.'s (2012) pilot had both a live and online component which was shown to enhance reflection. Salminen et al. (2014) also developed a virtual patient with reflective prompts built in. This pilot was well received by students.

Qualitative studies also pointed to a number of **benefits** of reflection. Findings indicated that students may become more self-aware which could increase self-care and decrease stress. Other possible outcomes of reflection included a decrease in bias and more creative ways to communicate in difficult clinical situations.

## Discussion

The focus of this review has been to determine what educational interventions are effective for teaching reflection to medical students. There has not been a systematic review of the reflection education literature since 2009 and none specific to undergraduate medicine (Mann et al. 2009). Over the past decade there has been a significant increase in the number of articles in the literature on the topic of reflection in medical education. The core articles used in the analysis of this review span the past twelve years, do not overlap with the previous review and highlight the changes in medical education since this last review.

Nguyen et al.'s analysis in 2014 provided a much-needed definition of reflection to address the lack of consensus in the literature on this topic as it proliferated. During the process of assessing articles for eligibility in this review, this lack of consensus was evident. The term reflection was used in the medical literature in a number of different ways. At the more simplistic end of things reflection was used to describe the process of reviewing learning material or previous tasks. More sophisticated uses of the term reflection indicated a process of self-evaluation associated with the review. And finally, in others reflective analysis was viewed as more iterative and self-intimate. Adopting a standard definition for future studies will be useful in guiding educators and researchers along common pathways.

Another insight from the process of assessing articles for eligibility in this review was that many educators were using reflection without teaching students how to reflect. Students' written reflections were used in many studies as a means of gaining insight into student experiences during educational interventions aimed at topics other than reflection. Significantly, there were numerous studies in which students were assessed on their ability to reflect with no intervention to teach or facilitate reflection. Findings from core articles in this review indicate that several types of interventions can improve reflective ability and therefore indicate that reflection is indeed a skill that can be developed rather than a purely innate ability. The process for this review has highlighted that this knowledge is not widespread in the medical education community and this is supported by a recent study done by Butani et al. with paediatric medical educators (Butani et al. 2017).

The core studies in this review were heterogeneous both in methodology and design of educational interventions. The heterogeneity of educational interventions may be due to the complexity of reflection itself, the lack of consensus around the definition of reflection and variability in educators understanding of the reflective process. However, other factors likely relate to the context that these studies took place in within the medical school curriculum. This review found that reflection is being used across the curriculum, in pre-clinical and clerkship settings, in core programs and electives and in association with many different faculties. These factors can all have an impact on the intervention design. However, for the purposes of studying the effectiveness of interventions on reflection, it is clear from this review that breaking down and simplifying the interventions can improve the quality of evidence. Aronson et al.'s (2012) randomized trial comparing the effect on reflective writing of providing guidelines versus control, and providing feedback versus control exemplifies the kind of study that will be most helpful for furthering our understanding of the efficacy of interventions to teach reflection.

Although research into interventions to teach reflection would benefit from concentration on one or two components, it is likely that a combination of these components will produce better results, due to the complexity and iterative nature of reflection itself (Nguyen et al. 2014). The review findings indicate that it is important for students to receive some assistance in navigating the complexity of reflection and that they benefit from learning about reflection through introductions, guidelines to writing and by receiving feedback on their work (Aronson et al. 2012, Devi et al. 2017). In addition, this review has highlighted that personal experience is an effective trigger to engage students in the process of reflection. This is not surprising considering “experience” as a trigger for reflection was identified by Boyd and Fales in the 1980’s and it is also an integral part of Nguyen et al’s recent definition of reflection (Boyd and Fales 1983; Nguyen et al. 2014).

There is some discussion in the literature about the vulnerability that is required for the reflective process (Arntfield et al. 2016). Supportive tutors and peers in small groups have the potential to enable student vulnerability and evolution through the reflective process as indicated by some studies in this review. However, it is important to note that tutor skills in managing group dynamics and process are likely to be key factors in whether small groups are successful in this regard (O’Neill et al. 2016). In addition, one of the most powerful influences on student achievement has been shown to be feedback between the instructor and learner (Hattie 2008). The positive outcomes of online interventions with reflection raises certain questions. Does the relative anonymity of on-line communication facilitate students sharing vulnerabilities? Or is it simply the ease, flexibility and support which on-line forums affords them which facilitates the lowering of their guards?

Another issue illuminated by this review was the landscape of assessment tools currently being used. Although validated tools were used in several studies, many authors were developing their own rubrics for assessing reflections. It can be argued that specific rubrics may need to be developed for multifaceted intervention in specific contexts (Cook et al. 2015). However, a systematic review concentrating on the validity and reliability of available tools to assess reflection would be useful to educators and researchers in this area. This is also compelling considering that there has been some concern in the literature around the reliability of some of these validated tools (Andersen et al. 2014; Moniz et al. 2015)

Finally, the level of evidence of studies in this review was limited to Kirkpatrick levels one and two. Future research is critical in this area to ascertain the value of reflection exercises. Studies need to demonstrate not only the impact on improving reflection in practitioners in their workplace, but also whether reflective practitioners can impact patient care itself. With respect to the Kirkpatrick framework itself, its suitability for appraising interventions in medical education has been questioned (Yardley and Dornan 2012). Although commonly used as an evaluation framework for educational interventions, the Kirkpatrick framework was originally constructed with a management lens for the evaluation of training inputs in an industrial or business setting, with a more instructional model of learning. When translating between complex environments the inherent assumptions behind Kirkpatrick’s framework means that its application in areas of education with multiple interfaces and influences can become difficult, and so the upper levels are rarely reached. Supporting learner centred



reflection is also deeply contextualised, dynamic and developmental, which can make it challenging to capture these elements using this framework.

### ***Strengths and limitations of the systematic review***

The research team for this review had a broad range of experience and expertise. Several of the reviewers were involved in teaching reflection to undergraduate students. In addition, a number of our reviewers had experience in BEME and/or Cochrane reviews. Consultation with two expert Medical Education librarians also ensured that our search strategy was systematic and rigorous. Finally, the review team performed checks for quality during several stages of the review to confirm that core articles met inclusion criteria.

Although a broad selection of interventions was captured, the heterogeneity of the interventions and of study designs limited the analysis to a narrative synthesis. Therefore, it was only possible to comment on the evidence for certain components. Issues of study quality in some of the studies also made it difficult to qualify their contribution to the literature.

### ***Implications for medical education, research and clinical practice***

The data synthesis of the core group of studies in this review, has provided a method of categorizing multifaceted education interventions to teach reflection. This may prove useful when designing new interventions to teach reflection in the undergraduate curriculum. In addition, this review has highlighted which intervention components have the best evidence for enhancing reflection to date. Evidence in this topic area is lacking and further studies looking at single component interventions would be of benefit to this body of knowledge. Furthermore, a systematic review of assessment tools for reflection is recommended. Studies in this review highlight possible benefits of reflection education for medical students, however, further research in this area is required. Recommendations for designing an education module to develop reflection are: i) introduce reflection and provide guidelines in order to increase the understanding of learners; ii) clearly describe the developmental process and provide feedback; and iii) assess the impact for learners. A realistic timeframe should be considered for all of these steps.

### **Conclusions**

This review highlights that there is a great deal of heterogeneity in the type of interventions and study designs looking at the teaching of reflection in medical students. In the studies reviewed, interventions to teach reflection had at least two of the following components: i) introduction, ii) trigger, iii) writing, iv) guidelines, v) small group discussion, vi) tutor and vii) feedback. The strongest evidence indicated that guidelines for, and feedback on, reflective writing improved student reflection. However, evidence is still lacking on how reflection exercises improve reflection in practice or impact patient care itself. There are many rubrics being used to assess reflection and in this review three rubrics used to assess reflective writing were validated. A future systematic review concentrating on validity and reliability of available tools would be useful for educators.

## Acknowledgements

We would like to acknowledge the irreplaceable expertise of Ms Grainne McCabe and Mr Paul Murphy, the information scientists who contributed to this review. We would also like to thank the BEME team for their guidance through the course of the review.

## Disclosure statement

The authors have no declarations of interest to report. This BEME review was completed without funding.

## References

- Andersen N, O'Neill L, Gormsen L, Hvidberg L, Morcke A. 2014. A validation study of the psychometric properties of the Groningen Reflection Ability Scale. *BMC Med Ed*. 14:214.
- Arntfield S, Parlett B, Meston CN, Apramian T, Lingard L. 2016. A model of engagement in reflective writing-based portfolios: Interactions between points of vulnerability and acts of adaptability. *Med Teach*. 38(2):196-205.
- Aronson L, Niehaus B, Hill-Sakurai L, Lai C, O'Sullivan PS. 2012. A comparison of two methods of teaching reflective ability in Year 3 medical students. *Med Educ*. 46(8):807-814.
- Aronson L, Niehaus B, Lindow J, Robertson PA, O'Sullivan PS. 2011. Development and pilot testing of a reflective learning guide for medical education. *Med Teach*. 33(10):e515-521.
- Association of American Medical Colleges, Recommendations For Clinical Skills Curricula For Undergraduate Medical Education.[Internet]. 2008. [cited 2017 Oct 30]. Available from: [https://www.aamc.org/download/130608/data/clinicalskills\\_oct09.qxd.pdf](https://www.aamc.org/download/130608/data/clinicalskills_oct09.qxd.pdf)
- Aukes L, Geertsma J, Cohen-Schotanus J, Zwierstra R, Slaets J. 2007. The development of a scale to measure personal reflection in medical practice and education. *Med Teach*. 29:177-182.
- Aukes LC, Geertsma J, Cohen-Schotanus J, Zwierstra RP, Slaets JP. 2008. The effect of enhanced experiential learning on the personal reflection of undergraduate medical students. *Med Educ Online*. 13:15.
- Bandura A. 1986. *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Barr H, Freeth D, Hammick M, Koppel I, Reeves S. 2000. *Evaluations of interprofessional education; a United Kingdom review of health and social care*. London: CAIPE/BERA.
- Barrett A, Galvin R, Steinert Y, Scherpbier A, O'Shaughnessy A, Horgan M, Horsley T. 2016. A BEME (Best Evidence in Medical Education) review of the use of workplace-based assessment in identifying and remediating underperformance among postgraduate medical trainees: BEME Guide No.43. *Med Teach*. 38(12):1188-1198.

Beylefeld AA, Nena KD, Prinsloo EA. 2005. Influence of community experiences on first-year medical students' reflective writing. *Med Teach*. 27(2):150-154.

Blatt B, Plack M, Maring J, Mintz M, Simmens SJ, 22, 49-5. 2007. Acting on Reflection: the Effect of Reflection on Students' Clinical Performance on a Standardized Patient Examination. *J Gen Intern Med*. 22(1):49-54.

Boud D, Keogh R, Walker D. 1985. *Reflection: Turning experience into learning*. London: Kogan Page.

Boyd R, Fales A. 1983. Reflective learning: key to learning from experience. *JHP*. 23(2):99-117.

Buckley S, Coleman J, Davison I, Kahn K, Zamora J, Malick S, Morley D, Pollard D, Ashcroft T, Popovic C et al. 2009. The educational effects of portfolios on undergraduate student learning: a Best Evidence Medical Education (BEME) systematic review: BEME Guide No. 11. *Med Teach*. 31(4):282-298.

Butani L, Bannister S, Rubin A, Forbes K. 2017. How educators conceptualize and teach reflective practice: a survey of North American pediatric medical educators. *Acad Pediatr*. 17(3):303-309.

Carr S, Carmody D. 2006. Experiential learning in women's health: medical student reflections. *Med Educ*. 40(8):768-774.

Chou CL, Johnston CB, Singh B, Garber JD, Kaplan E, Lee K, Teherani A. 2011. A "safe space" for learning and reflection: one school's design for continuity with a peer group across clinical clerkships. *Acad Med*. 86(12):1560-1565.

Chretien K, Goldman E, Faselis C. 2008. The reflective writing class blog: using technology to promote reflection and professional development. *J Gen Intern Med*. 23(12):2066-2070.

College of Family Physicians of Canada. Maintenance of Proficiency Credits. Understanding Mainpro-C Credits.; [accessed October 30, 2017]. Available from: <http://www.cfpc.ca/ProjectAssets/Templates/Column1b.aspx?id=698&>.

Cook D, Brydges R, Ginsburg S, Hatala R. 2015. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*. 49:560-575.

Covidence. [Internet]. 2018.[cited 2017 Oct 30]. Available from: <https://www.covidence.org/>.

Critical Appraisal Skills Program (CASP). [Internet]. 2013. [cited 2017 Oct 30]. Available from: <http://www.casp-uk.net/casp-tools-checklists>.

Cutrer WB, Miller B, Pusic MV, Mejicano G, Mangrulkar RS, Gruppen LD, Hawkins RE, Skochelak SE, Moore Jr DE. (2017). Fostering the development of master adaptive learners: a conceptual model to guide skill acquisition in medical education. *Acad Med*. 92(1): 70-75.

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.

Duke P, Grosseman S, Novack DH, Rosenzweig S. 2015. Preserving third year medical students' empathy and enhancing self-reflection using small group "virtual hangout" technology. *Med Teach*. 37(6):566-571.

EndNote X7 for Windows. 2013. Released May 20, 2013.

Epstein RM. 1999. Mindful practice. *JAMA*. 282(9):833-839.

Epstein RM. 2008. Reflection, perception and the acquisition of wisdom. *Med Educ*. 42(11):1048-1050.

Flanagan JC. 1954. The critical incident technique. *Psychol Bull*. 51:327-58.

General Medical Council. Supporting Information for Appraisal and Revalidation [Internet]. 2012. [cited October 30, 2017]. Available from: [http://www.gmc-uk.org/RT\\_\\_\\_Supporting\\_information\\_for\\_appraisal\\_and\\_revalidation\\_\\_\\_DC5485.pdf\\_55024594.pdf](http://www.gmc-uk.org/RT___Supporting_information_for_appraisal_and_revalidation___DC5485.pdf_55024594.pdf)

Gibbs G. 1988. *Learning by doing: a guide to teaching and learning methods*. Oxford: Further Education Unit, Oxford Polytechnic.

Gordon M, Gibbs T, BMC, 2014 M, 143. 2014. STORIES statement: Publication standards for healthcare education evidence synthesis *BMC Medicine*. 12:143.

Grant A, Kinnersley P, Metcalf E, Pill R, Houston H. 2006. Students' views of reflective learning techniques: an efficacy study at a UK medical school. *Med Educ*. 40(4):379-388.

Hattie J. 2008. *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. , Abingdon: Routledge.

Hayton A, Kang I, Wong R, Loo LK. 2015. Teaching Medical Students to Reflect More Deeply. *Teach Learn Med*. 27(4):410-416.

Learman L, Autry A, O'Sullivan P. 2008. Reliability and validity of reflection exercises for obstetrics and gynecology residents. *Am J Obstet Gynecol*. 198(4):461, e461-468.

Kirkpatrick DL, Kirkpatrick JD. 2006. *Evaluating Training Programs*, third edition. San Francisco: Berrett-Koehler

Liao HC, Wang YH. 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):234.

Lutz G, Roling G, Berger B, Edelhauser F, Scheffer C. 2016. Reflective practice and its role in facilitating creative responses to dilemmas within clinical communication - a qualitative analysis. *BMC Med Educ*. 16(1):301.

Lutz G, Scheffer C, Edelhaeuser F, Tauschel D, Neumann M. 2013. A reflective practice intervention for professional development, reduced stress and improved patient care--a qualitative developmental evaluation. *Patient Educ Couns*. 92(3):337-345.

Makoul G, Zick AB, Aakhus M, Neely KJ, Roemer PE. 2010. Using an online forum to encourage reflection about difficult conversations in medicine. *Patient Educ Couns*. 79(1):83-86.

Mamede S, Schmidt H, Penaforte J. 2008. Effects of reflective practice on the accuracy of medical diagnoses. *Med Educ*. 42(5):468-475.

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.

McHugh M. 2012. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 22(3):276-282.

Menard L, Ratnapalan S. 2013. Teaching moment: reflection in medicine: models and application. *Can Fam Physician*. 59(1):105-107, e157-109.

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. 2009. *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*. [Ann Intern Med, 151\(4\)](#).

Moniz T, Arntfield S, Miller K, Lindgard L, Watling C, Regehr G. 2015. Considerations in the use of reflective writing for student assessment: issues of reliability and validity. *Med Educ*. 49(9):901-908.

Murdoch-Eaton D, Sandars JAEC, 99:279-83. 2014. Reflection: moving from a mandatory ritual to a meaningful professional development. *Arch Dis Child*. 99:279-283.

Naeger DM, Hua EW, Ahearn B, Webb EM. 2015. Reflective Writing: A Potential Tool to Improve Interprofessional Teamwork with Radiologists. *Acad Radiol*. 22(10):1221-1225.

Nelson S, Purkis M. 2004. Mandatory reflection: the Canadian reconstitution of the competent nurse. *Nurs Inq*. 11:246-257.

Nguyen, QD., Fernandez N K, Charlin B. 2014. What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Med Educ*. 48:1176-1189.

O'Neill S, Foster K, Gilbert-Obrart A. 2016. The Balint group experience for medical students: a pilot project. *Psychoanal Psychother*. 30(1):96-108.

Özçakar N, Mevsim V, Güldal D. 2009. Use of portfolios in undergraduate medical training: First meeting with a patient. *Trakya Universitesi Tip Fakultesi Dergisi*. 26(2):145-150.

Patterson A, Sharek D, Hennessy M, Phillips M, Schofield S. 2016. Medical humanities: a closer look at learning. *Med Humanit*. 42(2):115-120.

Peterkin A, Roberts M, Kavanagh L, Havey T. 2012. Narrative means to professional ends: new strategies for teaching CanMEDS roles in Canadian medical schools. *Can Fam Physician*. 58(10):e563-569.

Plack MM, Driscoll M, Marquez M, Greenberg L. 2010. Peer-facilitated virtual action learning: reflecting on critical incidents during a pediatric clerkship. *Acad Pediatr*. 10(2):146-152.

Pope C, Mays N, Popay J. 2006. How can we synthesize qualitative and quantitative evidence for healthcare policy makers and managers? *Healthc Manag Forum*. 19(1):27-31.

Salminen H, Zary N, Bjorklund K, Toth-Pal E, Leanderson C. 2014. Virtual patients in primary care: developing a reusable model that fosters reflective practice and clinical reasoning. *J Med Internet Res.* 16(1):e3.

Saunders PA, Tractenberg RE, Chaterji R, Amri H, Harazduk N, Gordon JS, Lumpkin M, Haramati A. 2007. Promoting self-awareness and reflection through an experiential mind-body skills course for first year medical students. *Med Teach.* 29(8):778-784.

Shon D. 1983. *The reflective practitioner.* San Francisco: Jossey-Bass.

Sobral D. 2001. Medical students' reflection in learning in relation to approaches to study and academic achievement. *Med Teach.* 23(5):508-513.

Spampinato CM, Wittich CM, Beckman TJ, Cha SS, Pawlina W. 2014. "Safe Harbor": evaluation of a professionalism case discussion intervention for the gross anatomy course. *Anat Sci Educ.* 7(3):191-198.

Stanford School of Medicine, Graduate Medical Education, Core Competencies [Internet]. 2017. [cited 2017 Oct 30]. Available from: [http://med.stanford.edu/gme/current\\_residents/corecomp.html](http://med.stanford.edu/gme/current_residents/corecomp.html).

Steinert Y, Mann K, Centeno A, Dolmans D, Spencer J, Gelula M, Prideaux D. 2006. A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8. *Med Teach.* 28(6):497-526.

Teal CR, Shada RE, Gill AC, Thompson BM, Fruge E, Villarreal GB, Haidet P. 2010. When best intentions aren't enough: helping medical students develop strategies for managing bias about patients. *J Gen Intern Med.* 25 Suppl 2:S115-118.

Toy E, Harms K, Morris R, Simmons J, Kaplan A, Ownby A. 2009. The Effect of Monthly Resident Reflection on Achieving Rotation Goals. *Teach Learn Med.* 21(1):15-19.

Wald H, Borkan J, Taylor J, Anthony D, Reis S. 2012. Fostering and evaluating reflective capacity in medical education: Developing the REFLECT rubric for assessing reflective writing. *Acad Med.* 87:41-50.

Walk H, Davis S, Reis S, Monroe A, Borkan J. 2009. Reflecting on reflections: enhancement of medical education curriculum with structured field notes and guided feedback. *Acad Med.* 84:830-837.

Wen CC, Lin MJ, Lin CW, Chu SY. 2015. Exploratory study of the characteristics of feedback in the reflective dialogue group given to medical students in a clinical clerkship. *Med Educ Online.* 20:25965.

Wittich C, Pawlina W, Drake R, Szostek J, Reed D, Lachman N, McBride J, Mandrekar J, Beckman T. 2013. Validation of a method for measuring medical students'critical reflections on professionalism in gross anatomy. *Anat Sci Educ.* 6:232-238.

Yardley S, Dornan T. 2012. Kirkpatrick's levels and education 'evidence'. *Med Educ.* 46(1):97-106.



