



# Ten caveats of learning analytics in health professions education: A consumer's perspective

Olle ten Cate, Suzan Dahdal, Thomas Lambert, Florian Neubauer, Anina Pless, Philippe Fabian Pohlmann, Harold van Rijen & Corinne Gurtner

To cite this article: Olle ten Cate, Suzan Dahdal, Thomas Lambert, Florian Neubauer, Anina Pless, Philippe Fabian Pohlmann, Harold van Rijen & Corinne Gurtner (2020) Ten caveats of learning analytics in health professions education: A consumer's perspective, Medical Teacher, 42:6, 673-678, DOI: [10.1080/0142159X.2020.1733505](https://doi.org/10.1080/0142159X.2020.1733505)

To link to this article: <https://doi.org/10.1080/0142159X.2020.1733505>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 09 Mar 2020.



Submit your article to this journal [↗](#)



Article views: 1055



View related articles [↗](#)



View Crossmark data [↗](#)

## Ten caveats of learning analytics in health professions education: A consumer's perspective

Olle ten Cate<sup>a</sup> , Suzan Dahdal<sup>b</sup>, Thomas Lambert<sup>c</sup>, Florian Neubauer<sup>d</sup>, Anina Pless<sup>e</sup>, Philippe Fabian Pohlmann<sup>f</sup>, Harold van Rijen<sup>a</sup> and Corinne Gurtner<sup>g</sup>

<sup>a</sup>Center for Research and Development of Education, University Medical Center Utrecht, Utrecht, The Netherlands; <sup>b</sup>University Hospital Bern, Bern, Switzerland; <sup>c</sup>Kepler University Hospital Linz, Johannes Kepler University Linz, Linz, Austria; <sup>d</sup>Institute for Medical Education, University of Bern, Bern, Switzerland; <sup>e</sup>Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland; <sup>f</sup>Faculty of Medicine, University of Freiburg, Breisgau, Germany; <sup>g</sup>Institute of Animal Pathology, Vetsuisse Faculty Bern, University of Bern, Bern, Switzerland

### ABSTRACT

A group of 22 medical educators from different European countries, gathered in a meeting in Utrecht in July 2019, discussed the topic of learning analytics (LA) in an open conversation and addressed its definition, its purposes and potential risks for learners and teachers. LA was seen as a significant advance with important potential to improve education, but the group felt that potential drawbacks of using LA may yet be under-exposed in the literature. After transcription and interpretation of the discussion's conclusions, a document was drafted and fed back to the group in two rounds to arrive at a series of 10 caveats educators should be aware of when developing and using LA, including too much standardized learning, with undue consequences of over-efficiency and pressure on learners and teachers, and a decrease of the variety of 'valid' learning resources. Learning analytics may misalign with eventual clinical performance and can run the risk of privacy breaches and inescapability of documented failures. These consequences may not happen, but the authors, on behalf of the full group of educators, felt it worth to signal these caveats from a consumers' perspective.

### KEYWORDS

Portfolio; computer-based; information handling; trends

Recently, JAMA published a study showing how 50 participants, divided in medical students, junior residents, senior residents, fellows, and neurosurgeons, were 90% correctly classified by a machine learning algorithm after all had exercised 250 simulated neurosurgical operations, each collecting 270 metrics (totaling over 3 million data points). The authors suggest that, while traditional simulation requires learner feedback by skilled instructors, artificial intelligence may decrease the need for human evaluators (Winkler-Schwartz et al. 2019). In 2016, Warm et al. used 364,728 entrustment data of 189 residents in the Cincinnati Internal Medicine (IM) program, collected within 36 months, and created clear graphics of development and trends showing systematic differences between faculty, peers and allied health professionals' assessments (Warm et al. 2016). Another Cincinnati study showed that for every point increase in USMLE Step 2 Clinical Knowledge scores of IM residents, the odds of passing the American Board of Internal Medicine certification increased by 6.9% (Sharma et al. 2019). Still another Cincinnati study showed how 1511 patients' electronic records enabled predicting with great accuracy whether or not an intern was the primary caregiver of the patient (Schumacher et al. 2019). Saqr et al. predicted student's final grades with 63.5% accuracy and predicted at risk students to some degree based on their earlier online activity traced among 133 students (Saqr et al. 2017).

### Practice points

Learning analytics (LA) may bring many benefits to health professions education, but educators, programs, and schools should be aware of potential caveats and risks, among which:

- LA should not unduly suggest 'optimal' learning strategies and disregard students' individual learning pathways.
- LA may recommend efficient education but efficiency should not translate to increased stress on learners.
- Schools should be aware of privacy and protection of collected data and guard against the use of data for unintended purposes.
- LA data in e-portfolios may feel as unescapable documentation of all successes and failures and a 'right to be forgotten' rule may be needed.

These examples of *learning analytics* (LA) show how big data collected on learners and professionals serve to inform health professions educators and others about learners and their progress (Chan et al. 2018). Learning analytics is in an early phase of development and may bring many

**CONTACT** Olle ten Cate  [t.j.tencate@umcutrecht.nl](mailto:t.j.tencate@umcutrecht.nl)  Center for Research and Development of Education, University Medical Center Utrecht, P.O. Box # 85500, 3508 GA Utrecht, The Netherlands

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

advantages to education and individual students. While benefits may be undisputed, there may also be some downsides of the use of LA. We attempted to identify some of these caveats.

## Background, setting, and methods

In July 2019, a group of 22 educators, predominantly participants of the Master of Medical Education Course of the University of Bern, Switzerland, convened for a summer retreat at University Medical Center Utrecht, The Netherlands, to discuss topics of interest for health professions education. The participants included medical specialists, a primary health care scientist, a veterinary pathologist, a nurse anesthetist educator, two residents, and three senior educational scholars. One of the topics of interest was LA. All participants engaged in its discussion, which alternated between groups of three and plenary discussions. The richness of the discussion led to the decision to consolidate its results in an article, with a focus on caveats of LA, as viewed from the perspective of an educator using and adjusting, but not developing LA, i.e. a 'consumers' perspective'.

The discussion evolved around three questions: What is learning analytics? What are its purposes? and What are potential risks of learning analytics?

The group arrived at definition features that reflect the literature (*'The measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs'*) (Siemens 2012). However, the group also felt that, in addition to necessary concepts of learners, measurement, data and analysis, the definition should include *learner progress, learning process information, and behavioral aspects*, such as time spent on learning.

As a purpose, besides understanding and optimizing learning, the group agreed that LA should optimize teaching and education, and its research and development. Furthermore, LA in health professions education should not only optimize learning but eventually optimize clinical practice.

Finally, the core discussion focused on caveats, i.e. unintended adverse consequences or potential risks of using LA. Data are easily collected; the hard part is interpretation and translations to educational policy. As Gašević et al. contend, LA tools should be based on robust theoretical models of human behavior to be effective (Gašević et al. 2015).

These caveats were listed on flip-charts during the discussion, transcribed, elaborated, and fed back to the participants after the meeting. In a subsequent second and a third round within weeks after the meeting, comments and additions were provided, resulting in a series of caveats the group found worthy of sharing. Next, all who volunteered to be co-authors on the full paper were tasked to substantiate two or more of the caveats identified. These were edited and consolidated in 10 caveats by the primary author.

## Caveats of learning analytics from a consumer's perspective

### Decreased autonomy through standardized learning

Designed as a tool to facilitate individual learning by collecting large amounts of data and providing individualized

learning recommendations, LA may give insight into what optimal learning is. However, this insight may give rise to pressure among learners to meet that ideal, which can cause stress and could obliterate alternative learning paths (Gašević et al. 2015). One example is that of e-portfolios: they provide an opportunity for learners to receive individualized feedback by teachers based on data entered into the electronic data management system (van der Schaaf et al. 2017). On the other hand, learners could feel pressured to fill in a certain amount of data points for fear of being compared to peers and ranked. Consequently, they might change their learning style, even if it has proven to be effective for them so far, just because it does not reflect how (most) others learn. The resulting self-chosen decrease of autonomy can be considered an unintended side effect of LA (Buckley et al. 2009; van der Schaaf et al. 2017). Another example is the commercial platforms such as AMBOSS (for students; <https://www.amboss.com/us>) or the Medical Knowledge Self-Assessment Program (for clinicians; <https://mksap18.acponline.org>) offer individualized recommendations to learning specific topics based on answers to MCQ questions from past exams (Bientzle et al. 2019). These concise, specialized recommendations entail the risk that learners lose sight of the bigger picture and their sense of ownership of learning. A decrease in autonomy of learning may have adverse effects on motivation and academic performance (Artino et al. 2010; Kusrkar et al. 2011).

### A decrease in the variety of learning resources

While resources for learning keep increasing and students can choose from a much larger variety to match their preferences than in the past, LA may strongly direct students toward the most efficient or effective sources. A caveat is that 'the most efficient or effective' for the group as a whole (or the 'average' student) may conceal individual differences in style and preference (Newble and Clarke 1986; Newble and Entwistle 1986). To meet the demands of their assessment students often seek strategic learning approaches (Taylor and Hamdy 2013). To game the system students may quickly seek and exchange resources with a 'proven' effect, thus narrowing down the breadth of learning resources and impoverishing its variety. The seductive 'power of proof' from LA can mislead individual students who would be better served with a non-mainstream approach to learning.

### Unintended efficiency consequences

Once LA has defined the most effective and efficient learning behavior, regulators may be tempted to regard this behavior as the norm, reflecting a shorter than average learning trajectory. There may eventually be a push toward a decrease in time and money allotted to education (and thus to learning). Medical education is long and expensive, for the individual as well as for the community, so any increase in efficiency and decrease in length will be welcomed. Data analytics in almost any branch drive to increase efficiency. The question is whether learning can really be sped up (ten Cate et al. 2018). Pressure on students to increase efficiency, and to match or exceed the average student, can cause stress, and increased competition among students. Student wellbeing, already at stake (Mata et al. 2015; Rotenstein et al. 2016),

should not be further jeopardized. Wellbeing of health professionals, one of the 'quadruple aims' for improvement of health care (Bodenheimer and Sinsky 2014), should also regard students.

### **Increased time commitment from teachers**

LA requires massive amounts of data to be collected, thus making a large number of evaluations necessary. In the current competency-based medical education (CBME) programs much qualitative and quantitative data is already being produced in order to assess trainee competencies and to make decisions about their progression within the curriculum. As LA requires an increased number of data points; however, the purpose of evaluations may shift from a direct benefit for the learner and teacher toward the more managerial goal of amassing data in order to feed a database (Chan et al. 2018). The process of this data collection may burden teachers. Creating the data becomes a task in itself, which can take time and energy resources away from teaching and may lead to evaluation fatigue (Barrett et al. 2018). Although long-term benefits of LA may be present, the lack of an immediate visible value for one's own teaching could lead to frustration and loss of enthusiasm and actually impede the quality of feedback to students. To serve decision making by teachers, the data need to be comprehensible; the mere volume of it may not serve that purpose (Chan et al. 2018).

### **A decrease of 'valid' information sources to assess learner progress**

Paradoxically, while LA uses big data, preferably from many sources across large numbers of subjects, there may be a risk that learner assessment will focus on less rather than more of these data sources, because the analysis may attempt to optimize the sources that predict outcomes of learning best, at the cost of other sources. Of the 270 metrics measured by Winkler-Schwarz et al., machine learning algorithms chose 6 to 9 to include in performance evaluation (Winkler-Schwartz et al. 2019). However, inferences based on groups may not always be optimal for all individuals, just like evidence-based patient management protocols may not be optimal for all individual patients.

E-Portfolios that document learner behavior in the clinical workplace may suggest building an accurate picture of the learner, because of the large number of data points, but some of these data may not be accurate. 'Garbage-in-garbage-out' maybe a too cynical view, but we should be aware of the possibility of false accuracy of conclusions drawn from aggregated portfolio data. While aggregated portfolio data are not always well interpreted (Oudkerk Pool et al. 2018), the data itself may lack validity. Entries may have been selective if the student can choose observation moments and observers, and if clinical staff who rate are careless, hasty, unduly affected by mood and personality or cognitively overloaded (Paravattil and Wilby 2019). A good clerkship director knows the 'doves and hawks' among the staff who assess and should interpret portfolio information, rather than accept all information as equally valid. Without carefully weighing portfolio data, such corrections may not happen.

### **Learning analytics may misalign with eventual clinical performance**

Maslow's old saying 'if your tool is a hammer, everything looks like a nail' might hold true, to a certain extent, for LA. Learning analytics data such as online behavior and scores on written tests, which are easy to collect, may poorly reflect the ultimate purpose of training: performance in a professional environment. LA is focused on optimizing *learning* in educational programs. If such learning is defined as classroom learning with written examinations as the most desirable outcome measure, predictors of clinical performance may be missed. The optimal *learner* may not be the optimal medical *practitioner*. Data at lower levels of Miller's Pyramid ('Knows' and 'Knows how') (Miller 1990) are readily available in many faculty databases but do not reflect clinical performance very well. Likewise, at the other end of the educational continuum, maintenance of certification based on written exams developed by specialty boards has been criticized because of their low usefulness for the actual maintenance of practice quality (Weinberger 2019) and knowledge about the effectiveness of continuing medical education courses is mostly limited to data from the lowest levels of the Kirkpatrick evaluation hierarchy (Tian et al. 2007). Using these 'hammers and nails' for feedback to learners, physicians and educators may distract from important components of clinical practice, simply because appropriate data is not collected. There is a need for more and better metrics that truly focus on clinical performance. Aligning LA with the most important, health care related outcomes of education is still in its infancy (Bakharia et al. 2016).

### **Impact of LA on privacy and data protection**

To optimize recommendations about learning behavior and to predict success, LA must collect as much data about learners as possible. This may include very personal measures, like place and time of the day when learning occurs, duration and number of breaks taken, online keyboard behavior, possibly even physiological parameters. Eye tracking during laparoscopic surgery has now been used to assess surgical skills (Gunawardena et al. 2011). Face recognition and *Human Activity Recognition* can be used to identify extremely personal data of human behavior through various sensors and for various reasons (Kabassi and Alepis 2019; Lu et al. 2017; Ravi et al. 2017). The mere existence of such data may evoke interest to use them for unintended purposes. Rules determining the limits of LA data collection, storage, use, and deletion are needed to protect the privacy of learners. Ownership and sovereignty of data and its use are crucial issues and need to be addressed. Which person or board can decide which data will be analyzed when and with which intention? Can the intended use of data be adjusted after consent was given? Must consent be renewed on a regular basis? Can data collection be a condition for course enrolment or can learners refuse to give consent and what consequences will that have? Without clarification of these questions, a fear of data misuse and the development of new but possibly unwanted power relationships remains. It seems pertinent that students have a say in these regulations (West et al. 2020).

### **The inescapability of documented failures**

While all students learn through errors and mistakes and deserve to be corrected with constructive oral feedback by skilled teachers, the increased electronic documentation in workplace-based assessment and feedback may give undue weight to attempts, small mistakes, and errors. Many students feel documented feedback to be more summative than was intended (Bok et al. 2013; Heeneman et al. 2015). They cannot be blamed for that, as the flipside of documenting all of the behavior is that all of their observed failures will remain forever in the information cloud. The mere idea that mistakes in the past can be held against students in the future, true or not, can have detrimental effects on their openness and their fun of learning (Dyer and Cohen 2018). In a recent report on 'online self-guarding' for higher education, Bond and Phippen recommend a rule concerning the 'right to be forgotten' (Bond and Phippen 2019) as is the case for search engines in many jurisdictions. This is particularly relevant when massive data collection for student portfolios and monitoring of progress in programmatic assessment (Van Der Vleuten et al. 2015) becomes dominant. The option to 'clean' a student's record in order to grant a new start may be incorporated in measures after successful remediation.

### **Hawthorne and time-lag effects of LA**

In a complex adaptive system such as education, the information from LA measurement fed back to learners may change their behavior. Consequently, the predictive power of LA inferences may decrease, comparable to the Hawthorne effect in psychological studies, when individuals modify behavior in response to awareness of being observed. While Hawthorne effects have been discussed in the context of direct observations of clinical tasks (Kogan et al. 2017; Paradis and Sutkin 2017), the reaction of learners to longitudinal behavior recordings is less clear. However, it is not hard to imagine how students flagged as 'under-achievers' or 'at risk' (Saqr et al. 2017) based on their documented online behavior and receiving surprise calls from a tutor, change their online behavior to mimic engagement and appropriate self-regulation. Their 'online image' may no longer be valid.

Similarly, the predictive value of LA could fail if the training cohort from which the reference data originate and the investigated cohort deviate significantly. Due to a rapidly emerging technology-based internet society (Wartman 2019), a fast evolution of health care, and due to the emergence of new habits in learning behavior, this time-lag might render any LA-based standards for individualized learning invalid.

### **A gradual shift of the purpose of LA**

Tools developed for one purpose may gradually shift to being used for purposes that were not intended. Monitoring instruments to provide feedback to learners, or evaluation data to improve programs run the risk of becoming used for accreditation, quality comparisons or ranking of individuals or programs, purposes for which the data and analysis were not designed. Rankings, in

particular, carry the risk of methodological or ideological errors and unfairness (Powell and Steelman 2006). There are abundant examples of misuse in the literature (Cousins 2004) and missing data might have severe ramifications (Chan et al. 2018; McConnell et al. 2016). It is not unthinkable that tools designed for workplace-based assessment will become mandatory to obtain license certification. Such use was not the initial purpose of 'assessments for learning' (Rogausch et al. 2012). Educators, educational institutions, and accrediting bodies bear a high responsibility for the correct use of LA instruments and data to avoid misuse.

### **Discussion**

The use of LA for the purpose of improving curricula and empowering students is a laudable cause and many virtues have been highlighted elsewhere (Chan et al. 2018). Technology advancements have their own dynamics and there is no sense in questioning these advancements. The question is not whether we should use such technology and data, but how to optimize this use for meaningful learning and fulfill the needs and protect the rights of learners, educators, and administrators at the same time. The educators gathered in the Utrecht meeting felt that some concerns were worth sharing with a wider audience.

Our exercise has significant limitations. The group consisted of health professions educators with, or completing advanced academic degrees in health professions education, but did not specifically include LA experts. Our report should, therefore, be viewed as an overview of informed consumer concerns. It may be regarded as a starting point for more rigorous future studies with established consensus methods among a more diverse expert panel.

### **Acknowledgements**

The authors wish to acknowledge Lorenz Bärlocher MD, Anne-Kathrin Eickelmann MD, Regula Fankhauser MD, Tamara Guidi Margaris MD, Christoph Hauser MD, Nadine Hollenstein, David Hörburger MD, Henry Madlon MD, Isabelle Pramana MD, Tobias Ries RN, Benny Wohlfarth MD, Manfred Wieser MD, Maya Wolfensperger MD, and Sandra Trachsel PhD. All were present, in addition to the authors, at the July 2019 Utrecht meeting for Bern's Master of Medical Education program and have actively engaged in the discussions about this topic.

### **Disclosure statement**

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

### **Glossary**

**Learning analytics:** The measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs.

### **Notes on contributors**

*Olle ten Cate*, PhD, Senior Scientist at the Center for Research Development of Education, University Medical Center Utrecht, The Netherlands.

**Suzan Dahdal**, MD, Senior Nephrologist, Department of Nephrology and Hypertension, University Hospital Bern, Switzerland.

**Thomas Lambert**, MD, Senior Consultant Physician and Director of the Cardiac and Medical Intensive Care Unit, Department of Cardiology, Kepler University Hospital Linz. Coordinator and teacher at the Medical Faculty, University of Linz, Austria.

**Florian Neubauer**, MD, PhD, Scientist at the Department of Assessment and Evaluation, Institute for Medical Education, University of Bern, Switzerland.

**Anina Pless**, MD, Scientist at the Institute of Primary Healthcare of the University of Bern (BIHAM), Switzerland.

**Philippe Fabian Pohlmann**, MD, resident in the Department of Urology, Scientist in the Division of Urotechnology of the Department of Urology, Medical Center Freiburg i. Br., Germany.

**Harold Van Rijen**, PhD, Professor of Innovation in Biomedical Sciences Education, University Medical Center Utrecht, The Netherlands.

**Corinne Gurtner**, Dr. Med. Vet., Senior Veterinary Pathologist at the Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Switzerland.

## ORCID

Olle ten Cate  <http://orcid.org/0000-0002-6379-8780>

## References

- Artino AR, La Rochelle JS, Durning SJ. 2010. Second-year medical students' motivational beliefs, emotions, and achievement. *Med Educ.* 44(12):1203–1212.
- Bakharia A, Corrin L, de Barba P, Kennedy G, Mulder RA, Williams D, Dawson S, Lockyer L, Gasevic D. 2016. A conceptual framework linking learning design with learning analytics. LAK '16: 6th International Conference on Learning Analytics and Knowledge Edinburgh United Kingdom, 25–29 April. p. 329–338.
- Barrett M, Georgoff P, Matusko N, Leininger L, Reddy RM, Sandhu G, Hughes DT. 2018. The effects of feedback fatigue and sex disparities in medical student feedback assessed using a minute feedback system. *J Surg Educ.* 75(5):1245–1249.
- Bientzle M, Hirsin E, Kimmerle J, Knipfer C, Smeets R, Gaudin R, Holtz P. 2019. Association of online learning behavior and learning outcomes for medical students: large-scale usage data analysis. *J Med Internet Res.* 21(8):1–8.
- Bodenheimer T, Sinsky C. 2014. From triple to quadruple aim: care of the patient requires care of the provider. *Ann Fam Med.* 12(6): 573–576.
- Bok HG, Teunissen PW, Favier RP, Rietbroek NJ, Theyse LF, Brommer H, Haarhuis JC, van Beukelen P, van der Vleuten CP, Jaarsma DA, et al. 2013. Programmatic assessment of competency-based workplace learning: when theory meets practice. *BMC Med Educ.* 13(1):123.
- Bond E, Phippen A. 2019. Higher education online safeguarding self-review tool. Suffolk (UK): Office for Students, University of Suffolk.
- Buckley S, Coleman J, Davison I, Khan KS, 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.
- Chan T, Sebok-Syer S, Thoma B, Wise A, Sherbino J, Pusic M. 2018. Learning analytics in medical education assessment: the past, the present, and the future. *AEM Educ Train.* 2(2):178–187.
- Cousins JB. 2004. Commentary: minimizing evaluation misuse as principled practice. *Am J Eval.* 25(3):391–399.
- Dyer C, Cohen D. 2018. How should doctors use e-portfolios in the wake of the Bawa-Garba case? *BMJ.* 360(February):k572–k573.
- Gašević D, Dawson S, Siemens G. 2015. Let's not forget: learning analytics are about learning. *TechTrends.* 59(1):64–71.
- Gunawardena N, Matscheko M, Anzengruber B, Ferscha A, Schobesberger M, Shamiyeh A, Klugsberger B, Solleder P. 2011. Assessing surgeons' skill level in laparoscopic cholecystectomy using eye metrics. Conference Proceedings ETRA '19: 2019 Symposium on Eye Tracking Research and Applications, Denver, Colorado, June 2019.
- Heeneman S, Oudkerk Pool A, Schuwirth LWT, van der Vleuten CPM, Driessen EW. 2015. The impact of programmatic assessment on student learning: theory versus practice. *Med Educ.* 49(5):487–498.
- Kabassi K, Alepis E. 2019. Learning analytics in distance and mobile learning for designing personalized software. *Mach Learn Paradigms Adv Learn Anal.* 158:185.
- Kogan JR, Hatala R, Hauer KE, Holmboe E. 2017. Guidelines: the do's, don'ts and don't knows of direct observation of clinical skills in medical education. *Perspect Med Educ.* 6(5):286–305.
- Kusurkar RA, Croiset G, ten Cate TJ. 2011. Twelve tips to stimulate intrinsic motivation in students through autonomy-supportive classroom teaching derived from self-determination theory. *Med Teach.* 33(12):978–982.
- Lu Y, Zhang S, Zhang Z, Xiao W, Yu S. 2017. A framework for learning analytics using commodity wearable devices. *Sensors.* 17(6): 1325–1382.
- Mata DA, Ramos MA, Bansal N, Khan R, Guille C, Di Angelantonio E, Sen S. 2015. Prevalence of depression and depressive symptoms among resident physicians. *JAMA.* 314(22):2373.
- McConnell M, Sherbino J, Chan TM. 2016. Mind the gap: the prospects of missing data. *J Grad Med Educ.* 8(5):708–712.
- Miller GE. 1990. The assessment of clinical skills/competence/performance. *Acad Med.* 65(9):S63–S67.
- Newble DI, Clarke RM. 1986. The approaches to learning of students in a traditional and in an innovative problem based medical school. *Med Educ.* 20(4):267–273.
- Newble DI, Entwistle NJ. 1986. Learning styles and learning approaches: implications for medical education. *Med Educ.* 20(4): 162–175.
- Oudkerk Pool A, Govaerts MJB, Jaarsma DADC, Driessen EW. 2018. From aggregation to interpretation: how assessors judge complex data in a competency-based portfolio. *Adv Health Sci Educ.* 23(2): 275–287.
- Paradis E, Sutkin G. 2017. Beyond a good story: from Hawthorne effect to reactivity in health professions education research. *Med Educ.* 51(1):31–39.
- Paravattil B, Wilby KJ. 2019. Optimizing assessors' mental workload in rater-based assessment: a critical narrative review. *Perspect Med Educ.* 8(6):339–345.
- Powell B, Steelman L. 2006. Bewitched, bothered, and bewildering: the use and misuse of state SAT and ACT scores. *Harvard Educ Rev.* 76(4):453–456.
- Ravi D, Wong C, Lo B, Yang GZ. 2017. A deep learning approach to on-node sensor data analytics for mobile or wearable devices. *IEEE J Biomed Health Inform.* 21(1):56–64.
- Rogausch A, Berendonk C, Giger M, Bauer W, Beyeler C. 2012. Purpose and usefulness of workplace-based assessment in daily clinical practice. *Swiss Med Forum.* 12(10):214–217.
- Rotenstein LS, Ramos MA, Torre M, Segal JB, Peluso MJ, Guille C, Sen S, Mata DA. 2016. Prevalence of depression, depressive symptoms, and suicidal ideation among medical students a systematic review and meta-analysis. *JAMA.* 316(21):2214–2236.
- Saqr M, Fors U, Tedre M. 2017. How learning analytics can early predict under-achieving students in a blended medical education course. *Med Teach.* 39(7):757–767.
- Schumacher DJ, Wu DTY, Meganathan K, Li L, Kinnear B, Sall DR, Holmboe E, Carraccio C, van der Vleuten C, Busari J, et al. 2019. A feasibility study to attribute patients to primary interns on inpatient ward teams using electronic health record data. *Acad Med.* 94(9): 1376–1383.
- Sharma A, Schauer DP, Kelleher M, Kinnear B, Sall D, Warm E. 2019. USMLE Step 2 CK: best predictor of multimodal performance in an internal medicine residency. *J Grad Med Educ.* 11(4):412–419.
- Siemens G. 2012. Learning analytics: envisioning a research discipline and a domain of practice. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge; May; Vancouver, British Columbia, Canada. p. 4–8. <http://dl.acm.org/citation.cfm?id=2330605>.
- Taylor DCM, Hamdy H. 2013. Adult learning theories: implications for learning and teaching in medical education: AMEE Guide No. 83. *Med Teach.* 35(11):e1561–e1572.

- ten Cate O, Gruppen LD, Kogan JR, Lingard LA, Teunissen PW. 2018. Time-variable training in medicine: theoretical considerations. *Acad Med.* 93(35):S6–S11.
- Tian J, Atkinson NL, Portnoy B, Gold RS. 2007. A systematic review of evaluation in formal continuing medical education. *J Contin Educ Health Prof.* 27(1):16–27.
- van der Schaaf M, Donkers J, Slof B, Moonen-van Loon J, van Tartwijk J, Driessen E, Badii A, Serban O, ten Cate O. 2017. Improving workplace-based assessment and feedback by an E-portfolio enhanced with learning analytics. *Educ Technol Res Dev.* 65(2): 359–380.
- Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, Govaerts MJB, Heeneman S. 2015. Twelve tips for programmatic assessment. *Med Teach.* 37(7):641–646.
- Warm EJ, Held JD, Hellmann M, Kelleher M, Kinnear B, Lee C, O'Toole JK, Mathis B, Mueller C, Sall D, et al. 2016. Entrusting observable practice activities and milestones over the 36 months of an internal medicine residency. *Acad Med.* 91(10):1398–1405.
- Wartman SA. 2019. The empirical challenge of 21st-century medical education. *Acad Med.* 94(10):1412–1415.
- Weinberger S. 2019. Can maintenance of certification pass the test? *JAMA.* 321(7):641–642.
- West D, Luzeckyj A, Toohey D, Vanderlelie J, Searle B. 2020. Do academics and university administrators really know better? The ethics of positioning student perspectives in learning analytics. *Aust J Educ Technol.* 36(2):60–70.
- Winkler-Schwartz A, Yilmaz R, Mirchi N, Bissonnette V, Ledwos N, Siyar S, Azarnoush H, Karlik B, Del Maestro R. 2019. Machine learning identification of surgical and operative factors associated with surgical expertise in virtual reality simulation. *JAMA Netw Open.* 2(8):e198363.