**ABSTRACT**

Construct: This study presents a tool that can facilitate a conversation about students’ and supervisors’ expectations concerning responsibilities during workplace learning. Background: It is often unclear who is responsible for facilitating learning opportunities in the workplace. In order to increase learning opportunities, it is important that expectations are discussed and alignment is reached between the student’s and supervisor’s expectations. This study collected and interpreted validity evidence for a tool that aims to provoke such a conversation.

Approach: Three types of validity evidence were collected: response process, content, and consequences evidence. Educational leaders, medical teachers, and students of four medical schools were involved. The data collection consisted of cognitive interviews, a modified Delphi approach (with three rounds of inquiry), completed tools, and narrative comments.

Findings: This study showed that the expectations of most students and supervisors were not initially aligned. The conversation, for which the tool aims to be a catalyst, facilitated better alignment of expectations about responsibilities during workplace learning. Moreover, the students’ perceived degree of consensus and satisfaction after the conversation were very high.

Conclusions: This study underlined the relevance and usefulness of a tool that facilitates conversation about expectations regarding responsibilities, potentially enhancing learning opportunities at the workplace.

**KEYWORDS**

Workplace learning; communication; responsibility; expectations

**Introduction**

Authentic workplace-based learning experiences are increasingly included in medical curricula. They offer students the opportunity to apply and refine classroom-acquired competences in clinical practice. Although research shows many benefits of such experiences, the learning outcomes of these experiences strongly depend on the quality of the interaction between student and supervisor.

The understanding reached by student and supervisor regarding who owns responsibility for driving learning appears to be particularly important. Both students and supervisors have various, and often shifting, responsibilities during workplace learning. Moreover, it is often unclear who has responsibility for important aspects, such as:

1. Formulating goals about applying competences
2. Organizing opportunities to practice
3. Linking practice with classroom-based content
4. Taking initiative for feedback
5. Taking initiative for reflection

For each of these five aspects, a range of expectations exists among students and workplace supervisors. Research has revealed that this variety could be perceived as a continuum between, on the one hand, holding students responsible and, on the other hand, placing responsibility primarily on workplace supervisors.

When a student and a supervisor have different expectations about who is responsible for any of those five aspects, it can lead to misunderstandings and consequently learning opportunities can be missed. If a student assumes that it is the supervisor’s responsibility to take initiative for feedback, they might find themselves still waiting for feedback at the end of a workplace-based experience. Meanwhile, the supervisor may expect that the student takes initiative to ask for feedback and therefore may be disappointed.
with the lack of initiative. If these expectations were discussed, this misunderstanding could be avoided, with the student receiving earlier feedback and potentially learning more. Consequently, it can help if the roles and responsibilities of students and supervisors are clarified and agreements are made about responsibilities for each element within the context. Such agreements, along with an understanding around the specific actions expected of everyone involved, can enhance accountability.

It is expected that alignment of a student’s and supervisor’s expectations might increase collaboration between them and, therefore, enhance the student’s learning opportunities. However, expectations do not always match or align. In order to find alignment, it is important that the student and supervisor are aware of each other’s expectations about responsibilities and that these are discussed. Yet, such discussions often do not take place or are experienced as difficult due to perceived hierarchy or an uneven relationship because supervisors commonly evaluate students. Nevertheless, a conversation about expectations may be helpful as it is often difficult to figure out expectations. To facilitate this conversation, a specific tool, called the AIR (Aligning Ideas about Responsibility) tool, was developed based on a focus group study and literature review.

Using the AIR tool is a two-step process. Firstly, a student and his supervisor independently express their own expectations using the artifact shown in Figure 1. Afterwards, these expectations are shared and discussed. The conversation about expectations precipitated by the tool, and not the data points themselves, is the key. Preferably, this conversation takes place at the beginning of a workplace-based learning experience in order to make expectations explicit from the beginning. In addition, the AIR tool can be used at a later stage, e.g., to refer to agreements made earlier. Such conversations may enhance alignment of expectations and ultimately increase learning opportunities for students as well as supervisors. This tool highlights the importance of collaboration and partnership between student and supervisor in which there is a focus on mutual learning and fruitful engagement. This shift toward co-constructed learning environments emphasizes that students and supervisors should participate in dialog around learning. Conversations, such as those facilitated by the AIR tool, can help build a collaborative learning environment after students and supervisors become aware of expectations.

This article presents a validation study of the AIR tool. The study aimed to collect and interpret different types of validity evidence for the AIR tool across different medical schools. The investigation was set up from the perspective that validation is a process of collecting validity evidence. Following this line of thought, validity is seen as a hypothesis. Just as one can never prove a hypothesis, validity can never be proven but evidence can, as it accumulates, support or refute the validity hypothesis. In other words, collecting various types of validity evidence will help inform judgments regarding the usefulness and critical appraisal of an instrument. There are five possible sources of validity evidence: content, internal structure, relationships with other variables, response process, and consequences evidence. Three sources of validity evidence were relevant to this study: response process, content, and consequences evidence.

**Method**

**Collecting response process evidence**

One researcher (SP) conducted cognitive interviews with ten potential users of the tool (students and
supervisors) in order to collect evidence of response process validity.22 This type of validity evidence is defined as “the fit between the construct and the detailed nature of performance… actually engaged in”21(p3) This evidence is important to determining how potential users interpret the tool and whether this matches the designers’ intent.23,24 Participants were recruited via students’ study groups and supervisors’ meetings. The person who conducted the cognitive interviews did not know these participants. The cognitive interviews were based on a combination of the “think-aloud” technique and verbal probing.22 Participants were asked to read the AIR tool aloud and tell exactly what was on their mind. This allowed the researcher (SP) to comprehend the way in which the potential users understood and used the tool. Additional probe questions were designed to elicit detailed information to clarify what participants mentioned before. The analysis of the cognitive interviews was done by two researchers (SP and AR) and was based on coding and interpretation of written notes from the interview, using the thematic analysis approach.22,23 It was used to modify and improve the tool before the start of the modified Delphi procedure.

**Collecting content evidence**

A modified Delphi procedure, which is a hybrid of both the Delphi and Nominal Group Technique,26,27 was used to collect content evidence. Content evidence is described as “the relationship between the content of a test and the construct it is intended to measure.”21(p3) Content validation effort focused on reaching consensus across medical schools about the representativeness, clarity, and relevance of the tool. The modified Delphi procedure consisted of three rounds: (1) an initial questionnaire, based on the Delphi approach, followed by (2) a face-to-face expert panel consensus meeting conducted using the Nominal Group approach,27,28 and (3) requesting, via e-mail, the final approval from all participants for the design, which accommodated recommendations made in the previous phases.

The advantages of the Delphi approach in Round One included the potential inclusion of large numbers of participants who were geographically dispersed,29 its subject anonymity, which can reduce the effects of dominant individuals29,30 and greater reliability21 compared to the Nominal Group approach. These benefits were combined with the advantages of the Nominal Group approach in Round Two, namely a more nuanced understanding of reasons for disagreement and a higher chance for reaching consensus than with the Delphi approach.29,31 Despite the widespread utility of Delphi and Nominal Group methods and the variety of hybrid approaches available, there are no guidelines for conducting such studies within medical education.32 The present study was based on the recommendations formulated by Waggoner et al.32

**Selection of the expert panel**

The goal was to recruit from across different universities a representative sample of informants with relevant knowledge and experience to serve as content experts.30,33 The present study included three categories of experts: (1) general practice students3; (2) workplace supervisors with at least two years of experience supervising general practice students (3) coordinators of workplace-based learning experiences in general practice with at least one year of experience in this role. Although they are not the potential users of the tool, coordinators were included in this study for their expertise in workplace learning. Participants were sought from three universities in Belgium (Antwerp, Ghent and Leuven) and one in the Netherlands (Maastricht). Various universities were included in order to increase the generalizability of the results.

Panel members were identified through purposive sampling.34 Individuals were approached by the research team34 in combination with the snowball technique. Each prospective panel member was contacted electronically or face-to-face. They were provided with a letter detailing the purpose, procedures, and anticipated time commitment of the study. If he or she affirmed their ability to complete all three rounds, then they were invited to participate. Recruitment continued until ten panel members from each category of expert agreed to participate. Although there is no consensus on appropriate panel size, panels usually consist of 15 to 30 participants from the same discipline, or 5 to 10 per category from different professional groupings.30,33,35 Panel members were provided with a unique subject identification number to guarantee that each person’s identity remained confidential throughout the whole process.

**Round one: questionnaire**

The aim of Round One was to evaluate the following domains: representativeness, clarity, and relevance of the AIR tool. Representativeness was defined as how completely the items (as a whole) encompassed the overall topic of responsibility for learning.22 Clarity
related to how clearly the items were worded. Relevance was the extent to which each item actually related to specific aspects of responsibility for learning and the usefulness of the overall tool.\(^{22}\)

The questionnaire (Appendix 1) began with the original version of the tool and explanations about its purpose and use. This was followed by demographic questions of participants and questions about the tool’s representativeness, clarity, and relevance. The questionnaire contained yes/no-questions and open-ended questions. The responses on the yes/no-questions were descriptively analyzed (N and %). Consensus on these questions was reached if more than 80% of the participants indicated “yes.” The responses on the open-ended questions were analyzed by two researchers (SP & PP) using the thematic analysis.\(^{25}\) A summary of these responses was presented to the participants in Round Two in order to discuss in group and try to reach consensus on the issues mentioned. Additionally, some questions from the Technology Acceptance Model (TAM) rating scale\(^{36}\) were incorporated in the questionnaire. Each question of the TAM rating scale was translated into Dutch and then revised by two researchers using the translation/back-translation method in order to avoid semantic problems.\(^{37}\) Participants were asked to rate their level of agreement with each statement of the TAM using a six-point scale ranging from “strongly disagree” to “strongly agree.” A general standard of how to measure consensus does not yet exist.\(^{38}\) Therefore, the analysis of these questions was based on previous research in which consensus was deemed to have been reached when more than 80% of the participants “strongly agreed” or “agreed” with a statement.\(^{29}\)

The questionnaire was pilot-tested with ten potential respondents to ensure its accuracy and clarity, and was revised based on their feedback.\(^{39}\) The questionnaire was distributed electronically or in hard copy to the expert panel. Weekly reminders were sent to panel members who had not completed the questionnaire after two weeks until either experts responded or six weeks had elapsed.

**Round two: face-to-face consensus meeting**

The findings of Round One were presented to all participants in face-to-face meetings. The aim of Round Two was again to reach consensus about the representativeness, clarity, and relevance of the tool. The expert panel was invited to offer suggestions where they considered one of the three domains as needing improvement. The responses to the open-ended questions of Round One, which were printed for all participants, were used as a starting point of the discussion. The aspects that had already reached consensus in Round One were not discussed.

It was not feasible for every participant of Round One to gather at a single meeting. Therefore, several meetings were organized. Online participation in the meetings was possible. A moderator (AR or GC) facilitated the discussions. Online participation in the meetings was possible. A moderator (AR or GC) facilitated the discussions, and another researcher took notes (SP). Consensus was reached if all participants who were present at the meeting agreed upon an issue (e.g. the phrasing of a sentence or the lay-out of the tool). For each meeting, all items on which agreement was reached were summarized in a consensus document. This document was developed by a researcher (SP) at the end of the meeting. The meetings were video-recorded in order for a second researcher (PP, NM or LS) to check the accuracy of the consensus document. If consensus at one meeting would not be the same consensus achieved at another meeting, three researchers (SP, AR and GC) would discuss how to solve this issue.

**Round three: requesting final approval**

All consensus documents from Round Two were taken into account and thoroughly discussed and evaluated by three researchers (AR, GC & SP). Afterwards, adjustments were made to the tool. This last version of the tool was presented to all participants electronically, and in hard copy if necessary, in order to request final approval via e-mail from each individual. Participants also had the opportunity to provide final comments on the tool.

**Collecting consequences evidence**

Once the modified Delphi was completed, the AIR tool was included as compulsory task during a six-week internship for first-year students in post-graduate general practice (GP) training at the University of Leuven in Belgium. This post-graduate GP course comprises three years. Students can enter this course when they have successfully completed six years of general medical training. During the first year of the GP course, upon which this study focused, students follow an integrated curriculum in which they spend around 23 weeks at the university, punctuated by six weeks internship at a general practice and 14 weeks of internship across various departments at hospitals. During the six-week internship at a GP setting students were allowed to perform parts of a patient consultation (e.g. taking patient history, doing a clinical investigation of a child with fever) with constant
supervision and guidance from their supervisor. Consequences evidence, “the impact, beneficial or harmful and intended or unintended, of a tool,” was collected in the academic year 2017–2018.

First, all students (N = 156) and supervisors were asked to fill in the AIR tool independently at the beginning of the internship. For each item in the AIR tool a score was given between 1 and 5 depending on the position of the response/cross. A score 1 was given if the response was located on the left-hand side and a score 5 on the right hand side of the continuum. The comparison between students’ scores and supervisors’ scores allowed for the establishment of whether there was an initial alignment or not. An alignment was reached when students’ and supervisors’ responses were the same. This meant that they had the same scores for each item in the AIR tool.

After students and supervisors completed a conversation with the AIR tool, students rated the perceived degree of consensus reached through the conversation on a 10-point scale ranging from no consensus (1) to consensus (10). Students also rated their satisfaction with the conversation on a 10-point scale ranging from no consensus (1) to totally satisfied (10). These data were descriptively analyzed by SP.

The comparison between students’ scores and supervisors’ scores allowed for the establishment of whether there was an initial alignment or not. An alignment was reached when students’ and supervisors’ responses were the same. This meant that they had the same scores for each item in the AIR tool.

According to the Qualitative Analysis Guide of Leuven, the initial coding phase focused on small units of text to identify the most important ideas. In the second coding phase, conceptually related ideas were grouped into broader categories or themes. These broader themes were presented in the results section.

This study was approved by the Social and Societal Ethics Committee of the University of Leuven (Reference number G-2016 08 602/G-2017 08 886). Informed consent was obtained from all participants.

### Results

#### Response process evidence

Six students and four supervisors participated in the cognitive interviews. Their responses suggested that most items were clear and the participants used it as was intended by the researchers. Yet, some small changes to certain sentences were recommended. The suggestions related mainly to the phrasing of sentences, e.g., it was recommended to change the word “connecting” in Item 3, “connecting practice with classroom-based content” to “linking.”

### Content evidence

#### Round one: questionnaire

Sixty-two experts completed the questionnaire. Table 1 shows the composition of the expert panel. Four of the 62 experts did not provide their name or contact details in the questionnaire. Therefore, these experts could not be contacted for Round Two and Three. The completed questionnaires from these four individuals were still taken into account for the data analysis of Round One.

The analysis of the questionnaire data showed that a few participants (N = 15; 24.19%) indicated that the AIR tool could be enhanced by the addition of certain elements (e.g., adding “follow-up of learning goals”). A small number of experts (N = 4; 6.45%) thought that it would be beneficial to delete certain elements of the tool in order to increase the representativeness. A few participants (N = 15; 24.19%) indicated that certain elements of the tool could be adjusted in order to improve its clarity. Comments about clarity pertained to the visual elements of the tool (e.g., replacing the boxes with an arrow at each end of the row), the content of the tool (e.g., suggestions about phrasing), and the guidance given about how and when to use the tool (e.g., use the tool also as evaluation of the workplace-based learning experience). Nearly all experts perceived the tool to be relevant (N = 57; 91.94%) and feasible (N = 56; 90.32%) for use in general practice. The experts’ responses on the TAM

---

**Table 1. Composition of expert panel.**

| Category of experts: | N (%) | Mean (range) |
|----------------------|-------|--------------|
| **Age** | | |
| Female | 43 (69.35%) | 39 (23–63) |
| **Category of experts:** | | |
| Coordinators | 8 (12.90%) | | |
| • Amount of years of experience as a general practitioner | 27 (14–35) | |
| • Amount of years of experience as a workplace supervisor | 19 (1–30) | |
| • Amount of students they supervise per year | 3 (2–5) | |
| Workplace supervisors | 26 (41.94%) | | |
| • Amount of years of experience as a general practitioner | 23 (8–38) | |
| • Amount of years of experience as a workplace supervisor | 11 (1–25) | |
| • Amount of students they supervise per year | 3 (1–7) | |
| Students | 28 (45.16%) | | |
| **University:** | | |
| • Antwerp (Belgium) | 15 (24.19%) | | |
| • Ghent (Belgium) | 16 (25.81%) | | |
| • Leuven (Belgium) | 20 (32.26%) | | |
| • Maastricht (the Netherlands) | 11 (17.74%) | | |
Table 2. Percentage of experts voting strongly agree or agree with each statement of the TAM.

| Statement of the TAM rating scale                                                                 | Experts voting strongly agree or agree |
|---------------------------------------------------------------------------------------------------|----------------------------------------|
| 1. Using the AIR tool improves the alignment of students’ and supervisors’ expectations (about responsibilities for enabling learning at the workplace) | 82.26% (=consensus)                     |
| 2. Using the AIR tool increases the productivity of finding alignment between students’ and supervisors’ expectations (about responsibilities for enabling learning at the workplace) | 64.52%                                 |
| 3. I find the AIR tool useful for finding alignment between students’ and supervisors’ expectations (about responsibilities for enabling learning at the workplace) | 64.52%                                 |
| 4. Using the AIR tool enhances the effectiveness of finding alignment between students’ and supervisors’ expectations (about responsibilities for enabling learning at the workplace) | 61.29%                                 |

rating scale also indicated that the tool was relevant or useful for discussing expectations regarding responsibilities about workplace learning. Consensus (at least 80% of experts voting strongly agree or agree) was reached for one statement of the TAM rating scale (see Table 2). The statements of TAM rating scale used in this study showed a high reliability (Cronbach’s α 0.86).

Experts expressed that the tool is relevant as (1) it helps with explicating expectations, (2) it makes discussing expectations more approachable, (3) it supports finding alignment between different expectations, (4) it helps to reach agreements and (5) it contributes to better collaboration between students and supervisors. Some of the participants’ comments pertained to the implementation of the tool in medical training. Participants expressed the need to have the tool easily accessible (e.g. both online and in hard-copy) to make students and supervisors familiar with using the tool and to implement the tool in multiple workplace-based experiences over different years of the medical training.

Round two: face-to-face consensus meeting

Twenty-four of the 62 participants in Round One took part in a group discussion (38.71% of participants in Round One). Of the twenty-four participants, eleven were students (45.83%), eleven were workplace supervisors (45.83%) and two were faculty coordinators (8.33%), which reflected the distribution of participants in Round One. Many experts were not able to be present due to time limitations, conflicting agenda issues or illness. There was at least one group discussion at each university. A group discussion consisted of four to seven experts and took, on average, an hour. Each discussion concluded with a consensus document including all changes to the tool that participants agreed upon. The various consensus documents had a lot of overlap and did not contain contradictory items. There were no differences between students’, supervisors’, and coordinators’ comments. Overall agreements about changes in the tool are shown in Table 3.

Round three: requesting final approval

A revised tool (see Figure 2; Appendix 2), including a page of explanation, was developed based on Round One and Two. Fifty-six of the 62 experts responded (90.32% of participants from Round One) and all approved the revised tool. Consensus was reached about the representativeness, clarity, and relevance of the revised tool.

Consequences evidence

The data of 144 of the 156 students, and their corresponding supervisors, were used for the data analysis (response rate = 92.31%). Dropout was attributable to incomplete AIR tools and the absence of a signed informed consent from some students.

The scored AIR tools indicated that 127 of the 144 students (88.19%) had no initial alignment with their supervisor before the conversation with the AIR tool. The students’ perceived degree of consensus after the conversation with the AIR tool (mean score 9.00) and their perceived satisfaction about the conversation (mean score 9.03) was very high (Table 4).

The analysis of students’ comments showed that students found the conversation with the AIR tool very useful for several reasons. Firstly, students felt that it “gave permission” to seek and discuss feedback. It enhanced an open communication with the supervisor from the beginning of the internship. A student explained: "The AIR tool is a very useful instrument in

Table 3. Overall agreements about changes in the tool.

- Add an attachment with extra information and an example of each item in the AIR tool
- Indicate key words in bold in the instructions
- Explicitly emphasize that the AIR tool is only the beginning of a conversation
- In the instructions, add that the AIR tool can also be used as evaluation (halfway or at the end of the workplace-based experience)
- Replace the boxes by an arrow in each row
- Change the phrasing of item 1 ‘formulating goals about applying competences’ to ‘discussing and following up learning goals’
- Change the phrasing of item 3 ‘linking practice with classroom-based content’ to ‘linking of clinical experiences with classroom-based content’
- Clarify the difference between feedback and reflection
- Add an extra row as ‘optional item’ for specific needs of the student or supervisor
the discussion of expectations, which would otherwise maybe only become explicit during or at the end of the internship”. Another student described it as: “The tool facilitated the conversation about expectations, something that didn’t happen or was very difficult during other internships”. Secondly, the tool identified gaps in expectations and facilitated bridging those gaps. A participant mentioned: “Hearing each other’s point of view allowed us to reach consensus and make concrete agreements”. Moreover, students found the AIR tool very feasible and easy to use.

Students who did not find the conversation with the AIR tool useful commented that they already enjoyed open communication with their supervisor and that the AIR tool was unnecessary in that case. One student said: “There was only a small discussion about the AIR tool. Everything went well and quickly. We agreed immediately. It was a short conversation. There was limited added value of the AIR tool to this conversation.” Additionally, students mentioned that their supervisor did not say much during the conversation so it did not have an impact on the students’ learning opportunities or that the supervisor did not act upon the agreements made with the AIR tool.

Students also emphasized unintended consequences of using the tool. During the conversation with the AIR tool, students and supervisors sometimes made agreements about issues other than those included in the tool, e.g., who is responsible for completing internship tasks and about being open for questions and meetings with other team members in a group practice.

Discussion
This study collected and interpreted validity evidence for the AIR tool across four different medical schools. The completed AIR tools showed that most students and supervisors did not have an initial alignment of expectations. The conversation, for which the AIR tool aims to be a catalyst, appeared to facilitate this alignment of expectations about responsibility during workplace learning. Students’ high perceived degree of consensus and satisfaction after the conversation stressed the relevance of the AIR tool.

Table 4. Descriptive statistics.

|                                | N   | Minimum | Maximum | Mean  | Std. deviation |
|--------------------------------|-----|---------|---------|-------|----------------|
| Perceived degree of consensus  | 144 | 6       | 10      | 9.00  | 0.99           |
| Perceived satisfaction         | 144 | 5       | 10      | 9.03  | 1.21           |
Learning in the workplace requires students and supervisors to collaborate successfully, which can be a challenging task. Such collaboration, including understanding each other’s expectations and reaching a mutual understanding about how to learn and provide patient care, might happen spontaneously by working together. Others might have difficulties and need guidance. The AIR tool can help students and supervisors to overcome several problems within workplace learning, e.g., unclear roles and unknown, or different, expectations. This study indicated that students felt that the AIR tool lowered the threshold of discussing expectations, and it enhanced open communication with supervisors. In this way, the AIR tool embraces the importance of communication and discussions about learning in the workplace, which has been emphasized by other researchers.

This approach also is consistent with the literature regarding co-constructed learning environments. Conversations such as those primed by the AIR tool focus on collective learning rather than individual learning. Becoming aware of each other’s expectations is an important first step toward creating a collaborative learning environment. It is assumed that clarity about expectations and responsibilities will lead to students learning more during workplace-based experiences. Future research will have to investigate this assumption.

Although the strength of this study was that various groups of experts (students, supervisors, and coordinators) from multiple universities across borders were involved, a limitation was that the response rate was low in Round Two of the modified Delphi. Yet, the response rate was very high in Round Three. Another limitation may be that coordinators seem to be underrepresented in this study, but this is a reflection of the reality of this branch of the medical field. At each university there are fewer coordinators than students and supervisors. A strength of the study was its rigorous methodological procedures. Moreover, given that the AIR tool was based on previous literature and qualitative study in the medical context, it has a strong theoretical and practical foundation.

The AIR tool was used for one-on-one contact between student and supervisor, but it could also be used for more than two individuals, e.g., by one or more students with one or more workplace supervisors. In those cases, the conversations that follow the use of the AIR tool will become especially relevant to reaching agreement about each other’s responsibilities and the division of tasks. Moreover, the AIR tool could also be useful beyond general practice training (e.g. internships in internal medicine, surgery, and so forth) and perhaps even beyond a medical training context (e.g. training of pilots). The five items in the tool are generic and do not solely focus on general practice training or a specific medical context. By presenting the last item in the tool as blank and open for users to fill in, this allows students and supervisors to add an item based on their needs or specific character of their context. This study highlighted some positive consequences of using the AIR tool. Future research needs to investigate what are the learning outcomes of using the tool in the medical context and elsewhere.

Conclusion

Students, supervisors and coordinators from four different universities underlined the relevance and usefulness of a tool that facilitates a conversation about expectations regarding responsibility for enhancing learning opportunities at the workplace. It is expected that the tool can easily be implemented in medical training, as well as in other contexts.

Notes

1. All competences acquired on the university campus (e.g. at the skills lab or during simulation-based sessions)
2. This term refers to all individuals who collaborate with a supervisor
3. Depending on the context, the terms ‘trainees’ and ‘residents’ also were used.

Acknowledgement

The authors would like to thank Evi Grauwet for her help in the data collection.

ORCID

Sanne Peters http://orcid.org/0000-0001-6235-1752
Nele Michels http://orcid.org/0000-0003-1971-0793
Peter Pype http://orcid.org/0000-0003-2273-0250
Lorette Stammen http://orcid.org/0000-0003-1528-546X

References

1. Yardley S, Littlewood S, Margolis SA, et al. What has changed in the evidence for early experience? Update of a BEME systematic review. Med Teach. 2010;32(9):740–746. doi:10.3109/0142159X.2010.496007.
2. Blume BD, Ford JK, Baldwin TT, et al. Transfer of training: a meta-analytic review. J Manage. 2010;36(4):1065–1105. doi:10.1177/0149206309352880.
3. Burke LA, Hutchins HM. Training transfer: an integrated literature review. Hum Resour Dev Rev. 2007;6(3):263–296. doi:10.1177/1534484307303035.
4. Tuomi-Gröhn T, Engeström Y, Between School and Work: New Perspectives on Transfer and Boundary-Crossing. 1st ed. Amsterdam; Boston: Pergamon; 2003.

5. Peters S, Clarebout G, Van Nuland M, et al. A qualitative exploration of multiple perspectives on transfer of learning between classroom and clinical workplace. Teach Learn Med. 2018;30(1):22–32. doi:10.1080/10401334.2017.1339605.

6. Perrella A, Milman T, Ginsburg S, et al. Navigating tensions of efficiency and caring in clerkship: a qualitative study. Teach Learn Med. 2019;31(4):378–384. [published Online First: 2019/01/01] doi:10.1080/10401334.2018.1556667.

7. Henry-Noel N, Bishop M, Gwede CK, et al. Mentoring in medicine and other health professions. J Cancer Educ. 2019;34(4):629–637. [published Online First: 2018/04/25] doi:10.1007/s13187-018-1360-6.

8. Gaukel EH, Batalden M, Sands R, et al. The hidden curriculum: what can we learn from third-year medical student narrative reflections? Acad Med. 2010;85(11):1709–1716. [published Online First: 2010/10/01] doi:10.1097/AMJ.0b013e3181f57899.

9. Soo J, Brett-MacLean P, Cave MF, et al. At the precipice: a prospective exploration of medical students’ expectations of the pre-ckeropk to clerkship transition. Adv Health Sci Educ. 2016;21(1):141–162. [published Online First: 2015/07/15] doi:10.1007/s10459-015-9620-2.

10. Peters S, Clarebout G, Van Nuland M, et al. How to connect classroom and workplace learning. Clin Teach. 2017;14(5):313–319. doi:10.1111/ctc.12697.

11. Farrell L, Bourgeois-Law G, Buydens S, et al. Your goals, my goals: the complexity of co-constructing goals with learners in medical education. Teach Learn Med. 2019;31(4):370–377. [published Online First: 2019/03/16] doi:10.1080/10401334.2019.1576526.

12. Sundler AJ, Bjork M, Bisholt B, et al. Student nurses’ experiences of the clinical learning environment in relation to the organization of supervision: a questionnaire survey. Nurse Educ Today. 2014;34(4):661–666. doi:10.1016/j.nedt.2013.06.023.

13. Burke LA, Saks AM. Accountability in training transfer: adapting Schlenker’s model of responsibility to a persistent but solvable problem. Hum Resour Dev Rev. 2009;8(3):382–402. doi:10.1177/1534484309357632.

14. Brown J. Transferring clinical communication skills from the classroom to the clinical environment: perceptions of a group of medical students in the United Kingdom. Acad Med. 2010;85(6):1052–1059. doi:10.1097/ACM.0b013e3181db76f.

15. Yardley S. Sense made common: how to add value to early experience. Clin Teach. 2014;11(1):5–9. doi:10.1111/tct.12050.

16. Masters KS, Kreeger PK. Ten simple rules for developing a mentor–mentee expectations document. PLoS Comput Biol. 2017;13(9):e1005709. [published Online First: 2017/09/22] doi:10.1371/journal.pcbi.1005709.

17. Hattie J. Bringing it all together. In: Hattie JAC, ed. Visible Learning: A Synthesis over 800 Meta-Analysis Relating to Achievement. London: Routledge; 2009:257–261.

18. Ho A, Watkins D, Kelly M. The conceptual change approach to improving teaching and learning: an evaluation of a Hong Kong staff development programme. Higher Educ. 2001;42(2):143–169.

19. Geraci SA, Thigpen SC. A review of mentoring in academic medicine. Am J Med Sci. 2017;353(2):151–157. [published Online First: 2017/02/12] doi:10.1016/j.amjms.2016.12.002.

20. Reusser K. Co-constructivism in educational theory and practice. In: Smelser NJ, Baltes B, eds. International Encyclopedia of the Social and Behavioral Sciences. London, UK: Elsevier Ltd; 2001:2058–2062.

21. Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. Adv Simul. 2016;1(1):31. [published Online First: 2016/12/07] doi:10.1186/s41077-016-0033-y.

22. Artino AR, Jr., La Rochelle JS, Dezee KJ, et al. Developing questionnaires for educational research: AMEE Guide No. 87. Med Teach. 2014;36(6):463–474. doi:10.3109/0142159X.2014.889814.

23. Willis GB. Cognitive Interviewing: A Tool for Improving Questionnaire Design. Thousand Oaks, CA: Sage Publications; 2005.

24. Willis GB, Artino AR, Jr. What do our respondents think we’re asking? Using cognitive interviewing to improve medical education surveys. J Grad Med Educ. 2013;5(3):353–356. [published Online First: 2014/01/10] doi:10.4300/JGME-D-13-00154.1.

25. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77–101. doi:10.1191/1478088706qp063oa.

26. Campbell SM, Braspenninck J, Hutchinson A, et al. Research methods used in developing and applying quality indicators in primary care. BMJ. 2003;326(7393):816–819. doi:10.1136/bmj.326.7393.816.

27. Foth T, Efstadthiou N, Vanderspank-Wright B, et al. The use of Delphi and Nominal Group technique in nursing education: a review. Int J Nurs Stud. 2016;60:112–120. doi:10.1016/j.ijnurstu.2016.04.015.

28. Raine R, Sanderson C, Black N. Developing clinical guidelines: a challenge to current methods. BMJ. 2005;331(7517):631–633. doi:10.1136/bmj.331.7517.631.

29. Humphrey-Murto S, Varpio L, Gonsalves C, et al. Using consensus group methods such as Delphi and Nominal Group in medical education research. Med Teach. 2017;39(1):14–19. doi:10.1080/0142159X.2017.1245856.

30. Hsu CC, Sandford BA. The Delphi technique: making sense of consensus. Pract Assess Res Eval. 2007;12(10):1–8.

31. Hutchings A, Raine R, Sanderson C, et al. A comparison of formal consensus methods used for developing clinical guidelines. J Health Serv Res Policy. 2006;11(4):218–224. doi:10.1258/13558190678476553.

32. Waggoner J, Carlone JD, Durning SJ. Is there a consensus on consensus methodology? Descriptions and recommendations for future consensus research. Acad Med. 2016;91(5):663–668. doi:10.1097/ACM.0000000000001092.
33. de Villiers MR, de Villiers PJ, Kent AP. The Delphi technique in health sciences education research. Med Teach. 2005;27(7):639–643. doi:10.1080/1361126050069947.

34. Morgan PJ, Lam-McCulloch J, Herold-Mclroy J, et al. Simulation performance checklist generation using the Delphi technique. Can J Anesth. 2007;54(12):992–997. doi:10.1007/BF03016633.

35. Clayton MJ. Delphi: a technique to harness expert opinion for critical decision-making tasks in education. Educ Psychol. 1997;17(4):373–386. doi:10.1080/0144341970170401.

36. Venkatesh V, Morris MG, Davis GB, et al. User acceptance of information technology: toward a unified view. MIS Q. 2003;27(3):425–478. doi:10.2307/30036540.

37. Behling O, Law KS. Translating Questionnaires and Other Research Instruments: Problems and Solutions. Thousand Oaks, CA: Sage Publications; 2000.

38. von der Gracht HA. Consensus measurement in Delphi studies: review and implications for future quality assurance. Technol Forecast Soc Change. 2012;79:1525–1536. doi:10.1016/j.techfore.2012.04.013.

39. Haji FA, Khan R, Regehr G, et al. Operationalising elaboration theory for simulation instruction design: a Delphi study. Med Educ. 2015;49(6):576–588. doi:10.1111/medu.12726.

40. Savin-Baden M, Major CH. Qualitative Research: The Essential Guide to Theory and Practice. Abingdon: Routledge; 2013.

41. Dierckx de Casterlé B, Gastmans C, Bryon E, et al. QUAGOL: a guide for qualitative data analysis. Int J Nurs Stud. 2012;49(3):360–371. doi:10.1016/j.ijnurstu.2011.09.012.

42. Olmos-Vega FM, Dolmans D, Vargas-Castro N, et al. Dealing with the tension: how residents seek autonomy and participation in the workplace. Med Educ. 2017;51(7):699–707. [published Online First: 2017/04/25] doi:10.1111/medu.13326.

43. Teunissen PW, Scheele F, Scherpber A, et al. How residents learn: qualitative evidence for the pivotal role of clinical activities. Med Educ. 2007;41(8):763–770. doi:10.1111/j.1365-2923.2007.02778.x.

44. Berkhout JJ, Helmich E, Teunissen PW, et al. Context matters when striving to promote active and lifelong learning in medical education. Med Educ. 2018;52(1):34–44. [published Online First: 2017/10/07] doi:10.1111/medu.13463.

Appendix 1: Questionnaire in round one

1. Information about the AIR tool
   Why the AIR tool?
   During focus group discussions with students of the general practice training, teachers and supervisors at the KU Leuven, it was apparent that it is often unclear who is responsible for the learning process during the internship and what students’ and supervisors’ expectations are. It also seems that students and supervisors often do not communicate about these expectations. The AIR tool aims to deal with this issue.

   The aim of the AIR tool:
   It is a tool to support a conversation between the student and supervisor. The conversation deals with each other’s expectations about the responsibility for the learning process during the internship. The aim is to find alignment of expectations and make agreements about responsibilities. It is an advantage for both the student and the supervisor to know each other’s expectations from the start of the internship. Using this tool could possibly also lead to more learning opportunities for the student.

   The AIR tool:
   
   | Who is responsible for ...? |
   |-----------------------------|
   | Supervisor                  |
   | Student                     |
   | Formulating goals about applying competences |
   | Organizing opportunities to practice |
   | Linking practice with classroom-based content |
   | Taking initiative for feedback (on past performance) |
   | Taking initiative for reflection (on future performance) |

   How to use the AIR tool?
   1. The student and supervisor independently fill in the AIR tool (one cross on each row).
   2. They discuss both AIR tool responses (preferably at the beginning of the workplace-based learning experience). The AIR tool is solely a starting point to support the conversation.
   3. They try to find alignment of expectations.
2. Questionnaire

- Sex: M - F
- Age:
- Current position:
  - Student in general practice training
  - University coordinator of workplace-based learning
  - Supervisor at general practice
- Amount of years of experience as a GP:
- Amount of years of experience as a supervisor:
  - Average amount of GP students that you supervise per year:
- To which university are you connected?
  - University of Antwerp
  - University of Gent
  - University of Maastricht
  - University of Leuven

| Question                                                                 | Totally disagree | Disagree | Slightly disagree | Slightly agree | Agree | Totally agree |
|-------------------------------------------------------------------------|------------------|----------|------------------|---------------|-------|---------------|
| Using the AIR tool improves the alignment of students’ and supervisors’ expectations (about responsibilities for learning at the workplace). |                  |          |                  |               |       |               |
| Using the AIR tool increases the productivity of finding alignment between students’ and supervisors’ expectations (about responsibilities for learning at the workplace). |                  |          |                  |               |       |               |
| I find the AIR tool useful for finding alignment between students’ and supervisors’ expectations (about responsibilities for learning at the workplace). |                  |          |                  |               |       |               |
| Using the AIR tool enhances the effectiveness of finding alignment between students’ and supervisors’ expectations (about responsibilities for learning at the workplace). |                  |          |                  |               |       |               |

- Are there certain elements that might be adjusted in order to enhance the clarity and comprehensibility of the AIR tool?
  - Yes – No
  - If yes, what?
- Are there elements that you would add to the AIR tool?
  - Yes – No
  - If yes, which?
- Are there elements that you would delete?
  - Yes – No
  - If yes, which?
- Do you think it is feasible to use the AIR tool in practice?
  - Yes – No
- How might the feasibility be increased?
- Do you think that the AIR tool is relevant for general practice?
  - Yes – No
  - Why?
- How could the relevance be increased?
- What suggestions do you have regarding improving the representativeness, clarity or relevance of the AIR tool?
- Other remarks/suggestions?

Appendix 2

Aim:
Responsibility for the learning process at the workplace is often unclear for students and supervisors. A division of responsibilities and tasks is often not made. This AIR (‘Aligning Ideas about Responsibility’) tool aims to facilitate a conversation between student and supervisor. This conversation deals with each other’s expectations about responsibility during the learning process at the workplace. The aim is to get to know each other’s expectations from the start of the internship and make agreements about the division of responsibilities for each task. The AIR tool can also be used at the culmination of the internship (e.g. to evaluate the effectiveness of the tasks?).
Explanations for each item in the AIR tool:

| During the internship the student applies classroom-based learning content into practice. |  |
|---|---|
| **Who is responsible for ... during the internship?** |  |
| Discussing learning goals and follow up |  |
| The student chooses specific aspects on which he/she would like to focus during the internship (= learning goals). These learning goals focus on the deliberate application in practice of what the student learned at university. It is important to discuss and reflect on these goals in order to increase learning opportunities. |  |
| *Indicate who takes the initiative for this task.* |  |
| Organizing opportunities to practise |  |
| The student learns a lot by doing consultations themselves (history taking, clinical examination and management). |  |
| *Indicate who takes the initiative to create opportunities to practise.* |  |
| Linking practice with classroom-based content |  |
| The internship is a good opportunity to link practice with knowledge and skills acquired at university (e.g. there is a patient with diabetics in the general practice. What knowledge and skills about diabetics did the student learn? In which way can this applied to this patient? Why/ why not?). |  |
| *Indicate who is responsible for linking practice with classroom-acquired knowledge and skills.* |  |
| Taking initiative for feedback |  |
| Feedback involves comparing the student’s performance with how the supervisor would have done it or how it is described in Evidence-Based guidelines. |  |
| *Indicate who takes the initiative to discuss the question “How did the student perform the task?”.* |  |
| Taking initiative for reflection |  |
| Reflection entails considering past experiences in order to improve future performance. |  |
| *Indicate who takes the initiative to discuss “What did the student learn from this experience which prepares them for the performance of a similar task in the future?”.* |  |
|  |  |
| The student and supervisor can add an extra aspect that they would like to discuss, if desired. |  |