Article

Peer Assessment: Channels of Operation

Keith Topping

School of Education, University of Dundee, Dundee DD1 4HN, UK; k.j.topping@dundee.ac.uk

Abstract: The present paper offers a definition of peer assessment and then reviews the major syntheses on its effectiveness. However, the main part of this paper is preoccupied with how to do PA successfully. A typology of 44 elements explains the differences between the many types of peer assessment. Then a theoretical model outlines some of the processes which may occur during PA. Initially, only a few of these will be used, but as those engaged in PA become more experienced, an increasing number of elements will feature. However, these may not appear in the linear order set out here, and indeed may be recursive. The implications for the design and organisation of PA are outlined, as well as the implications for future research.

1. Peer Assessment: Channels of Operation

Feedback is widely considered important in education [1] and peer assessment (PA) is one method of enhancing the speed and quantity of feedback, if not the quality. Many professions may expect to engage in PA as part of their vocations, so its value goes beyond school and university.

2. What Is Peer Assessment?

A widely quoted definition of PA is: "an arrangement for learners to consider and specify the level, value or quality of a product or performance of other equal-status learners" [2] (p. 256). However, other similar terms (synonyms) are in the literature (e.g., peer grading/marking—giving a score to a peer product or performance; peer feedback—peers giving elaborated feedback; peer evaluation—more usually in workplaces regarding skill and knowledge; or peer review—more usually in academia regarding assessment of written papers).

3. Does Peer Assessment Work?

PA is not just for managing assessment burdens for teachers, but more importantly a mechanism for more effective learning, particularly with elaborated feedback. For the assessor, the intellectual demands of reflecting, making a balanced assessment, formulating and delivering feedback can all lead to learning gains [3]. For the assessee, the intellectual demands of receiving and evaluating the feedback, deciding what aspects to implement and what not, and reflecting on other issues prompted by the feedback (but not contained within it) can all lead to learning gains [4].

The evidence on PA with all kinds of learners is generally positive, from the earliest reviews (e.g., [5] on peer grades and feedback; [6] on peer grades) to the latest meta-analyses (e.g., [7,8]). An early systematic literature review on the effects of PA appeared in 2009 [9]. Fifteen studies from 1990 to 2009 dealt with effects on achievement. However, only one of these studies included students from a school, the remainder consisting of university students. PA had positive effects. The authors offered four underlying constructs: psychological safety, value diversity, interdependence and trust. Psychological safety was defined as a belief that it was safe to take interpersonal risks in a group of people. Value diversity referred to differences in opinion about what a team's task, goal or mission should

be—it should be low for PA to be effective. Interdependence has been long studied, but needs to be perceived by the participants rather than assumed by teaching staff. It requires that multiple perspectives are made explicit and students are individually responsible for an active contribution to group discussions. In respect of trust, several studies noted that students felt uncomfortable criticising one another's work, or at least initially found it difficult to rate their peers.

Another study [10] considered what quality criteria were specifically relevant to PA. One hundred and thirty-two studies of PA were selected, together with 42 studies for a qualitative analysis. Nowhere was any distinction made between studies based in school, higher education or other settings. Studies were evaluated with regard to two quality criteria: (1) the recognition of educational measurement criteria, and (2) the consideration of student involvement in the assessment of learning. Where emphasis was placed on authenticity and future learning needs across the lifespan, PA had much to recommend it in terms of generalisability, particularly utility in contexts beyond the present institution.

Only one review was solely concerned with PA in schools [11], analysing 26 studies of peer response on writing proficiency. The author noted that several studies had indicated that peer response was effective, but had not explored why. Many studies appeared to combine instruction in strategies, rules for interaction, and/or genre knowledge—and this seemed to be effective compared to individual writing.

The first meta-analysis of PA [12] studied PA in digital platforms since 1999, again mainly in universities, finding a moderately strong average correlation between peer and teacher ratings of 0.63. This correlation was higher when: (a) the PA was paper based rather than computer assisted; (b) the subject area was not medical/clinical; (c) the course was graduate level rather than undergraduate or in school; (d) individual work instead of group work was assessed; (e) the assessors and assessees were matched at random; (f) the PA was voluntary instead of compulsory; (g) the PA was not anonymous; (h) peer raters provided both scores and qualitative comments instead of only scores; and (i) peer raters were involved in developing the rating criteria.

Turning to the latest meta-analyses, one [7] found an overall effect size (ES) of 0.29 in 58 studies (an effect size is a number measuring the strength of the relationship between two variables, which can apply across all studies). Significant moderator variables were found of training and online/digital (moderator variables are third order variables that affect the size or nature of the relationship between an independent and dependent variable). Another meta-analysis [8] found an overall ES of 0.31 in 54 studies, but no significant moderator variables. In both cases the ESs were lower than previous studies.

In PA studies, it is often assumed that teacher "expert" assessment should be the criterion for validity, but both these studies showed PA was more reliable and had higher ESs than teacher assessment, although in fact teacher assessment is not very reliable [13].

4. Digital Peer Assessment

It is unsurprising that digital PA has been separately reviewed, given the widespread recent move towards online methods and the fact that PA in large university courses can only be managed by such means (e.g., [14,15]). The first meta-analysis [16] found 37 controlled studies from 1999 to 2018. Eight studies were in school and the rest in higher education, and again this mixing of contexts without discrimination is a weakness. Of the 37 studies, 19 examined outcomes (overall ES 0.58) and 17 the effects of extra supporting strategies (ES 0.54).

These ESs would be considered "moderate" by most researchers, but are larger than those reported most recently for PA in general (above), suggesting that (despite some disadvantages), digital PA has countervailing advantages that make it more effective than face-to-face PA. Training and anonymity improved outcomes, and duration of PA was also important (6–10 weeks being the optimum). However, direct comparison of online and offline learning was rare—most studies compared online PA to no PA.

However, here we are less concerned with whether PA works and more concerned with the how of PA, and we will consider a typology of PA and then a theoretical model of PA. Together, these should give practitioners a clearer idea of the how to successfully design and implement a PA project, and researchers a clear idea of the broad context of PA.

5. Typology of Peer Assessment

Several studies compare two or three types of PA, but the variety in types of PA goes far beyond that. Teachers need to be able to clearly categorise what they want to do—in a way which will also remind them of variables which they might have forgotten. It is important to be aware of what you are not doing as well as what you are. Different kinds of PA are more or less suitable for particular classroom contexts, different levels of maturity in the students, different subjects and assessed activities, and these are judgements the teacher must make.

A typology of relevant variables was first described in 1988 [2], Subsequently, a more developed inventory was offered [17]. Further developments [18] (pp. 12–13) in 2018 outlined 44 variables (see Table 1).

Table 1. Variations in Peer Assessment.

	Alternative A	Alternative B	Alternative C or Comment
1	Objectives: Cognitive Metacognitive	Objectives: Social Emotional	or both
2	Summative	Formative	or both
3	Quantitative grading	Qualitative feedback	or both
4	Voluntary	or Compulsory	
5	Digital technology used	No digital technology	or blended
6	Single product	Several products	
7	Same kind of product	Different products	
8	Same curriculum area	Different areas	
9	Individuals	Pairs	or groups
10	Assessment criteria clear	Not clear	
11	Students involved	Student not involved	in defining criteria
12	Rubric used	Rubric not used	
13	Training given to peers	Not given	
14	Feedback positive	Feedback negative	or both
15	Feedback → improvement	No improvement	
16	Product reworked	Not reworked	
17	Scaffolding given	Not given	prompts, cues, etc.
18	One-way	Reciprocal	or mutual in group
19	Matching deliberate	Matching random	or matching accidental
20	Matching academic	Matching social	or both
21	Same year of study	Different year of study	
22	Same class	Different class	
23	Same ability	Different ability	in this subject area
24	Previous experience of PA or peer learning	No previous experience	
25	Experience positive	Experience negative	or both
26	Cultural expectations positive	Cultural expectations negative	
27	Gender balance	Gender imbalance	ability, motivation, etc.?
28	In class	Out of class	or both
29	Length of sessions		
30	Number of sessions		
31	Arranged by peers	Arranged by teacher	
32	Justification to peer	No justification	
33	Confidentiality	No confidentiality	to pair + teacher + others
34	Anonymous	Non-anonymous	
35	Feedback expected	Not expected	quantity + quality
36	Feedback objective	Feedback subjective	or both
37	Revisions many	Revisions few	
38	Process monitored	Not monitored	
39	Reliability moderated	Not moderated	and validity

Table 1. *Cont.*

	Alternative A	Alternative B	Alternative C or Comment
40	Task simple	or complex	or simple → complex
41	Intrinsic rewards	Extrinsic rewards	neither
42	Aligned	Non-aligned	with other assessment
43	Transferable skills	None measured	
44	Evaluated	Not evaluated	

Proceeding through the list, firstly the objectives for the exercise may vary—the teacher may target cognitive and/or metacognitive gains, teacher time saving, or other goals. There may be other gains, such as social gains or attitudinal gains (e.g., better relationships, improved self-confidence, improved motivation). Do you see peer assessors and assessees talking more out of class? Do you feel that some students are more engaged in what they are doing as a result of PA?

A key difference is whether the PA is formative or summative or both. Will it serve to give students indications of how to improve their work (formative), so the final version can be better? Or will it just indicate to the students how good or bad their work was (summative), with no opportunity for improvement?

Similarly, the PA can be quantitative (assigning a number with respect to a grade) or qualitative (giving rich verbal feedback on positive and negative aspects and possibilities for improvement), or both. If students are merely to give a grade, they will need considerable experience in grading before their grades can be considered reliable. Further, even if they are reliable, they do not give the assessee any clues on how to improve their work the next time. By contrast, qualitative feedback gives rich ideas on how to improve the current piece of work, let alone future pieces of work. The assessee may not agree with all of these, but some negotiation of the nature of improvement can follow.

Will PA be voluntary or compulsory? When it is used in a class, it would be a normal expectation that all students would participate, but if it is compulsory from the beginning, some students might be very resistant to participation. It might be better to say that it will be voluntary at the beginning. So few students are likely to opt out, that after a short while those who have opted out will realise that their opposition is unusual if not a little bizarre, and agree to join in.

Will you use some form of digital technology? This could be all online or it could be blended, with some face-to-face contact. This could help even if the PA was mostly occurring in class, e.g., having students rehearse their oral presentations on video on their mobile phones until they are satisfied with the performance, then have them upload the final version to a common location (e.g., GoogleDocs) for everyone to see, then meet face to face to discuss and conduct PAs. For more remote students and during pandemics with lockdowns, all PA will have to be online. If neither of these is relevant, it all could be face to face unless the number of students is too large to allow this.

Other differences between types of PA are more subtle. For example, are the PAs on single pieces of work, or are they of several pieces of work? A piece of writing is relatively easy to assess, as it has a beginning and an end. But even here you should not assume that peer assessors are only relevant after the writing has been completed. They could for instance be involved again as the writer tries to improve the piece of writing. Other products of work may be more complicated. For example, in PA of a group presentation, should the quality of discussion prior to the presentation itself be peer assessed?

Are PAs on the same kind of product? The product or output assessed can vary—writing, portfolios, presentations, oral statements, and so on. Assessment of writing is very different to assessment of an oral statement, which is in turn very different to PA in music or physical education. Students will need some experience of each kind of PA before they have confidence that they can manage the necessary tasks.

PA can operate in different curriculum areas or subjects, which may impose different demands. For example, in physical education classes, can peers be trained to investigate

differences in the way the other student runs, or catches a ball, or throws a javelin, and so on? In foreign language learning, how quickly might students be able to accurately respond to the comments or questions of a peer in the foreign language?

The participant constellation can vary, with consequent variation in joint responsibility for the assessed product. Assessors and assessed may be individuals, pairs or groups. Will you have one assessor and one assessee in a pair? Or a small group where everyone assesses all the productions of the other members of the group? Will their PA be reciprocal? Or will you have one cooperative group assessing another cooperative group—again, reciprocal or not? Be careful in supposedly cooperative groups that all members of the group have contributed. You could invite the group to assess each of its members on the size of their contribution to the group proceedings. Then the responsibility for the finished product is not unfairly apportioned to the lazy members of the group.

Will it be anonymous or not? Of course, if you have reciprocal face-to-face PA in one classroom, it is impossible to make it anonymous. But if you have one class assessing the work of another class, and giving feedback in writing or over the internet, it might be much more possible. But will you actually want the feedback to be anonymous? Peer feedback from somebody you know might be more powerful than that from somebody who is anonymous. But if you do not know your assessor, you might feel safer initially if they were anonymous.

Clarification of the assessment criteria is essential, and peers may or may not be involved in establishing these criteria. In general, however, peers should always be involved in the development of the assessment criteria, even if the teacher has their own ideas or there is some external assessment system that needs to be acknowledged. The fact that the peer group will eventually come up with very similar criteria to those the teacher would have given does not take away from the value to the peers of feeling engaged in the process. As a result, they know the criteria better from the outset.

Rubrics or structured formats listing assessment criteria for feedback may or may not be provided. However, assessment rubrics almost always help the assessors and the assesses. As above, they should be developed by the peer group. But having these criteria written down will help add consistency to the PA.

Training in PA may be given to assessors and/or assesses to a greater or lesser extent. It is surprising how many projects in the literature appeared to give no training to the peer assessors. Some training will be needed—the only question is: how extensive will it be? It cannot go on too long or the peer group will become restless to get some "real" activity. However, it should not merely involve the teacher talking. Some encounters with real life examples and some practice in actually applying PA should certainly feature as part of the training.

Is any feedback provided expected to be balanced between positive and negative, or only one of these? When you are starting with PA, you might be inclined to ask the peer assessors to provide only positive feedback. Then you get them used to the idea of being positive. Later, you can also ask them to give "suggestions for improvement", which of course are open to discussion. Once students are competent with both aspects of feedback, you can give them free rein, except that every piece of assessed work should have some positives and some negatives.

Is feedback expected to lead to opportunities to rework the product in the light of feedback, or is there no opportunity for this? Of course, we all hope that the current version of our work is the final one, so there might be some resistance to (apparently endlessly) reconsidering—although this is almost always going to result in a better piece of work. Negative feedback indicates where the work needs improving, and hopefully there will be time available to achieve this. A related question here is that of audience—why should the peer assessee try to improve the work? Who will tell the difference? Students need to see what the point is of improving.

Is feedback expected to include hints or suggestions for improvement? Negative feedback will be much more acceptable if it is accompanied with some suggestions for

improvement, even if those suggestions are not accepted. They give the assessee something to think about, and maybe they will then come up with a completely different way of doing things.

The nature of subsequent PA activity may be very precisely specified or it may be left loose and open to student creativity. Again, this may be a developmental issue, in that at the beginning, peer assessors and assessees may need a fairly strict procedure. Later, however, this may become looser, so that assessors may begin to give more feedback in their own time, as they develop a sense of responsibility towards their assessee.

Does the interaction involve guiding prompts, sentence openers, cue cards or other scaffolding devices? At the beginning of PA, one, some or all of these are a good idea, as some students will have little idea how to begin a PA conversation. Giving them some questions to use to get them started is an excellent idea—they do not necessarily need to use them.

PA can be one-way, reciprocal, or mutual within a group. If you have an older class assessing a younger class, directionality is likely to be one-way. If you are working with same-ability pairs in one class, directionality is likely to be reciprocal. If you are working with groups, does the group decide on a mutually agreed assessment for another group, or are the separate PAs of the other group to be taken into account? (requiring an agreed group assessment gives the group another valuable learning experience).

Matching of students may be deliberate and selective or it may be random or accidental. If the teacher is new to the class, it may need to be random. If the teacher knows something about the class members, one can be more careful. Matching may take account only of academic factors, or also involve social differences. The most able assessing the least able is not recommended. You may decide that you want the top half of the class assessing the bottom half of the class. Or you may decide that you want students to be matched based on having similar abilities, especially if you are doing reciprocal PA. Or you may decide that while ability is relevant, personality and social issues are also relevant.

Assessors and assessed may come from the same year of study or from different years. If you have a colleague from a class of a similar age who is also interested, you could certainly see if the two classes could be matched up for the purposes of PA. If the classes are more or less of the same size, you have an ideal opportunity. But many teachers will want to experiment first within their own class.

The assessors and assessees may be of the same ability, or deliberately of different ability. If they are of the same ability, you can expect a rich dialogue between them. If they are of different ability, the flow may be more one way, with the more able child dominating the proceedings.

The amount of background experience the students have in PA can be very variable. PA may represent a considerable challenge to, and generate considerable resistance in, new initiates. If they have previous experience, it may have been positive, negative or both. So, bear in mind the previous experience that these students might have had in previous classes. You might want to ask them about that right at the beginning.

Students from different cultural backgrounds may be very different in acceptance of PA. In particular, students from a Middle Eastern or Asian background may have great difficulty accepting PA. In the case of Middle Eastern students, resistance might have a lot to do with gender, as boys might be very reluctant to accept advice from a girl. In the case of Asian students, the idea that there is not one right answer which the teacher already knows can be rather startling, and also lead to resistances.

Gender may thus make a difference, and thought should be given to the implications of same-sex or cross-sex matching. With Middle Eastern students, same-sex matching might be easier to start with. We have some evidence from peer tutoring that same-sex matching is generally more effective for boys, but of course that leaves you with the question of what to do with the girls. So, there is no easy answer here. Of course, if there is no face-to-face contact (as in an online environment), gender may not be apparent.

Place can vary: most PA is structured and can occur in class, but it can also be informal and occur outside of class. Once students become really involved in it, you may find they are having PA conversations in break time. Indeed, in some cases, taking PA into their homes and using it with older and younger siblings.

Similar variation occurs with respect to the time when the PA takes place: How long are the sessions, how many sessions? Generally, the morning is best for thinking activities, but maybe PA could also fit into the afternoon when the timetable perhaps feels a little looser. If a big and complicated piece of work is being peer assessed, a good deal of time might be needed, but this should be broken into smaller sections of no longer than one period, and some structure provided so that students do not go off track. Make sure you give enough time so that the PA is actually finished in the tine specified.

What degree of justification for opinions is expected of the assessor? In the beginning it will be hard enough to get peer assessors to give suggestions for improvement, without expecting them to say why they think what they think. But with experience, peer assessors may become more adept at this—and also be more careful about not giving an opinion until they are sure they can justify it.

Will all PAs be confidential to the assessing pair and the teacher, or will they be made publicly available? At the start you will want to keep the PAs confidential to each assessing group. Once you have checked some of them for reliability, and you are satisfied about reliability, you may wish to operate a more open system. This could of course become competitive, and you would not wish what you had hoped would be a positive social experience degenerate into a competition.

Another issue is the extent to which the process of PA is monitored by supervisory staff. With PA in one class, it is relatively easy for the teacher to keep an eye on the situation. But PA between classes can become tricky in terms of keeping an eye on the situation. Obviously, you will want to be alert to any problems and able to nip them in the bud.

The extent to which the reliability and validity of the PA is moderated by supervising teachers is also an issue. While this generally comes up mainly with summative quantitative PA, it can also be relevant where students are giving elaborated verbal feedback. Sometimes this feedback may seem so strange that you are tempted to intervene—but remember, it is for the assessee to comment first, so give them the chance to say that the PA is nonsense.

Inspecting a sample of the assessments is particularly important where the assessment is summative. Is the task a simple surface task requiring limited cognitive engagement, or a highly complex task requiring considerable inference of the part of assesses, or does a simple initial task develop into increasingly complex tasks? If it is complex, you might be particularly inclined to pay some attention to the process.

In relation to this, what quantity and quality of feedback is expected, and is this elaborated and specific, or more concise and general? Time will be a major factor here. Initially, you might want to ask your assessors to give two positive points of feedback and two points where improvement might help. Should this latter be about a minute point (such as a spelling) or much broader (such as the structure of a piece of writing), or do you want to say that one should be broad but the other can be small?

To what extent is the feedback tending toward the objective and definitive, as it might be in response to a simple task, or to what extent more subjective, as it might be with a more complex task? What effect might this have on the amount of disputation that ensues? Is there time for the assessees to actually make all the suggested improvements?

How are assessees expected to respond to feedback? Are their revisions to be none, few or many, simple or complex? Again, given the time constraints, you may wish to put some sort of quota on this—perhaps a maximum of three revisions to be done in 20 min, or some such.

What extrinsic or intrinsic rewards are made available for participants? The USA has been much criticised for its use of extrinsic rewards. First, it is worth thinking about what the students might get out of PA in intrinsic terms. Once over their first shock, do the assessors get more pride in what they are doing, more involvement as they engage their

assessee(s) in conversation, and so on? Do assessees seem to respond at all to the deeper and quicker suggestions for improvement they get from a peer assessor (as compared to a teacher)? Might this activity become self-sustaining without it having to be inflicted on the students? Can you reflect this back to the students?

Another issue is whether the PA is aligned with the traditional forms of assessment. Will the PA be taken into account when grading students at year end, for example, or does all of the assessment information for this have to be generated by the teacher? Do all students have to sit formal examinations irrespective? If so, is there any way you can use PA to help them prepare for these examinations?

What transferable skills relevant to other activities might be measured as by-products of the process? Are you seeing improved social or communicative skills which might generalise beyond the PA situation? Or writing skills or presentation skills? Or music skills or physical education skills? Might any of these endure beyond school or university? These are important by-products which should be taken into account when you are considering the success or otherwise of your PA project.

Finally, is the PA being evaluated, as one would hope with any new venture, or is its success or failure just assumed? Time spent evaluating is costly, and could be spent doing something else, but if you are to persuade the powers that be (within your school/university or wider than that) that PA is worthwhile, you are going to need some evidence that looks at least a little bit objective.

Thus, it is clear that PA is not just one method, but many. Labels can be given to some of these variations, distinguishing formative from summative PA, qualitative from quantitative, structured from unstructured, unidirectional from reciprocal or mutual, same-year from cross-year, and same-ability from cross-ability PA, for instance.

Using Table 1, teachers will be able to decide and see what kind of PA they intend to implement. Importantly, because all the variations are listed, teachers will not overlook any issue they should have considered. There are rather a large number of variables in the table, and some researchers have proposed clustering these. The difficulty is that different researchers propose different clusters, so I have left the list un-clustered. Now let us consider theory in PA, and see how we can relate that to the typology just explicated.

6. Theoretical Issues in PA

PA theory is rather scarce. An early contribution in 1998 [19] regarding peer learning in general introduced the idea of distributed cognition leading to distributed metacognition. Subsequently a conceptual framework for PA in teacher education was articulated [20]. Later, researchers used expectancy theory regarding students' motivation for PA, emphasising the belief that performance would lead to valued outcomes [21]. The cognitive underpinnings of PA were explored in 2010 [22]. In 2016, a model was advanced [23] describing how PA operated in marking/grading, analysis, feedback, conferencing and revision, noting that investigating learning opportunities was more useful than investigating student/instructor grade relationships. More recently, researchers [24] have explored the theoretical underpinnings of PA in a digital world.

However, a more comprehensive and integrated theoretical model of the cognitive processes involved in PA has been proposed [18] (chapter 4, pp. 103–109), encompassing: organisation; cognitive conflict; individualisation and engagement; scaffolding and error management; communication; affect; intersubjectivity; practice and generalisation; reinforcement; metacognition, self-regulation and self-efficacy; and levels of learning. This illuminates many of the processes which may occur during PA, either deliberately or accidentally. This model has since been simplified and further developed, as in Figure 1.
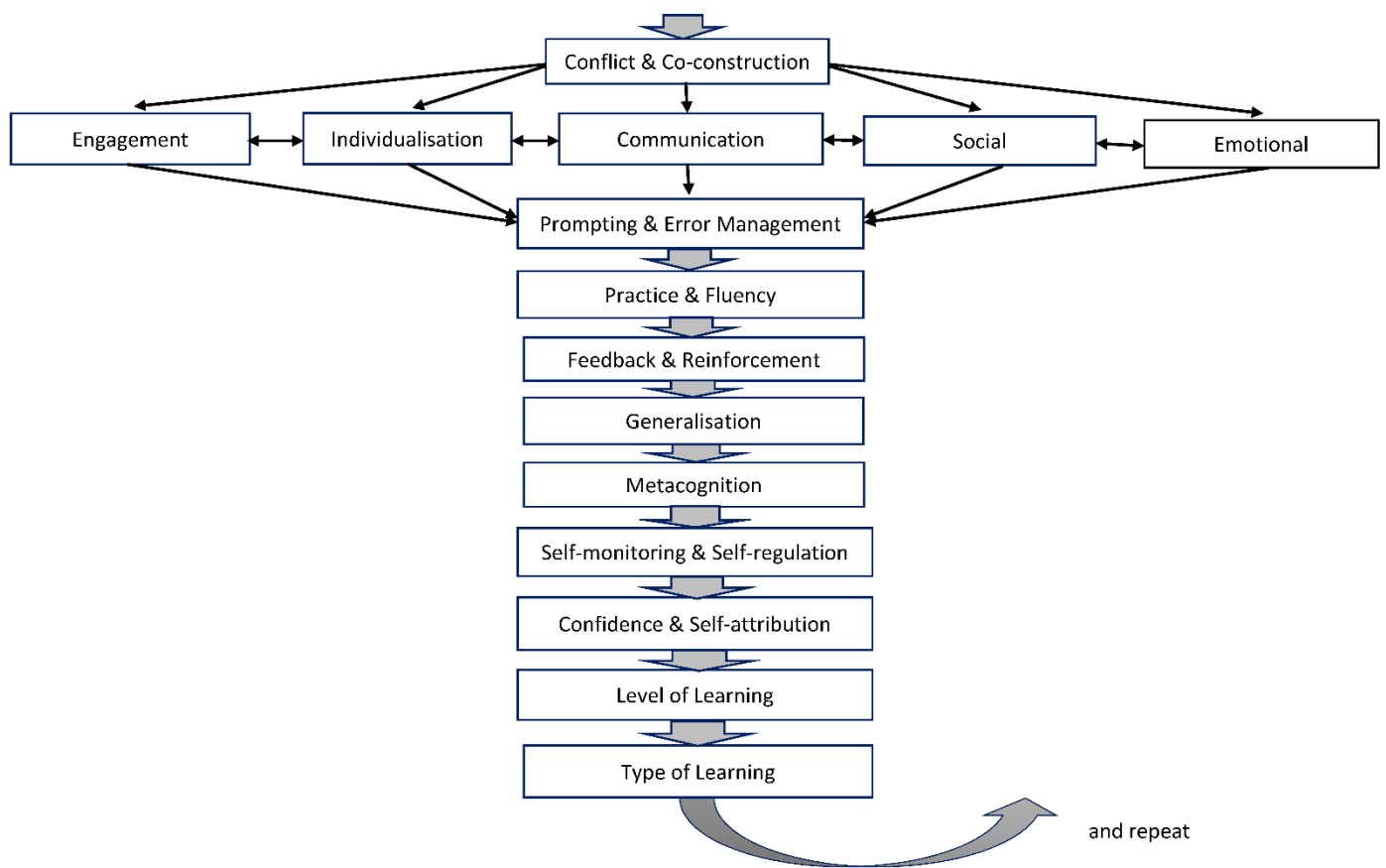
Figure 1. Theoretical Model of Peer Assessment.

Of course, not all partnerships will show all these features when they are first developing. Some may not show many features even when somewhat developed. The purpose of the theoretical model is to enable partners (perhaps with professional help) to see what new functional areas their relationship might develop into next. A more elaborate relationship is likely to be more satisfying for both assessor and assessee, and lead to enhanced educational outcomes. The model also gives professionals a framework within which they can counsel partners towards more effective experiences. Thus, it has strong practical implications for improvement of PA quality. A most important point is that both partners can be expected to benefit in all these ways—both as assessor and as assessee.

The model also has implications for research. The design of new PA interventions to be evaluated could be tested against this model, to ensure all aspects had been considered. The other question is which of these elements might be the most effective in any particular context. Research could possibly investigate the relative efficacy of each part of the model, while holding the other parts constant, but this would only be relevant to the context in which the PA was occurring. It may be that the whole proves greater than the sum of its parts.

6.1. Organisation

Many of the issues regarding organisation will have become clear from reading of the foregoing typology section—what kind of PA is proposed and what planning decisions do you need to make [25]? What organising time for PA will do is enable participants to get together and focus on the task(s) in hand. This has an effect on attention and concentration. Issues include the need and pressure inherent in PA toward increased time looking at the task and maybe thinking about it (time on task) and time observably involved in doing something active leading to task completion (time engaged with task)—the two being different concepts. The need for both helper and helped to elaborate goals and plans, the

immediacy of feedback possible within the small group or one-on-one situation, and the variety of a novel kind of learning interaction are also included in this category.

The issues of gender and race can also present problems. First, gender—should you pair with the same gender or mix genders, or does it not matter? There is no one right answer here. The main issue may be how to engage males, since often more females volunteer for such activities. Nonetheless, the presence of a male figure is particularly important for boys, but may also be important for girls.

Second, race—should you try to pair participants of the same race, or does it not matter? One problem here is determining exactly how the participants are located in terms of race, which might depend on how recently they have arrived in the host country (this might not be an issue for participants who have generations of experience of the host culture and speak the host language as a native would). However, many recent immigrants might call themselves citizens of their new country, but their culture and beliefs might still owe much to the country of origin of themselves or their parents. This may lead to issues of acceptance. Further, even within one country there are often a great many cultural and religious differences. So, you cannot assume that because both participants come from the same country, they will be well matched—indeed, sometimes quite the opposite.

6.2. Cognitive Conflict and Co-Construction

From Organisation, we proceed to more abstract and psychological variables. Conflict and Co-Construction are very much part of informal learning. Conflict is a clash of opposite opinions, which need to be worked through and resolution found. Co-construction is collaborating with others in building knowledge together—jointly investigating, analysing, interpreting and reorganising. Both are needed to liquify primitive cognitions and beliefs.

When the pair first meet, they will need to talk to decide their first area of inquiry. Then they will need to find out where each other is in their area of inquiry. What they will discover is not only that the knowledge of both is somewhat patchy, but that the assessee (and maybe the assessor!) holds some ideas very dear which are not helpful—in fact, they are wrong, or at best unduly simplistic. What will follow is a somewhat heated conversation where the pair try to determine a consensus on what they both already know about the subject which is actually correct. This is known as a period of "cognitive conflict"—disagreement about thinking.

Once the pair have established this baseline, they are in much better shape to proceed to build correct knowledge which is new for the assessee (and maybe for the assessor). However, this will be done gradually, and result in the assessee (and maybe the assessor) re-tuning their existing knowledge into something more complex and refined, adding new elements to it in a way that coheres rationally with what is agreed to be already known, or perhaps even restructuring existing knowledge to accommodate the new knowledge. This kind of "cognitive co-construction" by mutual agreement leads to a state known as "intersubjectivity" or shared understanding (for this area of inquiry) between the pair. Intersubjectivity is the sharing of subjective states by two or more individuals—they agree on a given set of meanings or a definition of the situation [26].

The notion of cognitive conflict reflects Piagetian schools of thought [27]. It concerns the need to loosen cognitive blockages formed from old myths and false beliefs by presenting conflict and challenge via one or more peers. Teachers focus on learning as if the pupil was a blank slate. But, in fact, the pupil's head is full of all kinds of stuff, much of it factually or conceptually erroneous. So, unlearning wrong stuff is as important as learning new stuff. Peers can be good at rooting out misconceptions in their partner—they certainly have more time for it than the teacher does.

The Russian psychologist Vygotsky [28] was famous for investigating cognitive co-construction between more able and a less able participants. He found that it was important that the level of challenge was appropriate for the assessee—within their "zone of proximal development" (the level where the assessee could not perform unaided but could perform successfully with some help from a more knowledgeable other).

From Conflict and Co-Construction, there are five different options, all of which interact with each other and have an influence on the linear steps which follow them (see Figure 1) [29]. We will take these five variables in the order in which they appear.

7. Engagement

Engagement describes intensity of arousal and involvement with the task. It encompasses curiosity, interest, attention, responsiveness, investigation, discovery, anticipation, persistence and initiation [30]. Any activity which is of interest to the pair will result in a focus of attention on the joint interactive task (and pairs should not try to engage with activities which are only of interest to one member of the pair). There will be concentration and arousal gains. Of course, if one member of the pair becomes too much like a teacher (didactic—maybe even bossy), the concentration and response of the assessee may suffer. So, some form of equal sharing of the interactivity is needed, which can be helped if the assessor is not an "expert" in the field (or is pretending not to be).

A great advantage is the immediacy of response from one to another [31]—especially high in face-to-face contact, albeit rather slower in messaging at different times (asynchronously) via the internet. This keeps the interaction speeding along at a good pace, even if there are diversions where the members of the pair do not agree and a compromise has to be negotiated. As the relationship develops, pairs are able to make goals and plans for the future about issues they will explore in future meetings [32]. Of course, there will be lots of talking, so any hope that PA will be quiet is unrealistic—there will be noise—but of course it will be productive noise and it is unlikely to disturb the pairs.

8. Individualisation

An immediate benefit of PA to both members of the pair is that both are receiving more than usual individual attention, intended to be specifically relevant to their immediate concerns. This might mean that the assessee gets more attention than in a regular class, while the assessor also gets more attention than in the course of usual everyday events, and in both cases this attention is closely focused on the mental activity of the other person, i.e., it avoids other distractions, is highly engaging, and requires new thinking. Nonetheless, more individual attention would rapidly lose its appeal if it had no active content.

An associated advantage is Individualisation—the content, pedagogy and pace of learning are based upon the unique abilities and interest of each learner—and perhaps their culture, socioeconomic status, language, gender, motivation, ability/disability, personal interests and so on (this is also known as Differentiation). Each member of the pair will increasingly respond to their partner in a way which is tailored to the needs of that partner. As time goes by, the assessor will modify the difficulty and other characteristics of the material under discussion so that the individual assessee can readily understand it—although this will take some time to develop [33,34]. Of course, the partner should not be "dumbing down" the issue too much so it is too easy—a certain amount of challenge is always needed.

Various forms of interactivity will take place. There will be many opportunities to question—from both members of the pair. A question is any sentence which has an interrogative form or function. Assessor questions act as instructional stimuli suggesting elements to be learned. Young children are often very good at asking questions, especially if they are encouraged—although sometimes their questions are too big to find an answer [35]. Equally, the assessor can question strategically—not offering just a closed question or one where the answer is self-evident, but asking a question which leads the student on from where their thinking has got to. Learning skilful questioning is highly desirable in assessors—and of course assessees will learn it and use it with their eventual assessees in later years. A good question promotes a high quality of answer—not just "yes" or "no", but an elaborated statement which indicates the reasoning behind the student's opinion. Of course, the opinion may be quite wrong, and the partner then has to skilfully question to get the student to see alternative perspectives.

8.1. Communication

Much of PA is about communication—the act or process of using words, sounds, signs, or behaviours to express or exchange information or to express your ideas, thoughts, feelings, etc., to someone else. Listening, explaining, questioning, summarising, speculating and hypothesising are all valuable skills of effective PA which should be transferable to other contexts. PA pairs will communicate in the common vocabulary of everyday people, not in the rather technical and complex language teachers sometimes use. This enables children to be much more talkative than they might otherwise be. Vygotsky [27] said that you only really know something when you have the language to express it to another person, and PA gives students the chance to develop the language to express their thoughts—including their deepest thoughts, which might quite surprise their partner. Both parties also need to carefully listen to the other as they attempt to explain their point of view, then ask questions which lead to further elaborations—or maybe a realisation that the first view was wrong or incomplete.

Of course, there needs to be care that a given explanation is not too abstract for the assessee to grasp. Exemplification can be very helpful here—a concrete example often works wonders. Students often make their initial stumbling explanations too long-winded and partners can help them by encouraging them to clarify, simplify or summarise. Summarising teaches assessees how to discern the most important ideas in a text, how to ignore irrelevant information, how to integrate the central ideas in a meaningful way and improves their memory.

Some students will be reluctant to offer half-formed thoughts, and the partner will encourage them to say something, because everything can be revised and improved later once you have something to start with. Similarly, the idea of rehearsing an idea should be shared (not just repeating it but adapting it at each stage), so that with continuous improvement it will eventually be worth sharing with other pairs or the whole group [36]. As an idea develops pairs can speculate freely or hypothesise, allowing their imagination to run riot, then later bring their ideas back and rationalise or summarise them for wider consumption. Needless to say, this process presents many learning opportunities for the assessor as well.

8.2. Social

Every learning interaction requires the use of social skills by both members of the pair [9]. Social skills are the skills we use to communicate our messages, thoughts and feelings and interact with each other, both verbally and non-verbally, through gestures, body language and our personal appearance. At a more advanced level, such skills include empathy and self-control. If they do not already know each other, at first meeting both assessor and assessee might need some way of introducing themselves and beginning to talk about what might be learned first. If need be, they can be given some training and a list of tips about this. The notional assessor will need to learn not to be bossy and not to talk too much of the time—in other words, not be too much like a professional teacher. The assessee will need not to be over-powered by their partner and be prepared to expose their initially rather faulty thinking, as well as accepting both criticism and praise without becoming upset or over-excited. Both members of the pair will need to show some social tolerance of the peculiarities of their partner. Of course, social skills developed with one partner will only partly transfer to interaction with a new partner. Apart from these functional issues, the participants should develop a sense of social connectedness and trust in each other.

9. Emotion (Affect)

Emotion has a particularly strong influence on selectivity of attention, as well as motivating action and behaviour. A trusting relationship with a peer who holds no position of authority might facilitate self-disclosure of ignorance and misconception, enabling subsequent diagnosis and correction that could not occur otherwise. Modelling of enthusiasm and competence and belief in the possibility of success by the helper can influence the

self-confidence of the helped, while a sense of loyalty and accountability to each other can help to keep the pair motivated and on-task.

Negative emotions such as anxiety, depression, anger and frustration can be the cause or effect of problems with learning and lead to a maladaptive and self-defeating pattern of behaviour which prevents learning [37]. At first meeting, a degree of anxiety is normal. Both partners are entering a new situation, which is unknown. As the pair get to know each other better and learn to trust each other (bearing in mind the assessor is not the same kind of authority figure as a teacher), their anxiety about each other should reduce and their self-esteem (or self-confidence) should grow. Of course, for some pairs, there might be a longer period of social as well as cognitive conflict before things settle down.

In the longer run, other emotional factors come into play. The assessee might be anxious about the material to be assessed. Here it will be important that the assessor is positive and encouraging and reassures the assessee that they felt the same way before they learned it, but now they are quite happy and confident with it. In other words, the assessor should be encouraging and demonstrate a model of coping and confidence. As the assessee becomes more confident, they will feel more able to disclose their thinking, which may well be faulty, and this will enable diagnosis and correction by their partner.

As time goes on, both members of the pair should develop more certainty about what is being assessed, and with that will come higher desire and confidence (motivation) to proceed to the next thing [38]. Added to this is the fact that the partners come to be accountable to each other—because they have a better and better relationship, they do not want to let their partner down. This gives them a stronger sense of responsibility for their learning. This responsibility leads to a stronger sense of ownership of their own learning—it is truly theirs rather than being inflicted upon them by an outside organisation.

10. Prompting (Scaffolding) and Error Management

Once assessees have the confidence to express their thinking out loud, it will become evident that they are making errors, or perhaps leaving gaps in their line of reasoning. How should assessors intervene? Particularly when one partner is more able in the area of interest than the other, they are likely to be involved in "prompting"—saying something to encourage or remind someone to do or say something, without telling them what they have to say [39]. Prompting is definitely not just telling them the "right answer"—if assessors do this, they are paying too much attention to correctness and not enough to the development of the thought processes required for the assessee to arrive at the right answer by themselves. Of course, the latter takes longer, but the assessee learns the thinking involved and can then use these skills to solve other similar problems.

"Scaffolding" is another word sometimes used in this context [40]. When grasping the concept is just too difficult for the assessee, the assessor provides some steps which lead the assessee in the right direction—without giving the answer. Like prompting, this is a skill that assessors have to develop over time—another of the benefits for them.

Error management is directed at dealing effectively with errors after they have occurred, with the goal of minimising negative and maximising positive error consequences. One of the major issues is the question of how errors should be corrected [41]. Even when the assessee's error seems glaring to the assessor, the assessee may be very emotionally attached to it, so it is no use just saying that is "wrong". The first issue is identifying the error—sometimes the assessor will miss errors without noticing or at first may choose to concentrate on major errors and overlook minor errors. When the assessor spots an error, they should not immediately go into a mini-lecture about it. Instead, they should wait till the end of the sentence then simply point to or say what the error was, and see how the assessee responds—they may be able to self-correct, which is a much more productive way of progressing [42].

The other issue is diagnosing the kind of error—what does it tell us about the assessee's faulty thinking and what might we need to address to resolve that faulty thinking? It follows from what has been said above that errors need to be discussed between the

partners, so they can arrive at a newly constructed form of truth before going on. If the assessee still cannot grasp the concept, the assessor may have to resort to giving a more concrete example or modelling or demonstrating how that bit of the problem can be solved. Again, skill development for the assessor. One of the great advantages is that errors can be corrected almost immediately. In a classroom, students might have to wait much longer, unless they were using some kind of computer application which offered corrective feedback.

Generally, errors should be corrected in a positive way through discussion, prompting, scaffolding and if necessary, modelling—demonstrating the relevant behaviour [43]. Assessors should also remember that once the error has been identified, they should pause or allow some "wait time" to allow the assessee to try to self-correct [44]. With a bit more thinking they might manage it on their own, and that would make for better learning than too much interference by the assessor. However, particularly with very difficult concepts, the assessor will need to monitor and control the flow of information so that the assessee is never presented with too large a chunk of material which they cannot assimilate. The concept of zone of proximal development is again highly relevant. The cognitive demands upon the assessor in terms of monitoring learner performance and detecting, diagnosing, correcting and otherwise managing misconceptions and errors are great. Herein lies much of the cognitive exercise and benefit for the helper.

The greater the differential in ability or experience between the assessor and the assessee, the less cognitive conflict and the more scaffolding might be expected. Too great a differential might result in minimal cognitive engagement for the assessor and unthinking but encapsulated acceptance with no co-construction. Of course, if the assessor is older, more experienced, and therefore more credible but actually has no greater correct knowledge or ability than the helped, then a mismatch and faulty learning might occur in a different way.

11. Practice and Fluency

PA enables and facilitates a greater volume of engaged and successful practice, leading to consolidation, fluency and automaticity of thinking, and social, communicative and other core skills. Much of this might occur implicitly, i.e., without the assessee or assessor being fully aware of what is happening.

PA might occur more frequently than interaction between the teacher and each student in the classroom. So, there are more opportunities to repeat similar tasks until the principles are really well understood. This also enables and facilitates a greater volume of engaged and successful practice—the actual application or use of an idea, belief, or method, as opposed to theories relating to it. Of course, the practice needs to be correct practice, or the assessee will overlearn mistakes! [45].

This more frequent practice leads to greater consolidation and fluency in understanding and performance. Someone is said to be fluent if their use of the language appears fluid, smooth, natural, coherent, and easy. Fluency is characterised by the language user's automaticity, their speed and coherency of language use, and the length and rate of their speech output. The flow is smoother because some of the learning has become automatic—it does not have to be consciously remembered but is put into operation without really thinking about it. The more learning is at the automatic level, the greater will be the retention of that learning—it is truly embedded in the assessee's consciousness. Much of this automaticity is implicit, i.e., the assessee is not really consciously aware of it.

12. Feedback and Reinforcement

Another great benefit of PA is feedback—information about their performance given to learners to praise positive aspects and point out areas needing improvement—which of course then needs to be acted upon. PA increases the quantity and immediacy of feedback to the learner very substantially. Feedback can also help develop the leaner's capacity to monitor, evaluate and regulate their own learning [46]. Feedback from assessors is more

frequent than with classroom learning [47]. Assessees can be frequently encouraged as they struggle with difficult concepts. As they say things that are partially right, the assessor can start by pointing out what they have said which is good and useful, then move on to point out where their reasoning is less good (always positive before negative) [48].

Positive reinforcement is the action or process of encouraging or strengthening a pattern of behaviour by associating some positive event with the behaviour, so it is more likely to occur again in the future. Usually positive reinforcement will be praise, but this should clearly specify exactly what is being praised. The role of praise is an interesting issue. One might say that all students should be praised as much as possible. However, even some young children are not happy with an excess of praise, perhaps because they feel they have to learn trust in the person who is praising before they can accept it [49]. Praise may not be appropriate in that context at that time. However, it can be helpful if assessees can be encouraged to give praise to assessors also, since the feedback process should be two-way.

Where praise is given a variety of forms of verbal praise are need, not just a routine and repetitive "good". In addition, the praise needs to be accompanied by non-verbal signals, so that the assessee is convinced that the assessor actually means it. Giving a variety of both verbal praise and non-verbal praise is a skill that has to be developed, and this extends the assessor's repertoire. Beyond the partnership, there may be explicit reinforcement for the pair in the form of social acknowledgement and status, official accreditation, or even more tangible reward. However, tangible reward which is not necessary is not likely to act as a reinforcer.

Some of this feedback and reinforcement will be implicit (the partners not consciously aware of it), but some will be explicit (the partners are consciously aware of it). However, indiscriminate reinforcement which is not linked directly with good performance or is predominantly for effort rather than performance will not be nearly as effective in promoting good learning. Explicit reinforcement might stem from within the partnership or beyond it, by way of verbal and/or non-verbal praise, social acknowledgement and status, official accreditation, or even more tangible reward. However, reinforcement should not be indiscriminate or predominantly focused on effort.

13. Generalisation

Generalisation accepts that humans recognise the similarities in knowledge acquired in one circumstance and that this enables transfer of that knowledge into new and somewhat different situations. Once the assessee has really learned a concept, they can begin to apply it to other similar problems. An obvious example would be in mathematics, where once a principle is grasped, it can be applied to many similar problems. PA can lead to generalisation from the specific example in which a concept is learned, extending the ability to apply that concept and its developmental variants to an ever-widening range of alternative and varied contexts [50]. In the first instance, much of this would be supported by the assessor, but as time goes on, it should become increasingly independent—the assessee managing this without much scaffolding. Likewise, in the first instance, it would be implicit, but as time goes on and the assessee is made aware of what is happening, it should become increasingly explicit.

14. Metacognition

Metacognition is awareness and understanding of one's own thought processes—thinking about thinking—which leads to the ability to control and direct those thought processes (see self-regulation in the next section). In a learning situation, it means becoming sharply aware of how you are thinking to learn, and consequently how that thinking can be made more efficient [51]. Assessors will usually become more metacognitively aware first, then the assessee may then follow. Metacognition is always explicit—it is always fully in consciousness and is intentional. It can be summarised in the catch phrase: I know I know; I know I know how; I know I know when and if.

As the learning relationship develops, both assessor and assessee should become more consciously aware of what is happening in their learning interaction, and more able to monitor and regulate the effectiveness of their own learning strategies in different contexts. Development into fully conscious explicit and strategic metacognition not only promotes more effective onward learning, it should make the assessor and assessee more confident that they can achieve even more, and that their success is the result of their own efforts. In other words, they attribute success to themselves, not external factors, and their self-efficacy is heightened.

15. Self-Monitoring and Self-Regulation

As learners become more sophisticated, they become more metacognitively aware and, through this, more able to self-monitor their own thinking. Self-monitoring can be defined as the process of attending to one's own actions and noting or recording the presence or absence of a specified relevant behaviour [52]. This of course requires multi-tasking—not only thinking, but also thinking about thinking. So, it is not easy. Beyond this, the learner should become more able to self-regulate or control their thinking about similar and then new topics in different contexts, so that many false paths are avoided and the logical consistency of their reasoning improves [53]. Self-regulated learning refers to one's ability to understand and control one's learning environment and includes goal setting, self-instruction, and self-reinforcement. This self-regulation can be both implicit and explicit.

16. Confidence (Self-Efficacy) and Self-Attribution

As the learner develops metacognition and self-regulation, an emotional change is likely to occur. Because the assessee is so much more aware of their thinking and in control of it, they feel increasingly confident about their mastery of this area of inquiry. Confidence means believing in your own ability, skills and experience and your ability to succeed. Of course, some students are over-confident, but many are under-confident. As assessors become increasingly competent, they also become increasingly confident. Confidence is also known as self-efficacy [54,55].

Furthermore, the students attribute this improvement to their own ability, rather than to the support of the assessor or any even more distant external factors. Self-attribution bias refers to an individual's tendency to attribute successes to their own personal skills and any failures to factors beyond their control [56]. Assessors attribute success to themselves as well as the efforts of the assessee. This self-attribution can be summarised in the catch phrase: I want to know; I want to know how, when, if; I believe I can know how, when, if.

17. Level of Learning

Surface learners have an unreflective approach—there is a focus on memorising and reproducing the learning material, knowledge is fragmented, facts are not elaborated upon and there is no real interaction with or connection between with ideas. The underlying argument is not comprehended and the learning task is treated as a monotonous chore. The learning is driven by external incentives or punishments, such as an impending test, i.e., is extrinsic. The aim is to recite and regurgitate the material inactively, forgetting it as soon as the external accountability requirement has passed [57].

By contrast, deep learners relate the topic and its ideas to past knowledge and experiences. They think critically about newly learned material and tie it in with information from other sources. They recognise a structure in the content. Their motivation comes from within and is intrinsic—they want to learn. They aim to understand the meaning behind the material and can create new arguments based on the new information. They retain much of what they learn.

Obviously, the aim is to enhance deep learning and reduce surface learning. However, all learners may need to engage at the level of surface learning before they can develop into deep learners, in relation to any particular topic of inquiry. The role of the assessor in

encouraging deep learning is cognitively challenging for them, and enhances their own level of thinking. As the PA relationship develops, the model continues to apply as the learning moves from the shallow, instrumental, surface level to the strategic level and on to the deep level as the students pursue their own goals rather than merely those set for them.

18. Type of Learning

Learners need to possess and be aware of three kinds of knowledge: declarative, procedural, and conditional. Declarative knowledge is factual information that one knows; it can be declared—spoken or written. Procedural knowledge is knowledge of how to do something, formed by doing, of how to perform the steps in a process; for example, knowing how to pronounce a multi-syllabic word. Conditional knowledge is about when to use a procedure, skill, or strategy and when not to use it—why a procedure works and under what conditions; and why one procedure is better than another. For example, learners need to know under what conditions to draw a diagram to more effectively illustrate points that they are making. In PA, all these kinds of learning are needed [58]. However, the usual tendency is to over-emphasise declarative knowledge at the expense of the two other kinds, so this needs to be struggled against.

These affective and cognitive outcomes feed back into the originating subprocesses—a continuous, iterative process and a virtuous circle. Of course, it is unlikely that PA in practice will neatly follow these linear stages. Some may be missing (and the teacher can prompt for their insertion). Sometimes one will occur before another which appears to follow it in the model. Most likely a number of events will occur which seem to be combinations of items. Even where students work through the whole model on one task, they may begin again at the outset on a new task.

19. Conclusions

Different individuals within the same learning partnership, and with different partner relationships, are likely to follow somewhat different pathways to the same learning goals. If one characteristic of the assessors and assessees is that they are developmentally young or slow learners themselves, then few of the channels in the model will develop automatically, intersubjectivity is likely to be primitive, and more training and closer monitoring, coaching and management will be necessary. Although all channels in the model might be eventually utilised to some extent by both members of a pair, their different learning styles might lead them to use some channels more than others in ways unique to themselves. This highlights the individualisation which is inherent in PA, but takes the notion much further than the mere individualisation of learning tasks or surface learning behaviours.

The point of the model is to enable learners (whether assessees or assessors) to see what channels they are currently not using enough or not using at all and encourage them to use additional channels as suits their personal learning styles to maximise the effectiveness of their learning. For professionals, this theoretical model is something of a mixed blessing. Just when they thought they knew how PA should work, along comes a model that makes everything seem rather more complicated. Of course, professionals should be encouraged to think of the model in terms of a step-wise progression for each pair. Having identified which elements any PA pair are not doing, the professional selects the one most obviously missing and desirable element and advises the pair to engage in it. Later, she/he selects the next most obviously missing element, and so on . . . So, professionals are never actually faced with trying at one moment to get the PA pair to engage in all the elements, which is likely to be too complex and counter-productive.

In that respect, a useful task for professionals is to explain the model to users in simple terms, discuss how it applies to present learning and how future learning might take advantage of additional opportunities. In this respect, it may become a feature in initial training in PA—or perhaps in a second phase of training after some initial experience. It can provide a framework for helping the learning partners themselves to reflect upon their own process—a tool for self-assessment or PA of the process which might further

enhance metacognition [18]. The model can also be used profitably as a template (or observational checklist) for monitoring PA as it is happening—a tool to structure monitoring and diagnostic fault finding.

For future research, the template provided by the model should prove useful for the design of new PA methods. Further research might seek to explore the validity of the model empirically, or of the relative effectiveness of different elements of the model with different learners. Research into the impact of the use of the model in monitoring implementation integrity (quality of delivery of an intervention) also would be worthwhile.

References

1. Brooks, C.; Carroll, A.; Gillies, R.M.; Hattie, J. University of Queensland; University of Melbourne A Matrix of Feedback. *Aust. J. Teach. Educ.* **2019**, *44*, 14–32. [CrossRef]
2. O'Donnell, A.M.; Topping, K.J. Peers assessing peers: Possibilities and problems. In *Peer-Assisted Learning*; Topping, K., Ehly, S., Eds.; Lawrence Erlbaum: Mahwah, NJ, USA, 1998.
3. Yu, F.-Y. Multiple peer-assessment modes to augment online student question-generation processes. *Comput. Educ.* **2011**, *56*, 484–494. [CrossRef]
4. Li, L.; Liu, X.Y.; Zhou, Y.C. Give and take: A re-analysis of assessor and assessee's roles in technology-facilitated peer assessment. *Br. J. Educ. Technol.* **2012**, *43*, 376–384. [CrossRef]
5. Topping, K.J. Peer assessment between students in college and university. *Rev. Educ. Res.* **1998**, *68*, 249–276.
6. Falchikov, N.; Goldfinch, J. Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Rev. Educ. Res.* **2000**, *70*, 287–322. [CrossRef]
7. Li, H.; Xiong, Y.; Hunter, C.V.; Guo, X.; Tywoniw, R. Does peer assessment promote student learning? A meta-analysis. *Assess. Eval. High. Educ.* **2019**, *45*, 193–211. [CrossRef]
8. Double, K.S.; McGrane, J.A.; Hopfenbeck, T.N. The Impact of Peer Assessment on Academic Performance: A Meta-analysis of Control Group Studies. *Educ. Psychol. Rev.* **2019**, *32*, 481–509. [CrossRef]
9. Van Gennip, N.A.E.; Segers, M.; Tillema, H.M. Peer assessment for learning from a social perspective: The influence of inter-personal variables and structural features. *Educ. Res. Rev.* **2009**, *4*, 41–54. [CrossRef]
10. Tillema, H.; Leenknecht, M.; Segers, M. Assessing assessment quality: Criteria for quality assurance in design of (peer) assessment for learning–A review of research studies. *Stud. Educ. Eval.* **2011**, *37*, 25–34. [CrossRef]
11. Hoogeveen, M.; Van Gelderen, A. What Works in Writing With Peer Response? A Review of Intervention Studies with Children and Adolescents. *Educ. Psychol. Rev.* **2013**, *25*, 473–502. [CrossRef]
12. Li, H.; Xiong, Y.; Zang, X.; Kornhaber, M.L.; Lyu, Y.; Chung, K.S.; Suen, H.K. Peer assessment in the digital age: A meta-analysis comparing peer and teacher ratings. *Assess. Eval. High. Educ.* **2015**, *41*, 245–264. [CrossRef]
13. Johnson, S. On the reliability of high-stakes teacher assessment. *Res. Pap. Educ.* **2013**, *28*, 91–105. [CrossRef]
14. Tenório, T.; Bittencourt, I.I.; Isotani, S.; Da Silva, A.P. Does peer assessment in on-line learning environments work? A systematic review of the literature. *Comput. Hum. Behav.* **2016**, *64*, 94–107. [CrossRef]
15. Fu, Q.K.; Lin, C.J.; Hwang, G.J. Research trends and applications of technology-supported peer assessment: A re-view of selected journal publications from 2007 to 2006. *J. Comput. Educ.* **2019**, *6*, 191–213. [CrossRef]
16. Zheng, L.; Zhang, X.; Cui, P. The role of technology-facilitated peer assessment and supporting strategies: A meta-analysis. *Assess. Eval. High. Educ.* **2020**, *45*, 372–386. [CrossRef]
17. Gielen, S.; Dochy, F.; Onghena, P. An inventory of peer assessment diversity. *Assess. Eval. High. Educ.* **2011**, *36*, 137–155. [CrossRef]
18. Topping, K.J. *Using Peer Assessment to Inspire Reflection and Learning*; Student Assessment for Educators Series; MacMillan, J.H., Ed.; Routledge: New York, NY, USA; London, UK, 2018; ISBN 978-0-8153-6765-9 (pbk).
19. King, A. Transactive Peer Tutoring: Distributing Cognition and Metacognition. *Educ. Psychol. Rev.* **1998**, *10*, 57–74. [CrossRef]
20. Sluijsmans, D.; Prins, F. A conceptual framework for integrating peer assessment in teacher education. *Stud. Educ. Eval.* **2006**, *32*, 6–22. [CrossRef]
21. Friedman, B.A.; Cox, P.L.; Maher, L.E. An Expectancy Theory Motivation Approach to Peer Assessment. *J. Manag. Educ.* **2007**, *32*, 580–612. [CrossRef]

22. Kollar, I.; Fischer, F. Peer assessment as collaborative learning: A cognitive perspective. *Learn. Instr.* **2010**, *20*, 344–348. [CrossRef]
23. Reinholz, D. The assessment cycle: A model for learning through peer assessment. *Assess. Eval. High. Educ.* **2015**, *41*, 301–315. [CrossRef]
24. Tai, J.; Adachi, C. The future of self and peer assessment: Are technology or people the key? In *Re-Imagining University Assessment in a Digital World. The Enabling Power of As-Sessment*; Bearman, M., Dawson, P., Ajjawi, R., Tai, J., Boud, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; Volume 7.
25. Topping, K.J. Peer assessment: Learning by judging and discussing the work of other learners. *J. Interdiscip. Educ. Psychol.* **2017**, *1*, 7. [CrossRef]
26. Zlatev, J.; Racine, T.P.; Sinh, C.; Itkonen, E. *The Shared Mind: Perspectives of Intersubjectivity*; Ben-jamins: Amsterdam, The Netherlands, 2008.
27. Weber, C.O.; Piaget, J.; Warden, M. The Language and Thought of the Child. *Am. J. Psychol.* **1927**, *38*, 299. [CrossRef]
28. Vygotsky, L.S. *Mind in Society: The Development of Higher Psychological Processes*; Harvard University Press: Cambridge, MA, USA, 1978.
29. Gagne, R.M.; Wager, W.W.; Golas, K.C.; Keller, J.M. *Principles of Instructional Design*, 5th ed.; Wadsworth: Belmont, CA, USA, 2004.
30. Engagement for Learning. *The Engagement for Learning Framework Guide*; Department for Education: London, UK, 2011.
31. Witt, P. *Communication and Learning*; De Gruyter Moutom: Berlin, Germany, 2016.
32. Rutherford, P. *Active Learning and Engagement Strategies: Teaching and Learning in the 21st Century*; Just ASK Publications: Alexandria, VA, USA, 2012.
33. Yeh, S. Understanding and addressing the achievement gap through individualized instruction and formative assessment. *Assess. Educ. Princ. Policy Pract.* **2010**, *17*, 169–182. [CrossRef]
34. Joseph, S.; Thomas, M.; Simonette, G.; Ramsook, L. The Impact of Differentiated Instruction in a Teacher Education Setting: Successes and Challenges. *Int. J. High. Educ.* **2013**, *2*, 28. [CrossRef]
35. King, A. Facilitating Elaborative Learning Through Guided Student-Generated Questioning. *Educ. Psychol.* **1992**, *27*, 111–126. [CrossRef]
36. Horinouchi, T.; Wakita, S.; Anse, M.; Tabe, T. A Study of an Effective Rehearsal Method in e-Learning. In *Constructive Side-Channel Analysis and Secure Design*; Springer International Publishing: Berlin/Heidelberg, Germany, 2007; Volume 4558, pp. 328–336.
37. Tyng, C.M.; Amin, H.U.; Saad, M.N.M.; Malik, A.S. The Influences of Emotion on Learning and Memory. *Front. Psychol.* **2017**, *8*, 1454. [CrossRef]
38. Heckhausen, J.; Heckhausen, H. *Motivation and Action*, 3rd ed.; Springer: New York, NY, USA, 2018.
39. Sitzmann, T.; Ely, K. Sometimes you need a reminder: The effects of prompting self-regulation on regulatory processes, learning, and attrition. *J. Appl. Psychol.* **2010**, *95*, 132–144. [CrossRef]
40. Gibbons, P. *Scaffolding Language, Scaffolding Learning*, 2nd ed.; Teaching English Language Learners in the Mainstream Classroom; Heinemann: Portsmouth, NH, USA, 2014.
41. Frese, M.; Keith, N. Action Errors, Error Management, and Learning in Organizations. *Annu. Rev. Psychol.* **2015**, *66*, 661–687. [CrossRef] [PubMed]
42. Ramdass, D.; Zimmerman, B.J. Effects of Self-Correction Strategy Training on Middle School Students' Self-Efficacy, Self-Evaluation, and Mathematics Division Learning. *J. Adv. Acad.* **2008**, *20*, 18–41. [CrossRef]
43. Haston, W. Teacher modeling as an effective teaching strategy. *Music Educ. J.* **2007**, *93*, 26–30. [CrossRef]
44. Forbes, S.; Poparad, M.A.; McBride, M. To err is human; To self-correct is to learn. *Read. Teach.* **2004**, *57*, 566–572.
45. Allington, R.L. *What Really Matters in Fluency: Research-Based Practices Across the Curriculum*; Pearson: New York, NY, USA; London, UK, 2008.
46. Nicol, D. From monologue to dialogue: Improving written feedback processes in mass higher education. *Assess. Eval. High. Educ.* **2010**, *35*, 501–517. [CrossRef]
47. Gielen, S.; Peeters, E.; Dochy, F.; Onghena, P.; Struyven, K. Improving the effectiveness of peer feedback for learning. *Learn. Instr.* **2010**, *20*, 304–315. [CrossRef]
48. Hattie, J.; Clarke, S. *Visible Learning: Feedback*; Routledge: New York, NY, USA; London, UK, 2018.
49. Dweck, C.S. The perils and promises of praise. *Educ. Leadersh.* **2007**, *65*, 34–39.
50. Polit, D.F.; Beck, C.T. Generalization in quantitative and qualitative research: Myths and strategies. *Int. J. Nurs. Stud.* **2010**, *47*, 1451–1458. [CrossRef] [PubMed]
51. Hacker, D.J.; Dunlosky, J.; Graesser, A.C. *Metacognition in Educational Theory and Practice*; Routledge: New York, NY, USA; London, UK, 1998.
52. Joseph, L.M.; Eveleigh, E.L. A Review of the Effects of Self-Monitoring on Reading Performance of Students with Disabilities. *J. Spéc. Educ.* **2009**, *45*, 43–53. [CrossRef]
53. Vohs, K.D.; Baumeister, R.F. *Handbook of Self-Regulation*, 3rd ed.; Research, Theory, and Applications; The Guilford Press: New York, NY, USA, 2017.
54. Zimmerman, B.J. Self-Efficacy: An Essential Motive to Learn. *Contemp. Educ. Psychol.* **2000**, *25*, 82–91. [CrossRef]

55. Schunk, D.H.; Zimmerman, B.J. *Motivation and Self-Regulated Learning: Theory, Research, and Applications*; Routledge: New York, NY, USA; London, UK, 2012.
56. Booth, M.Z.; Gerard, J.M. Self-esteem and academic achievement: A comparative study of adolescent students in England and the United States. *Comp. A J. Comp. Int. Educ.* **2011**, *41*, 629–648. [CrossRef] [PubMed]
57. Lindblom-Ylänne, S.; Parpala, A.; Postareff, L. What constitutes the surface approach to learning in the light of new empirical evidence? *Stud. High. Educ.* **2018**, *44*, 2183–2195. [CrossRef]
58. Ormrod, J. *Human Learning*, 7th ed.; Pearson: New York, NY, USA; London, UK, 2015.