The Journal of Education in Perioperative Medicine

Original Research

Characterization of Reflective Capacity of Anesthesiology Trainees in an Irish Tertiary Referral Teaching Hospital

Hassan M. Ahmed, MD
Audrey Dunn Galvin, PhD
Aoife O'Loughlin, MBBCH, MSc
Aisling O'Meachair, BA, MBBCH BAO
Jeffrey B. Cooper, PhD
Richard H. Blum, MD, MSE, FAAP
George Shorten, MD, PhD

Introduction

Reflection is a person's deliberate thinking, either "in the moment" or subsequently, on different experiences with a goal of better understanding of self and situation in order to enhance one's performance in the future.[1,2] Reflective capacity is one's ability to reflect, whereas reflective practice is the manner and extent to which one actively participates in reflection.[1] Reflective practice is an essential characteristic of professionally competent clinical practice.[3]

Previous experience is one aspect of professional development, but is not necessarily applied in a way that supports learning.[4] The deliberate fostering of reflective practice (reflecting on one's experiences regularly) by health care trainees and practitioners has been shown to result in superior critical thinking, enhanced professionalism, improved accuracy of medical diagnosis, and strong performance in complex situations, especially those associated with morally ambiguous circumstances.[5–9] Conversely, failure to engage in reflection can result in errors in diagnosis and lesser quality of patient management and care.[7] In line with this, failure to reflect on one's own practice and ideas was recently described as one of the reasons underlying physician overconfidence.[10]

Professional and regulatory bodies, such as the Accreditation Council for Graduate Medical Education (United States) and General Medical Council (United Kingdom), have recognized the integral role of reflective practice and have recently required evidence of reflective practice proficiency for revalidation of health professionals.[11,12]

Wald and Reis[3] described reflection as "not necessarily intuitive" and suggested that "educational interventions are warranted." To date, reflective practice is not embedded in most medical training programs; no consensus exists as to the optimal form of training in reflective practice. We have recently outlined a framework[13] in which the concepts set out by Kolb and Kolb,[14] Mann et al.,[15] and others are configured specifically with the practice of medicine and anesthesiology in mind. We have pointed out the need to characterize reflective practice among anesthesiology trainees to provide a basis for effective support within the training paradigm and to enable future comparison if constructive changes in training are introduced. To those ends, we have begun to conduct a "baseline" study of reflective capacity among anesthesiology trainees in Cork, Ireland.

The principal objective of this study was to characterize the reflective capacity of anesthesiology trainees in a tertiary referral Irish teaching hospital. Secondary objectives were (1) to determine if an association exists between years of medical training and reflective capacity; and (2) to compare (within and between individuals) levels of reflective capacity evaluated using responses to investigator-written vignettes provided to trainees with those to vignettes written by the trainees themselves and based on their own experience.

Materials and Methods

With institutional ethical approval by the Clinical Research Ethics Committee of the Cork Teaching Hospitals (May 21, 2019) and having obtained written informed consent from each participating trainee, the study was carried out at Cork University Hospital (CUH) and Cork University Maternity Hospital (CUMH) between May and December 2019.

In the Rater training section of the methods described below, we ensured that the following criteria were met to ensure reliability in training of assessors and scoring: (1) that elements are objectively defined, (2) that observers are adequately trained in the techniques being used, and (3) that understanding and agreement on scoring was assessed.

Evaluation of Reflective Capacity

Reflective capacity was evaluated and scored using the Reflection Evaluation for Learners' Enhanced Competencies Tool (REFLECT), which was developed and validated by Wald and colleagues.[5] This rubric was developed based on the theories of Schon and DeSanctis,[16] Boud and colleagues,[17] Moon,[18] and Mezirow.[19] Based on responses obtained from US medical students, Wald and colleagues[5] reported good reliability of the REFLECT rubric (Inter Class Correlation [ICC] single

continued on next page

continued from previous page

measure: 0.632). Standardized vignettes of patient situations that have elements of imperfect care were used and subjects were asked to answer specific questions that required them to write about their perspectives and opinions of the event. The elicited responses were scored by trained raters using the REFLECT rubric, which sets out the dimensions of the construct to be scored, defining benchmarks for each that represent advancing levels of reflection: Level 1, Habitual action (Nonreflective); Level 2, Thoughtful action or introspection; Level 3, Reflection; and Level 4, Critical reflection. Each of these levels, and the 5 criteria that constitute them, were clearly defined by Wald and colleagues.[5] Thus, each vignette is scored on (1) writing spectrum, (2) presence, (3) description of conflict or disorienting dilemma, (4) attending to emotions, and (5) analysis and meaning making.[5] We directed raters to use Wald's 4-step process for applying the REFLECT rubric: (1) read the response entirely, (2) zoom into details of the response through fragmentation into sentences and phrases to evaluate the presence and quality of each criteria, (3) zoom out to level the theme of the whole response taking into consideration the analysis in step 2, and (4) defend the level of assignment using examples from the response text. Raters first analyzed the text according to criteria 2 to 5 and finished their rating using criteria 1.

In our study, several investigators (ADG, RB, JBC, AO'L) created a library of 9 brief (300 words) vignettes (see Appendix 1 for example with response and supplementary material that contains all of the vignettes) that were plausible as encounters experienced by anesthesiology trainees of any level of training. Each vignette was developed based on strict criteria for content, and on relevant theory.[13] The vignettes used were selected based on an extensive review by an expert in psychometrics (ADG) to ensure they met the predefined goals set out in Appendix 2. Guidance on how we produced vignettes is attached as Appendix 2. Each participant was presented with 2 investigator-written clinical vignettes that contained a challenge in terms of decision making. These were selected randomly from the library of 6 vignettes (9 in total with 3 randomly selected for training purposes; see "Rater Training"). Each vignette was accompanied by the 3 questions (What happened? What are your thoughts and feelings about why this event happened? What could you have done differently?), to which the subject was invited to write a response, with no limitation in length. In answering the questions, the subjects were instructed to imagine that they were the clinician faced with the challenging event. Subjects could use either an electronic or paper version of the instrument and were requested to refrain from discussing the vignettes with others.

In addition, each participant was invited to write 1 or more vignettes based on his or her own experience and to provide written answers to the same 3 questions.

Subjects

All anesthesiology trainees, ranging from interns (within 1 year's training of graduation from medical school) to fellows (who had completed basic professional training and were undertaking higher subspecialty training), employed at CUH/CUMH for any part of 2019 were invited to participate in the study.

Rater Training

Two investigators (anesthesiology trainees, year 6 and year 1) underwent rater training during three, 90-minute face-to-face sessions with an expert in psychometric evaluation (ADG). The training included a detailed description of the 5 categories of the REFLECT rubric[5] and instruction on how to score each category.

Step 1: The raters were familiarized with the theories underlying the REFLECT rubric and scoring.

Step 2: The instructor provided examples of responses of different levels of reflection (and the elements on which each is scored) to illustrate correct scoring.

Step 3: Raters each scored exemplar vignettes (previously and independently rated by the instructor) separately and discussed discrepancies in a follow-up training session. This was continued in iterative fashion until agreement across raters and instructor was <0.5 on a 10-point scale on vignette scoring.

Step 4: Instructor and raters each scored 3 novel vignettes drawn from the pool of 9.

The scores in step 4 were examined for interrater reliability (between the 2 raters and between the average of 2 raters' scores and instructor score), using weighted kappa.

Vignette Scoring in Present Study

The 2 independent raters were blinded to the identity of the participating trainees whose responses they scored (by redaction of respondent name, age, sex, country of origin, and level of expertise). The order of responses scored by each rater was randomly generated (raters did not score responses to different vignettes written by an individual trainee in sequence, to avoid priming bias). A participant's responses for a given vignette were presented to raters and evaluated as a unit, according to the REFLECT rubric. The average of 2 raters' scores was used in the subsequent analyses.

Data Analysis

Each trainee was invited to respond to (1) 2 investigator-written vignettes prepared by the investigators and suitable for evaluation using REFLECT, and (2) to generate a written vignette based on their own experience. All responses were evaluated by 2 raters independently. Interrater reliability was assessed by weighted kappa coefficient. Internal or scale reliability was assessed by Cronbach's alpha. A mean score was calculated for each criterion. Because there was a strong positive correlation between each of the criteria ($r = 0.71$ to 0.85), a mean total score was also calculated. Because Wald suggested that the criteria 2 to 5 contributed to criteria 1 (writing spectrum), we analyzed the reliability of the scale with 5 and 4 criteria (ie, with "writing spectrum" and without). Because Cronbach's alpha for the scale was stronger with 5 criteria ($\alpha = 0.93$) compared with 4 criteria ($\alpha = 0.93$), the findings of the 5 criteria were used for subsequent analyses. Tests for difference (2- tailed t-tests or analysis of variance, as appropriate) were used to evaluate REFLECT scores, according to gender, years of training in medicine, and investigator-written versus self-written vignettes. Bonferroni correction was used to adjust *P* values because of the increased risk of type 1 error when using multiple statistical tests. Association between years of training in medicine and level of reflective

continued on next page

continued from previous page

capacity was examined using correlation and multiple regression analyses. Results were taken as significant if $P < .05$.

Results

Participant Characteristics

We invited 61 trainees to participate, of whom 29 complied. The participants were aged between 23 and 40 years; 70% male. Training in medicine ranged from 25 to 156 months, and in anesthesiology from 4 months to 132 months.

Analysis of the Responses

The distribution of the scores was normal, with only 0.1 difference between the mean and 10% trimmed means. Cronbach's alpha for the 5 elements of the REFLECT rubric was $\alpha = 0.93$, indicating excellent internal reliability for the scale. Weighted kappa and ICC were $k = 0.81$; $r = 0.86$, respectively, demonstrating very satisfactory interrater reliability. Because of the excellent agreement between the rater, mean of ratings was used here. The mean total REFLECT score for participants was 2.16 ± 0.7, corresponding to "thoughtful action," and indicating low to moderate reflective capacity. Participants provided 86 responses (58 responses to investigator-written vignettes and 28 responses to self-written vignettes). The individual scores for the REFLECT rubric criteria were writing spectrum 2.2 ± 0.9, presence 2.3 ± 1.0, description of conflict 2.4 ± 0.8, attending to emotions 1.6 ± 0.8, and analysis and meaning making 2.4 ± 0.8. Data were normally distributed and assumptions for regression analysis were met. Table 1 shows the mean scores for each of the 5 criteria of the Rubric, for investigator-written ($n = 58$) and self-written ($n = 28$) vignettes.

Women (2.6 ± 0.6) reported significantly greater total reflective capacity than men (2.0 ± 0.7) ($t(1) = 4.3$, $P = .001$).

The mean scores of participants' responses to the first and second investigator-written vignettes exhibited a remarkable consistency (except in 2 participants) (Appendix 3). There was no difference noted between the scores for responses to self-written vignettes and those for investigator-written vignettes (2.2 ± 0.7 vs 2.1 ± 0.7, respectively; $P > .05$). Scores of the "Presence" component were significantly greater for self-written vignettes versus for investigator-written vignettes (2.7 ± 1 vs 2.0 ± 0.9, respectively; $P < .05$) (Table 1); t-tests with Bonferroni correction identified a significant difference only on "presence" (mean = 2.0 vs 2.7, respectively [$t(4) = -3.04$, $P < .01$]).

Multiple regression analysis was used to test if "years of training in medicine" significantly predicted participants' total score (dependent variable), controlling for age and self- versus investigator-written vignette (Table 2). The results of the regression indicated "experience in medicine" significantly predicted the total REFLECT score (F 4.9, $P = .004$), demonstrating a decrease in reflective ability as experience increased. Age and self- versus investigator-written vignettes were not significant, although age trended in the same direction. The same pattern held (only "years of training in medicine" significant, $P < .05$) when each of the 5 criteria were entered as dependent variables in separate models, except for "presence," where age ($t = -2.06$, $P = .43$) and self- versus investigator-written vignettes ($t = 3.5$, $P = .001$) were significant predictors.

Next, "years of training in medicine" was re-coded according to quartiles (months experience 0-32; 33-65; 66-98; 99+). Analysis of variance showed that respondents with fewer years of training in medicine had greater means (indicating a greater level of reflective capacity) than those with longer experience ($F = 5.5$, $P = .02$). Post hoc analysis showed a mean difference of 0.7 ($P = .03$) between ≤32 months versus ≥99 months, and a mean difference of 0.8 ($P = .02$) between 33 to 65 versus ≥99 months (Table 3).

Discussion

This study is the first (to our knowledge) in which the reflective capacities of anesthesiology trainees are systematically examined and describes the use of the REFLECT rubric for their evaluation. We believe that the approach we describe here enables the establishment of a "baseline" from which the effect of future interventions to improve reflection can be evaluated. Overall, the reflective capacity score of participating anesthesiology trainees in CUH using the REFLECT rubric was 2.16 (SD 0.7), which corresponds to the "thoughtful action" level.

During the first iterative cycle of developing the REFLECT rubric,[5] Wald and colleagues proposed 5 overall levels of reflection that were progressive from nonreflective/habitual action (level 1) to transformative learning (level 5). Using randomly selected second-year medical students' field notes as examples of reflective writing, the distribution of students by overall level was level 1 = 0, level 2 = 17, level 3 = 38, level 4 = 28, and level 5 = 10. Refinement of the REFLECT tool decreased the number of overall levels to 4: Nonreflective: habitual action; Nonreflective: thoughtful action; Reflective; and Critically Reflective. Against this background, we note the large proportion of trainees in our study whose responses were classified as nonreflective. This apparently concerning result may be influenced by a lack of full investment by the subjects in the writing exercise or that the tool itself may not be measuring what we intend. The fact that participating trainees (29 of 90 invited) "opted in" and that almost all (28/29) completed 3 responses, including an optional self-written vignette, indicates that they were engaged by the task of writing. Based on our interpretation of the original validation of the REFECT rubric, we believe that its validity was established rigorously and, in a setting, largely equivalent to that of our study. It is important to note that in the absence of a "gold standard" for reflective capacity, we cannot interpret a particular evaluated level with complete confidence. Understanding the practical importance of a lesser level (using the REFLECT rubric) will require further work.

Although it might be expected that more exposure to patients enabled by more years of training in medicine would result in greater levels of reflective capacity, our results showed lesser reflective capacity was associated with more years of medical experience. This result indicates that years in training in a hospital environment tends to decrease reflective capacity. Although our data cannot provide an explanation for such an effect, it is broadly consistent with our previous observation that medical students over the course of the

continued on next page

continued from previous page

final undergraduate year (largely based in acute hospitals) appear to decrease in their reflective ability and moral reasoning.[20] Furthermore, Boenink and colleagues[21] reported lesser reflection level in medical students who had experience in health care work compared with those who did not. Possible explanations include the practical demands of the clinical training (exhausting as it can be attentional and emotional reserve), fatigue, or the tendency to avoid contemplating one's own underperformance in order to maintain self-confidence or disruption of peace of mind. This latter possibility might be greater as one assumes greater clinical responsibility and approaches independent practice.

Our results are consistent with findings of Blum and colleagues,[22] who identified critical gaps in anesthesiology training. Two of these gaps were reflection-dependent (planning for self-improvement and recognition of one's own limits), which did not improve with years in training. From a theoretical perspective, the dual systems theory of Tversky and Kahneman[23] could provide an explanation. Although reflection is often used intuitively, skills for deliberate reflective practice may need to be enhanced through the teaching of strategies and informed experiential learning. If reflective practice is not a habitual strategy used by an individual, then with increasing experience it becomes more likely that System 1 (fast, intuitive; according to Tversky and Kahneman[23]) reasoning will be used, which depends on the usual practice/unchallenged strategies that have become reinforced through years of experience.[16] Alternatively, such an effect may be a function of one's progress in training toward independent practice (eg, reluctance to dwell on poor performance) or of training-program ethos (eg, errors considered markers of substandard individuals and to be hidden or denied). We suggest that if reflective practice were to be embedded in the medical (or anesthesiology) curricula, it is more likely that System 2 reflective reasoning would be used appropriately in more complex and challenging situations.[13]

Although we do not have strong evidence of what interventions would be effective to improve reflective practice among medical (or anesthesiology) trainees, some interventions have been shown to improve reflective abilities for medical students. Early implementation of reflective practice teaching, particularly when associated with simulation excercises,[24] providing critical reflection guidelines for medical students,[25] and providing formative rather than summative feedback,[26] are all techniques that seem to be effective in promoting reflective abilities among medical students that could be adopted in anesthesiology training.

The average scores of each participant for responses to the 2 investigator-written vignettes are similar (except in 2 participants) (Appendix 3), which indicates remarkable consistency across vignettes. Interrater agreement was satisfactory, which could be attributed to using only 2 raters and intensive training. Our results are consistent with the findings of Wald and colleagues,[5] who reported a single ICC of 0.63 while scoring 60 narratives by 4 raters independently. In contrast, Grierson and colleagues[27] reported poor reliability of the REFLECT rubric when used to score the reflective writing of first-year medical students and clinician-faculty from the Department of Family Medicine by 2 raters independently. However, this could be explained by the cohort mixture in the work of Grierson and colleagues[27] to which the rubric was applied (medical students and family medicine clinicians) and insufficient rater training (2-hour workshop), which precluded achievement of a shared understanding of the REFLECT rubric criteria. The difference in findings underscores the importance of training quality in the evaluation of any new such instrument.

Our results indicate that women have a greater reflective capacity in comparison with men, which is consistent with the findings of Boenink and colleagues.[21] There was no statistically significant difference noted in reflective capacity of the anesthesiology trainees in self-written vignettes and investigator-written vignettes (2.24 ± 0.74 vs 2.12 ± 0.71, respectively). Of note, scores of the Presence component were significantly greater in self-written vignettes when compared with investigator-written vignettes (2.7 ± 1 vs 2.0 ± 0.9, respectively; $P < .05$). Our finding that scores on the "Presence" domain of REFLECT in self-written compared with investigator-written vignettes could imply that the cognitive challenge created through this task creates a deeper immersion in recalling events that might limit the ability to analyze these events effectively.

This is a pilot study that has several limitations: (1) a relatively small number of participants based at a single center and relatively low participation rate (29/61), which creates a risk of (self) selection bias. (2) We used only the REFLECT rubric to score the reflective capacity and did not compare it with other reflective assessment tools.[28–32] (3) The REFLECT rubric was validated initially for medical students and not for postgraduate clinicians.[5] (4) We made several modifications to the original vignette design developed and validated by Wald and colleagues[5] (ie, added self-written vignettes, altered questions to suit anesthesia-resident context, added vignette writing guide). (5) The external validity of the findings presented here may be limited by differences in the medical education/training systems across jurisdictions. For instance, an Irish trainee might enter a medical training program at a younger age than is usual in the United States. (6) The raters were anesthetic trainees in the same institution in which the study was performed; this might result in identification of the raters to a clinical situation that participants were involved in and could result in priming bias; however, our raters reported that they never recognized a personal vignette among the 28 vignettes that they rated. In future work, if raters are from the same program as subjects, they should be instructed not to rate any personal vignette that they recognize and that identifies the subject to them. Perhaps the greatest limitation is that there is no gold standard against which to determine external validity of the REFLECT methodology we used.

Future research is needed to determine the reflective capacity characteristics among various cohorts of medical trainees and practitioners, some of which we are currently undertaking. Additional work is required to explore various techniques to improve reflective capacity of health

continued on next page

continued from previous page

care professionals; for example, delivering targeted educational programs using real-life scenarios and high-fidelity simulation sessions, and providing positive environment that could trigger the reflective process through regular feedback and questions from supervisors and peers to learners.

Conclusion

Among anesthesiology trainees at an Irish tertiary referral teaching hospital, overall reflective capacity is characterized by introspection or thoughtful action with little or no evidence of critical reflection. Reflective capacity of the anesthesiology trainees appears to decrease as years of medical training progress; however, our respondents were not sampled over time to fully support this conclusion. Overall, female trainees demonstrate superior reflective capacity compared with their male colleagues. Our findings indicate that specific educational interventions are required within training programs to improve the reflective capacity of trainees. More work is required to determine how representative our findings are of trainees' reflective capacity in other settings.

Acknowledgments

Authors' contributions

Study design: GS, ADG

Data collection: HM, A'OL, A'OM

Data analysis: ADG

Data interpretation: HM, ADG, GS, JBC, RHB

Drafting of paper: HM, ADG, JBC, RHB, GS

Critical revision of paper: ADG, JBC, RHB, GS, A'OL, A'OM

Final approval of paper: all authors.

References

1. Sandars J. The use of reflection in medical education: AMEE Guide No. 44. *Med Teach*. 2009;31(8):685-95.
2. Murdoch-Eaton D, Sandars J. Reflection: moving from a mandatory ritual to meaningful professional development. *Arch Dis Child*. 2014;99(3):279-83.
3. Wald HS, Reis SP. Beyond the margins: reflective writing and development of reflective capacity in medical education. *J Gen Intern Med*. 2010;25(7):746-9.
4. Dewey J. *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. Boston: D.C. Heath and Company; 1933.
5. Wald HS, Reis SP, Anthony D, et al. Fostering and evaluating reflective capacity in medical education: developing the REFLECT rubric for assessing reflective writing. *Acad Med*. 2011;87(1):41-50.
6. Stern DT, Papadakis M. The developing physician — becoming a professional. *N Engl J Med*. 2006;355(17):1794-9.
7. Mamede S, Schmidt HG. Reflection in medical diagnosis: a literature review. *Heal Prof Educ*. 2017;3(1):15-25.
8. Shapiro J, Kasman D, Shafer A. Words and wards: a model of reflective writing and its uses in medical education. *J Med Humanit*. 2006;27(4):231-44.
9. Costa Filho GB, Moura AS, Brandão PR, et al. Effects of deliberate reflection on diagnostic accuracy, confidence and diagnostic calibration in dermatology. *Perspect Med Educ*. 2019;8(4):230-6.
10. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. *Am J Med*. 2008;121(5 Suppl):S2-23.
11. Accreditation Council for Graduate Medical Education 2020. Common Program Requirements (Residency). https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRResidency2020.pdf. Accessed May 20, 2021.
12. General Medical Council 2018. Outcomes for Graduates. https://www.gmc-uk.org/media/documents/dc11326-outcomes-for-graduates-2018_pdf-75040796.pdf. Accessed May 20, 2021.
13. DunnGalvin A, Cooper JB, Shorten G, Blum RH. Applied reflective practice in medicine and anesthesiology. *Br J Anaesth*. 2019;122(5):536-41.
14. Kolb AY, Kolb DA. Learning styles and learning spaces: enhancing experiential learning in higher education. *Acad Manag Learn Educ*. 2005;4(2):193-212.
15. Mann K, Gordon ÆJ, Macleod ÆA, et al. Reflection and reflective practice in health professions education : a systematic review. *Adv Health Sci Educ Theory Pract*. 2009;14(4):595-621.
16. Schon DA, DeSanctis V. The reflective practitioner: how professionals think in action. *The Journal of Continuing Higher Education*. 1986;34(3):29-30.
17. Boud D, Keogh R, Walker D. *Reflection: Turning Experience Into Learning*. London: Routledge; 1985.
18. Moon JA. *Reflection in Learning and Professional Development: Theory and Practice*. London: Routledge; 1999.
19. Mezirow J. Transformative dimensions of adult learning. *Adult Educ Q*. 1992;42(3):195-7.
20. Chalmers P, Dunngalvin A, Shorten G. Reflective ability and moral reasoning in final year medical students: a semi-qualitative cohort study. *Med Teach*. 2011;33(5):e281-9.
21. Boenink AD, Oderwald AK, de Jonge P, et al. Assessing student reflection in medical practice. The development of an observer-rated instrument: reliability, validity and initial experiences. *Med Educ*. 2004;38(4):368-77.
22. Blum RH, Boulet JR, Cooper JB, Muret-Wagstaff SL. Simulation-based assessment to identify critical gaps in safe anesthesia resident performance. *Anesthesiology*. 2014;120(1):129-41.
23. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. In: Wendt D, Vlek C, eds. *Utility, Probability, and Human Decision Making. Theory and Decision Library (An International Series in the Philosophy and Methodology of the Social and Behavioral Sciences), vol 11*. Dordrecht: Springer; 1982:141-62.
24. Ker JS. Developing professional clinical skills for practice— the results of a feasibility study using a reflective approach to intimate examination. *Med Educ*. 2003;37:34-41.
25. Arnson L, Niehaus B, Hill-Sakurai L, et al. Comparison of of two methods of teaching reflectiveability in Year 3 medical students. *Med Educ*. 2012;46:807-14.
26. Reis SP, Wald HS, Munroe AD, Borkan JM. Begin the BEGAN – a framework for enhancing educational impact of faculty feedback to students' reflective writing. *Patient Educ Couns*. 2010;80:253-9.
27. Grierson L, Winemaker S, Taniguchi A, et al. The reliability characteristics of the REFLECT rubric for assessing reflective capacity through expressive writing assignments: A replication study. *Perspect Med Educ*. 2020;9(5):281-5.
28. Plack MM, Driscoll M, Blissett S, et al. A method for assessing reflective journal writing. *J Allied Health*. 2005;34(4):199-208.
29. Aukes LC, Geertsma J, Cohen-Schotanus J, et al. The development of a scale to measure personal reflection in medical practice and education. *Med Teach*. 2007;29(2-3):177-82.
30. Kember D, McKay J, Sinclair K, Wong FKY. A four-category scheme for coding and assessing the level of reflection in written work. *Assessment & Evaluation in Higher Education*. 2008;33(4):369-79.
31. O'Sullivan P, Aronson L, Chittenden E, et al. Reflective ability rubric and user guide. *MedEdPORTAL*. 2010;6(1):mep_2374-8265.8133.
32. McNeill H, Brown JM, Shaw NJ. First year specialist trainees' engagement with reflective practice in the e-portfolio. *Adv Health Sci Educ Theory Pract*. 2010;15(4):547-58.

continued on next page

continued from previous page

Hassan M. Ahmed *is Lecturer in the Department of Anesthesia and Intensive Care, Faculty of Medicine, Cairo University, Cairo, Egypt, and Consultant in the Department of Anesthesia and Intensive Care, Leeds Teaching Hospitals, Leeds, United Kingdom.* ***Audrey Dunn Galvin*** *is Programme Director Early Years and Childhood Studies in the School of Applied Psychology, Cork University Hospital, University College, Cork, Ireland; and Visiting Professor in the Department of Paediatrics and Child Infectious Diseases, I.M. Sechenov First Moscow State Medical University, Moscow, Russia. The following authors are in the Department of Anesthesia and Intensive Care, Cork University Hospital, Cork, Ireland:* ***Aoife O'Loughlin*** *is Specialist Registrar,* ***Aisling O'Meachair*** *is Senior House Officer, and* ***George Shorten*** *is Consultant.* ***Jeffrey B. Cooper*** *is Professor in the Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, MA. The following authors are at the Harvard Medical School, Boston, MA:* ***Jeffrey B. Cooper*** *is Professor of Anesthesia and* ***Richard H. Blum*** *is Associate Professor of Anesthesia. The following authors are at the Center for Medical Simulation, Boston, MA:* ***Jeffrey B. Cooper*** *is Executive Director Emeritus and Senior Fellow and* ***Richard H. Blum*** *is Senior Faculty.* ***Richard H. Blum*** *is Senior Associate in Perioperative Anesthesia in the Department of Anesthesiology, Critical Care, and Pain Medicine, Boston Children's Hospital, Boston, MA.* ***George Shorten*** *is Professor of Anesthesia and Intensive Care Medicine at the School of Medicine, University College of Cork, Cork, Ireland; and Funded Investigator, at the Insight II SFI Research Centre, Cork, Ireland.*

Corresponding author: *Hassan M. Ahmed, MD, Department of Anesthesia, Leeds General Infirmary Hospital, Great George Street, Leeds, LS1 3EX.*

E-mail address: *Hassan M. Ahmed: hassan.ahmed7@nhs.net*

Conflicts of Interest: *The work was done in the Department of Anesthesia and Intensive Care, Cork University Hospital, Cork, Ireland. The authors have no conflicts of interest to report.*

Financial Disclosures: *This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Preliminary results were presented in American Society of Anesthesiologists Annual meeting 2020.*

Abstract

Background: Reflective practice is associated with improved accuracy of medical diagnosis and superior performance in complex situations. Systematic observation of trainees' reflective capacities constitutes a basis for an effective support of reflective practice within the training paradigm. We set out to examine the reflective capacity among anesthesiology trainees in a tertiary referral hospital.

Methods: We invited 61 anesthesiology trainees in Cork University Hospitals, Ireland, to participate. Each trainee was invited to respond to 2 investigator-written vignettes prepared by the investigators and suitable for evaluation using the Reflection Evaluation for Learners' Enhanced Competencies Tool (REFLECT) and to produce and then respond to a written vignette based on their own experience. All responses were assessed by 2 independent assessors who had undergone training in the application of the REFLECT rubric, which gives quantifiable scores. Interrater reliability was assessed by weighted kappa coefficient. Association between years of training in medicine and level of reflective capacity was examined using correlation and multiple regression analyses, controlling for age.

Results: Twenty-nine trainees agreed to participate, the overall REFLECT Level was 2.16 (SD 0.7), corresponding to "thoughtful action," indicating low to moderate reflective ability. Cronbach's alpha for the 5 items of the REFLECT scale was excellent (r = 0.92). Weighted kappa was very satisfactory (k = 0.81). A strong association was demonstrated between years in medicine and scores on REFLECT, controlling for age of participant (F = −2.57, Beta coefficient = −0.30). Respondents with less experience had greater mean REFLECT scores than respondents with more experience (F = 5.5, P = .02; post hoc mean difference = 0.7, P = .03 for ≤32 months vs ≥99 months). There was a significant effect for gender (t = −4.3, P = .001), with women's responses receiving greater REFLECT scores than men's responses (mean difference = 0.67, P = .001).

Conclusions: Overall, participants demonstrated low to moderate reflective capacity, as assessed by the REFLECT rubric. Reflective capacity of the anesthesiology trainees appears to decrease as years of medical training progress. However, our respondents were not sampled over time to fully support this conclusion. Further research is needed on the psychometric properties of the REFLECT rubric and the generalizability of our findings.

Keywords: Anesthesia trainees, medical experience, reflective capacity, REFLECT, tertiary Irish hospital

continued from previous page

Tables

Table 1. *Mean Scores for Each of the 5 Criteria, for Investigator-Written and Self-Written Vignettes*

Criterion	Writing Spectrum	Presence	Description of Conflict/ Dilemma	Attending to Emotions	Analysis & Meaning Making
Investigator-written	2.2 (0.9)	2.0 (0.9)	2.4 (0.8)	1.5 (0.7)	2.4 (0.7)
Self-written	2.2 (0.9)	2.7 (1.0)[a]	2.4 (0.7)	1.7 (0.7)	2.3 (0.8)

Values are mean (SD).
[a] Denotes statistical significance; $P < .05$.

Table 2. *Multivariate Regression Predicting Participant Total Score on REFLECT Rubric*

Model	B	SE	Beta	t	Significance
(Constant)	2.683	0.174	-	15.386	.0001
Years of training in medicine	−0.006	0.002	−0.30	−2.62	.010
Age	−0.003	0.003	−0.132	−1.145	.256
Self-written vs investigator-written	0.152	0.154	0.100	0.985	.327

Dependent variable: Total score on REFLECT rubric.

Table 3. *Mean Scores on REFLECT Rubric for Duration of Training in Medicine, Expressed as Months Quartiles*

Duration of Training in Medicine (mo) Quartiles	Mean (SD)
0-32	2.3 (0.8)
33-65	2.4 (0.7)
66-98	2.2 (0.7)
99+	1.6 (0.5)

continued on next page

continued from previous page

Appendices

Appendix 1

REFLECT Vignette – requesting senior assistance

Read the vignette that describes an event in the OR. Imagine that it is your case and you are really there. Where we haven't included details, please feel free to imagine them and incorporate them in your take on the scenario (e.g., specific conversations or wordings, tone of voice, prior relationships, etc). Then take a few moments to reflect on the scenario, and to say what you actually thought and felt.

You are rostered in a day case surgery theatre. Your consultant has told you that he is responsible for two theatres today and that the first case in the other list is a complex case, so he plans to concentrate on that case initially. He has made clear that if you have any queries or concerns, you should call him immediately. This is the first occasion on which you are responsible for a theatre list without senior supervision present.

The first case on the list is maxillofacial surgery (removal of a wire from the palate). When you review the patient, you assessed his upper airway as being difficult to manage. The medical record contained accounts of many previous general anaesthetic, with no reference to difficult airway management.

You elect to proceed with iv induction of anaesthesia without notifying your consultant of your concerns. When you attempt bag mask ventilation, you encounter high airway resistance, positive pressure ventilation is very difficult. The patient arterial saturation declined rapidly, and he becomes cyanosed. The head nurse sends for a senior anaesthetic assistance urgently. When a consultant arrives, he takes over management of the airway and after substantial difficulty manages to intubate the patient's trachea.

The surgery proceeded without further incident. Postoperatively, you review the patient and identify no neurological deficit.

You inform your consultant about what happened, and he expresses frustration with your management of the case. You enquire about your responsibilities regarding documenting the events and explaining them to the patient.

What happened?

I ignored my own concerns and gut feeling about a case and placed the patient in danger due to my own reluctance to bother the consultant anaesthetist who I knew was in a difficult case. This may have partly been due to the lack of mention of difficult airway management in the past although that could well have been poor documentation or prior to whatever maxillofacial issue this patient was having. I then failed to ventilate the patient, causing them to become hypoxic. It is not clear if I followed the difficult airway algorithm and tried a supraglottic airway. It is also not clear if I called for senior assistance or if the nurse saw I was struggling and took it upon herself- that would have been a second error on my part to not call for help. Following this, I did do the right thing by following up the patient, coming clean to the consultant, documenting and openly disclosing to the patient.

What are your thoughts and feelings about why this event happened?

My initial thought was that my handling of the case was due to a reluctance to bother the consultant although actually on second reading I realise that was not explicitly said and that may be just my take on it because that would be in keeping with how I might find myself in a situation like that. Other possible reasons why a person at my level would proceed alone with this case on spite of their concerns could be either that they underestimated the risk to the patient, that they felt themselves capable of managing the case alone or that they didn't want to "blow" their first chance managing a case alone. However the first scenario is the one I could most likely see myself in.

What could you have done differently?

I should have checked in with my consultant and voiced my concerns prior to starting the case, I could have taken more care with my anaesthetic by pre-oxygenating the patient and using an airway adjunct at an early stage, I could have done a fiberoptic intubation (although at my level that would require the consultant anyway) and I should have called for help earlier, before the patient became cyanosed and before the nurse had to take it upon herself.

continued on next page

continued from previous page

Appendices *continued*

To what extent do you think your written response here <u>is similar</u> to how you'd respond if this were your actual clinical case?

My written response here is [not at all similar =1 *to* very similar = 7] to how I would respond if this were my actual clinical case- 6

Not at all similar						Very similar
1	2	3	4	5	6	<u>7</u>

Please give your opinion as to how useful this vignette was as a means of prompting you to reflect, including suggesting how it might be improved.

This vignette was very useful because there were a number of reasons why a person might chose to proceed with a case they felt might be above their experience level and without the vignette giving any one in particular, I immediately jumped to the motivation that correlated best with my own personality and experiences. However on second reading, I opened out the scenario a bit more and reflected in a more abstract way about all the different reasons something like that could happen. In that way, the vignette made me think more about my own attitudes as well as those of others.

continued on next page

continued from previous page

Appendices *continued*

Appendix 2

Vignette Design Notes and Guidelines

Important Characteristics and Considerations in Vignette Design

The residents/subjects will explicitly be told that this is their case and respond in the first person using "I" and not he/she to hopefully allow them to more deeply engage with the scenario/dilemma.

Vignettes should depict medical knowledge and experience that would be expected of a CA-1 resident as these will be used on all level of residents and possibly fellows and attendings.

Clinical history should be clear and unambiguous to avoid confusion and possible discussion of complex medical issues such as the differential diagnosis and management decisions which is the objective.

Residents should not be primed in terms of the emotions he/she are or should be feeling, e.g. I felt really bad after…

Vignettes should not address issues related to moral or ethical issues, e.g. not telling the truth to avoid potential punishment or other negative consequences.

Examples for Essential Elements for REFLECT Vignette Design

Essential elements	
Background/Setting	Late arrival to work from unexpected alarm clock failure.
Procedure/Anesthesia Plan	Spine fusion under GA with large bore IV access and A-line.
Main Proponent Actions/Behaviors	Stress from being late led to rush to set up OR and preparation for case.
Challenge/Conflict/Dilemma/Confrontation	Patient with severe clinical deterioration/arrest from venous air embolism from IV line not properly deaerated. Attending noted air and attempted to stop infusion.
Consequences	Case cancellation with concern for possible neurologic or other complications. Unanticipated admission to ICU for further workup and close monitoring.
Follow up (if relevant)	Staff anesthesiologist very upset about medical error, you were relieved from clinical duties without appropriate feedback and were required to speak with the family with attending.

Scenarios: what to reflect on…...

Things that go wrong. These situations 'stay in your head' and force us to think about whether we could have done anything differently.

For example:

- Peri-operation complications.
- Communication problems with patients and families e.g. anxious/aggressive patient or family member.
- Communication problems with superiors and/or disagreeing with decisions made by superiors.
- Personal medical error or mistake in judgement of the practitioner or someone they observe not acting correctly/making a mistake in their professional practice.
- A dissatisfied patient
- Failed procedure

continued on next page

continued from previous page

Appendices *continued*

However, reflecting on things that went well can also be useful, particularly in terms of learning.

For example:

- A well-managed crisis e.g. cardiac arrest.
- A patient thank you letter.
- Coping well with anxious/aggressive patient or family member.
- A difficult but well performed case.

Claims: Procedures and contributory factors

Dental Damage

Spinal and epidural anesthesia

Interscalene brachial plexus blocks were the most common site, followed by femoral nerve blocks. Common themes in these claims included:

- Failure to obtain informed consent with respect to risks and benefits, and to clearly document that such a discussion had occurred.
- Failure to use either a nerve stimulator or ultrasound guidance during the performance of the block.
- In cases where a nerve stimulator was used, failure to record sufficient detail, such as loss of twitch on initial test injection of local anesthetic.
- Failure to record that intravascular injection was unlikely as there was no aspiration of blood. And similarly, there being no record of resistance on injection to demonstrate that intraneural injection was unlikely.

Nerve damage due to malpositioning

Awareness during general anesthesia

Never events

Complaints: Common themes

In relation to anesthesia, complaints included:

- Poor manner and attitude, including rudeness and inappropriate remarks made to patients during preoperative assessment and in the anesthetic room. Complaints by colleagues relating to the manner and attitude displayed towards other staff members were also made.
- Inadequate postoperative analgesia.
- Ineffective spinal or epidural anesthesia and complications of postdural puncture headache, hematoma, infection and neural damage.
- Painful or repeated attempts at cannulation.
- Failure to provide adequate sedation.
- Cancellation of surgery.
- Aspiration pneumonia following anesthesia.
- Failure to obtain informed consent for a procedure, such as a nerve block.

For critical care medicine, complaints often related to discussions with relatives, including allegations of poor communication with respect to treatment withdrawal. Complaints specifically relating to pain medicine included those relating to a lack of empathy shown during consultations and inappropriate comments by the anesthesiologist. Some complaints related to delays in providing treatment, misdiagnosis of the source of pain, and continuation of pain after treatment.

Regulatory (General Medical Council) and disciplinary cases: Common themes

- Performance concerns, such as technical skills including poor airway management, clinical judgement and communication.

continued on next page

continued from previous page

Appendices *continued*

- Probity issues, for example:
  - » not reviewing patients prior to a theatre list
  - » leaving anesthetized patients unattended
  - » prescribing for family members, colleagues or self
  - » being unavailable when on call or refusal to attend a patient when requested
  - » conducting private practice in NHS time
  - » fraudulent information included on job application forms or appraisals.
- Health issues, including alcohol or drug misuse, and the theft of drugs from hospital premises.
- Inappropriate personal behavior or misconduct, such as inappropriate comments made to patients, bullying and harassment of colleagues.
- Inappropriate delegation, like leaving inexperienced junior trainees alone.

continued on next page

continued from previous page

Appendices *continued*

Appendix 3

Scores of participants' investigator-written vignettes

Participants	First vignette (mean rated score)	Second vignette (mean rated score)	Difference
1	2.4	2.1	0.3
2	1.8	1.7	0.1
3	2.6	2.6	0.0
4	2.7	2.3	0.4
5	1.5	1.9	0.4
6*	1.6	3.6	1.0
7	1.5	1.7	0.2
8	3.3	3.8	0.5
9	2.3	1.8	0.5
10	1.4	1.7	0.3
11	3.3	3.8	0.5
12	1.1	1.4	0.3
13	1.8	1.7	0.1
14	1.4	1.4	0.0
15	1.2	1.4	0.2
16	2.7	2.7	0.0
17	1.7	1.7	0.0
18	1.5	1.7	0.2
19	1.4	1.9	0.5
20	1.0	1.3	0.2
21	1.8	1.8	0.0
22	1.5	1.3	0.2
23*	2.6	3.3	0.6
24	1.9	2.1	0.2
25	2.0	2.3	0.2
26	2.4	2.6	0.2
27	3.3	3.6	0.3
28	2.5	2.8	0.3
29	2.9	2.9	0.0

* Denotes lack of consistency between 2 responses