education sciences

MDPI

Article

Advantages and Disadvantages of Online and Face-to-Face Peer Learning in Higher Education: A Review

Keith James Topping

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

Abstract: During the pandemic, many institutions shifted to online teaching, and in some cases, this included existing peer learning programs. As the pandemic receded, some of these peer learning programs returned to face-to-face operation and others adopted a blended format, while others remained online. Interestingly, the literature suggests that online peer learning is somewhat more effective than face-to-face peer learning. This might be because online peer learning enables responses at any time (which might be more thoughtful), anonymity, and a wider nexus of relationships, although it can create issues regarding the initial development of trust. There are a great many studies of both face-to-face and online peer learning, but relatively few that directly compare both. By way of addressing this gap, this paper aims to systematically review 17 papers that directly compare both, informed by and updating the only previous review in this area. Online performs better than offline learning in terms of cognitive outcomes, with a small to moderate effect size. However, the associated socio-emotional issues are more complex. Online learning offers flexibility regarding response time, but sacrifices the immediate dialogue of offline learning. Some cultures found accepting peer learning more difficult. Few studies undertook longer-term follow-up, although with more practice motivation, this might well improve. The results have implications for the type of peer learning chosen by pedagogical designers as appropriate to their own learning context.

Keywords: peer assisted learning; peer tutoring; peer assessment; online; offline

check for updates

Citation: Topping, K.J. Advantages and Disadvantages of Online and Face-to-Face Peer Learning in Higher Education: A Review. *Educ. Sci.* **2023**, *13*, 326. https://doi.org/10.3390/educsci13040326

Academic Editors: David Arendale and James Albright

Received: 2 February 2023
Revised: 10 March 2023
Accepted: 21 March 2023
Published: 23 March 2023

Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

What is peer learning? A definition of peer assisted learning (PAL) which is still widely quoted is: "PAL is the acquisition of knowledge and skill through active helping and supporting among status equals or matched companions. PAL is people from similar social groupings, who are not professional teachers, helping each other to learn and by so doing, learning themselves" [1] (p. 1). PAL includes peer tutoring, peer modelling, peer education, peer counselling, peer monitoring, and peer assessment, both reciprocal and non-reciprocal, in schools and institutions of higher education, as well as in the workplace. This definition can also clearly include all forms of cooperative learning, in which students work together in small groups, sometimes with the specification of different roles. However, this special issue includes many other contributors discussing cooperative learning, so in this paper, the focus is largely on peer tutoring and (particularly) peer assessment.

Peer tutoring has been demonstrated to be effective over many years, with a great number of reviews, systematic analyses, and meta-analyses supporting this, even before 1996 [2] and continuing through 2019 [3]. There are recent special editions regarding 51 studies of peer tutoring in music education [4], a systematic review of 16 reviews and meta-analyses on peer tutoring with students with behavioural problems [5] and a systematic review purely on online PAL [6].

Parallel, but somewhat later, interest in peer assessment has also spawned a large number of reviews and meta-analyses, especially in recent years (e.g., [7–19]). However, none of these have specifically explored the differences between online and offline learning.

Additionally, there have been many recent reviews of online and blended learning, particularly discussing 'emergency remote learning', such as that which occurred during the pandemic, which was often completely online. As we emerge from the pandemic, the nature of future pedagogy is being more widely discussed. There are many reviews asserting that online learning is at least as effective as face-to-face learning (e.g., [20]). There are also systematic analyses revealing that blended (a mixture of face-to-face and online) is even more effective than purely online learning (e.g., [21]).

While there is a great number of papers on online PAL, unfortunately, there are almost no reviews which directly compare online learning, blended learning, and face-to-face learning, at the same time and in the same context as peer tutoring/peer assessment. Even simultaneously comparing online with face-to-face learning proves difficult enough. This paper aims to address this gap by systematically reviewing papers that directly compare online learning with face-to-face learning in the context of peer tutoring or peer assessment, and is informed by and updates the only previous review in this area. During the preparation of this paper, there were initially no reviews, but one meta-analysis appeared by virtue of diligent searching. However, systematic analyses and meta-analyses are very much determined by the search terms used and the databases targeted for this purpose. Changes in either can substantially affect the outcomes, despite such searches typically encompassing hundreds of potential papers.

The present paper thus updates the paper by Jongsma et al. in 2022 titled, "Online versus Offline Peer Feedback in Higher Education: A Meta-Analysis" [22]. Peer feedback is a common element in many PAL projects. The search terms used in this paper were: 'peer assessment' OR 'peer feedback' OR 'peer review' OR 'peer evaluation' OR 'peer rating' OR 'peer scoring' OR 'peer grading' AND 'learning outcome' OR 'learning achievement' OR 'achievement' OR 'outcome' OR 'learning performance' OR 'academic achievement' OR 'academic performance'. 'Peer tutoring' and other forms of PAL were not included, and the inclusion of outcome/achievement keywords seems likely to have limited the search.

Jongsma borrowed five papers from another author, then selected only five papers of her own, so lack of coverage is a possibility. From a parallel search (discussed below), I found thirteen other papers up to 2020 (when Jongsma et al.'s search ended) which seemed relevant. Additionally, I go beyond 2020 to consider four relevant papers which appeared in 2021–2023. Then, I offer some discussion of socio-emotional issues stemming from PAL, since such supplementary gains paralleling cognitive gains may exist, and socio-emotional factors may partially determine the longevity of any cognitive effects (longer-term follow-up is also largely absent from the literature).

2. Methodology

This paper studies formal learning in higher education, including peer interaction for all students to enhance learning in the pursuit of higher academic achievement. As peer interaction often occurs outside the classroom, it can be somewhat difficult to monitor. This paper is not about informal learning, which typically is engaged in more by some students than others, and not at all by some, and is never monitored or directly assessed; thus, it is excluded here.

In this systematic review, two research questions were identified:

1. Which research studies on peer tutoring, assessment, and feedback directly compare the effectiveness of online and offline teaching and learning in the same study?
2. Is there evidence of effectiveness, and if so, what proportion of this research is solely dependent on student and teacher perceptions, and how much of it uses other indicators?

What search terms were employed in this attempt to parallel Jongsma's systematic analysis? First, I tried: "peer learning" OR "peer assessment" OR "peer feedback" OR "peer review" OR "peer evaluation" OR "peer rating" OR "peer scoring" OR "peer grading" AND online AND offline OR face-to-face OR "face to face". There were no date restrictions. Five databases were searched: Web of Science, Scopus, JSTOR, ERIC, and Google Scholar

(these were quite different from Jongsma, whose meta-analysis came up in none of these searches). These keywords did not yield many hits. Seeking to obtain more hits, I tried these more general terms: "peer learning" AND "peer assessment" AND online versus offline. These broader and fewer keywords yielded more hits. Given that fewer and broader keywords yielded more hits, I then tried: "peer learning" AND offline versus online AND review OR analysis. Then I tried: "peer learning" OR "peer tutoring" OR "peer assessment", which of course, generated many hits and required considerable inspection of titles and abstracts.

The inclusion criteria were that the paper needed to: (1) directly compare online and offline learning, simultaneously and with the same course content, (2) include some kind of data (but not necessarily be experimental), and (3) relate to higher education (college or university). The four search strategies in five databases yielded a total of 724 hits (excluding replications). Google Scholar yielded the highest number and JSTOR the lowest. Obviously, the more generic search terms yielded many more hits, which required more time to inspect, but most of these putative hits proved irrelevant. Eventually, I selected only 13 papers from 2020 and before, and four papers from 2020–2023. Not all of these were experimental, and their quality varied significantly.

Papers were reported in relation to whether they were experimental or not. This might be taken as implying that experimental papers are always of superior research design than non-experimental papers, but that is not the implication here. Rather, the results here are so divided to enable ease of relating to the purely experimental results of Jongsma et al. (2022) [22]. Critique of the quality of each paper will be found in the results.

3. Results: Cognitive Outcomes: Papers up to 2020

Thirteen papers were found which appeared to compare online to face-to-face peer learning up to 2020, but in relation to Jongsma et al. (2022) [22] only four of these were experimental studies which might have appeared in their meta-analysis. A further four had interesting research designs which were somewhat rigorous, but which would not have appeared in to Jongsma et al. A further five were not remotely experimental, but were nonetheless interesting. We will take these in reverse order.

3.1. Not at All Experimental

A survey with medical students at several universities in China was undertaken [23], self-evaluating learning effectiveness, learning efficiency, learning atmosphere, and other issues associated with online and offline learning. Most students expressed the view that offline learning was superior to online learning in terms of effectiveness, efficiency, and atmosphere. However, online learning was better in terms of the acquisition of learning resources and flexibility. The responses of undergraduates and postgraduates were largely similar. However, undergraduates valued offline learning significantly more than postgraduates. This study only explored perceptions, and we do not know how much experience these students actually had with different kinds of online learning. The cultural context of medicine as a subject as well, as the cultural context of China, might have affected the results.

In 2004, a method for peer assessment in the UK using web-based technology in an undergraduate computer programming course was described [24], and the advantages and disadvantages of the process were discussed. However, PAL was used to discuss test results provided automatically, and the focus was thus more instrumental than is usual. In addition, the work was graded by peer assessors, and the grades were discussed in groups of four. Student perceptions of the process were reported, which suffers from the same issue as those related above [23].

A novel application was described in StudioBRIDGE [25], an awareness system based on instant messaging, developed for students working in open studio spaces in the Architecture Department at the Massachusetts Institute of Technology. It was intended to help students initiate online and offline interactions by giving them an awareness of nearby

people, groups, locations, and events in the community. The authors believed that this awareness would lead to increased peer learning and expertise sharing by encouraging informal social communication. They describe the user community and the motivation, design, and initial pilot deployment of StudioBRIDGE, but say little about effectiveness, other than again reporting participant perceptions.

Twenty-two volunteer Japanese students of English as a Foreign Language in Canada on an exchange programme volunteered to participate [26] (out of 60 enrolled in the course). Online peer feedback pushed students to write balanced comments, with an awareness of audience needs, but with an anonymity that allowed them to make critical comments. Two face-to-face peer discussions were coupled with an online assessment of 500-word essays, and the assessees could confer with the assessors to receive clarification. The quality of online peer feedback was studied, along with comparisons of initial and revised drafts. Positive and negative comments were relatively balanced, and many negative comments included suggestions for improvement. A total of 13 of 22 students did revise their papers in light of the feedback. However, this is a small proportion of 60 course members, so generalisations are difficult. Again, only volunteers were studied, which may have introduced bias, and cultural issues may have intervened.

A mixing of offline and online feedback with an English for Academic Purposes programme at a UK university was reported in 2014 [27], but the number of participating students was not reported, although 358 written papers were assessed. A rubric was provided, and potential supportive comments were listed to promote positive feedback. The author explored both student and tutor perceptions drawn from focus groups, highlighting the respective strengths and weaknesses of both offline and online interactions. Students tended to focus on grammar, punctuation, spelling and vocabulary, rather than on broader aspects, such as style. Nonetheless, student engagement was boosted. The difficulty here again is that only student perceptions were explored, and we do not know the number of students participating.

3.2. More Rigorous Research Designs

A comparison of face-to-face (FtF) peer review and computer mediated peer review was conducted in an English as a Foreign Language academic writing context (n = 37 students, mostly female) [28]. Two classes were involved, and in the senior class, peer partners were assigned, while in the junior class, peers could choose their own partners. The senior class conducted most of their activities outside of class, while the junior class mainly participated in class. Both of these issues may have affected the results, which aggregated the data from both classes. Online feedback largely involved using Track Changes in Word, which is actually relatively uncommon in peer feedback research. A questionnaire of 30 items with a 5-point Likert scale was used to explore student feelings about both peer review methods, but reliability was generally low, and the FtF section had particularly low reliability. Additionally, adjectivally labelled Likert responses were analysed with parametric statistics, which is highly questionable. Moreover, the paper did not include the questionnaire, which was not helpful.

Nonetheless, the students agreed that peer review was very useful. However, participants preferred the FtF method, partially because they preferred speaking to writing but also because FtF enabled them to talk in their native tongue rather than English. Online review offered more flexibility, and 67% of the learners liked using Track Changes, but only 46% said they wanted to continue online. However, stress from confronting another's mistakes was an issue (possibly a culturally specific issue), and most learners reported that they felt more comfortable and less social pressure online.

A study with 24 undergraduates in Taiwan examined how a combination of three peer learning modes (FtF, synchronous online, asynchronous online methods) influenced students' peer review in second-language writing courses [29]. FtF peer review took place in class, and online peer review occurred outside of class, which may have affected the results. There was an initial peer review online (asynchronous), a second was FtF, and a third

was online (synchronous). Varieties of peer feedback were demonstrated by the teacher. Data came from audio files and transcripts recorded in FtF sessions, online logs, drafts of writing tasks, two questionnaires, and interviews with nine students. The affordances of the interplay of the three modes influenced students' task engagement and perception of peer review. Satisfying individual preferences for modes might be important, but of course, these preferences were likely to change over time. Equally, arranging various modes appropriately at different stages of drafting might maximize the effects of peer review.

A comparison and contrast of gifted children's use of critical thinking in a summer talented writer's workshop was performed in two situations [30]: while giving peer feedback using a structured paper guide and while giving peer feedback using online social commenting on blogging and digital storytelling websites. Of 34 children, only 10 volunteered to participate. Online, students wrote daily Kidblog entries on current events in their journalism class and wrote picture books using Storybird.com. These very different forms of writing were combined under the social media or online category. Offline, they rewrote fiction stories in their creative writing class, assessed by a rubric and categorized as offline. The outcome data were the children's peer feedback. Ten student writings were analysed for the degree of critical thinking. Inter-rater reliability was quite low, but the nature of the task was complex. Critical thinking was more evident in offline responses structured by the rubric compared to those in social media contexts, which tended to feature informal language which was often vague and displayed little critical thinking.

Another author [31] also focused on students of English as a Foreign Language, studying only 13 students, again in Taiwan. All participants were required to write four multi-draft expository essays on topics given to them throughout the semester. Peer review training was given to the participants, a rubric was used, and sample essays were assessed as practice. Two papers were peer reviewed FtF and two online, but second drafts were all followed by teacher feedback, which might have confused the results. There were three sources of data: first, all drafts, written and electronic comments, transcripts of face-to-face discussions, and online chat logs; second, a questionnaire used to elicit student perceptions; and third, follow up interviews. The peer review mode did affect some types of peer comments, to a certain extent. There were significantly more global alteration comments and fewer local alteration comments in the FtF than in the online mode. While the participants preferred detailed online comments over handwritten comments, they felt face-to-face discussions were more effective regarding issues that could not be easily replaced by electronic chat (e.g., immediacy and paralinguistic features).

3.3. Experimental Studies

Turning to the four experimental studies, in the first, a traditional FtF instruction group (n = 35) was compared with a computer mediated group (n = 35) in the peer assessment of writing in Iran [32]. All participants were male. The groups were taught by the same teacher, and English tests demonstrated that they were equivalent at the onset. Students who did not show the required level of digital literacy were assigned to the control group. The treatment consisted of two sessions per week over eight weeks. Training was given to the experimental group. The TOEFL writing post-test of 40 items included structure questions (13 items) and written expression questions (27 items). The computer-mediated group did highly significantly better than the FtF group. In this study, the groups were relatively small and the intervention period brief. The assignment of less digitally literate students to the control group indicates a deviation from random assignment, which is understandable, but not desirable.

Pre-service teachers were engaged in a peer assessment activity in two classes [33], a distance learning and an FtF class. Although there was no control group, there were two clearly separate alternative treatments, but it is not clear whether the two groups were equal at the outset. Students' work samples were collected to examine feedback function ratings and participants completed an online questionnaire. For both of these, there was

no statistically significant difference between the FtF and online groups. However, the provision of feedback was significantly better in the FtF class.

A somewhat similar study [34] investigated peer assessment, peer feedback quality, and grading accuracy among 339 engineering students in a project-based course. Participants studied the same course, but in three different modes: on-campus (n = 77), online (n = 110), and in a massive open online course (MOOC) (n = 152). A content analysis of feedback comments identified four categories: reinforcement, statement, verification, and elaboration. The results showed that the MOOC participants provided more feedback comments and volunteered to assess more projects than participants in the other groups. However, campus students provided higher quality feedback and their peer grading correlated better with grades assigned by the teaching assistants.

Other authors studied 71 mostly female English as a Second Language students mainly from Korea [35], using peer assessment of writing in offline and online modes. Online students not only engaged in peer review, but completed quiz activities on exemplary errors in an online café, so the treatments were not at all similar. Student writings were examined as outcome measures. Online students did significantly better than offline and control groups in writing performance in both the short- and long-term. When teaching intermediate level students, if a teacher's concern was for global errors, which would eventually impede communication, the order of sentence type > main verb > usage > word choice > redundancy > word order > logic seemed optimum. Logic was the most difficult area of acquisition, with its own unique characteristics.

4. Results: The Jongsma Meta-Analysis of Online/Offline Peer Feedback

Meta-analysis is sometimes held up as the "gold standard" of evidence synthesis, but it does possess many problems, not least of which is the exclusion of an enormous proportion of the papers focusing on a phenomenon in the cause of research design rigour, which sometimes leaves a ridiculously small number of papers as the focus of the analysis. Of the various critical reports, some authors, for example [36] (p. 2), comment that: "Most meta-analyses include too few randomised participants to obtain sufficient statistical power and allow reliable assessment of even large anticipated intervention effects. The credibility of statistically significant meta-analyses with too few participants is poor, and intervention effects are often spuriously overestimated (type I errors) or spuriously underestimated (type II errors). Meta-analyses have many false positive and false negative results."

However, Jongsma et al. (2022) [22] are to be congratulated on publishing a meta-analysis which conforms to many of the official requirements for such an undertaking. Here, we will critique their work, but that is not to denigrate the enormous contribution they have made. These authors point out that online peer assessment offers the possibility of anonymity, can save teacher time, and more easily allows the teacher to monitor the peer feedback comments (assuming these are written). Disadvantages include the fact that dialogue can be difficult in an asynchronous context. In feedback dialogue, students have the opportunity to receive feedback on the feedback they have given, clarifying or negotiating the meaning of the received feedback.

Jongsma et al. (2022) [22] only studied university students. They drew five studies from Zheng et al. (2020) [16], and their own search resulted in the addition of five more, so the total number of studies was small. Their search terms have been critiqued earlier. They only searched two databases (following [16]), a small number given the variability between databases. Meta-analyses are often critiqued for combining studies that are fundamentally unalike, and this applied here as well: the sample group size ranged from n = 11 to n = 65, the subject domain was mainly the English language, but included graphic design and other studies which merely mentioned "teaching"; the assessed task was mainly a form of writing, but included graphic design elements and a reading test; in only one study was the peer assessment anonymous (despite the assertions noted above); training was mostly provided (although its intensity was unclear, and in two studies, even the access to training was unclear); the technology for the online element varied greatly, from online blogs, to

Google Docs, to videos, to "learning environments"; and assessments mostly comprised grades for performance, most often in writing, rather than more elaborated feedback. The duration of the studies varied from less than 5 weeks, to 5–10 weeks, and on up to more than 10 weeks.

On the positive side, the studies were assessed for risk of bias using the Cochrane Risk of Bias tool for randomized controlled trials (which few of the studies were). Only two of the ten studies found no effect, while the other eight all had at least one finding concluding that online learning was better. No study found that offline learning was better. The meta-analysis provided an overall effect size (Hedges) of g = 0.33, $p < 0.05$ for online peer feedback, which is somewhere between small and moderate.

5. Results: Cognitive Outcomes: Papers after 2020

All four of these papers appeared in 2021, but only one was experimental in design.

5.1. Non-Experimental Papers

A qualitative research method was used [37] with 58 second-year mostly female pre-service primary teachers in Turkey. They were divided into 11 groups, each consisting of five to six students, who all participated in both online and offline peer assessment. Peers had to grade their colleague's work on an eight-factor scale, with a maximum of three points per factor. The online peer assessment activity was anonymous, but of course, the FtF activity was not, so the treatments were not equivalent. Six students who worked in different groups were randomly selected in order to conduct semi-structured interviews, and these results were coupled with observation data. The inter-rater reliability was at least 80%. There was considerable student concern about the unfairness of the grading system. Anonymity fostered giving more objective grades. The study concluded that a combination of peer and instructor evaluation could provide better validity and objectivity of assessment.

In Indonesia, the impact of classroom peer feedback and asynchronous online communication via Facebook on the quality of students' writing revisions was compared [38] with 25 participating students, of whom 11 agreed to be interviewed. Training was given, and participants completed one activity FtF followed by another online (whether the reverse order would have had the same effects is an interesting question). Students' essay writing scores were analysed, along with a qualitative analysis of observations and interviews. Peer feedback through asynchronous online interactions was significantly more effective than that conducted offline. Six themes emerged: (1) peer feedback increased student autonomy, (2) teacher involvement was beneficial to improve consistency, (3) asynchronous peer feedback provided extra time, which yielded more elaborate feedback, (4) asynchronous peer feedback gave more opportunity to be audience members, (5) Facebook made the students feel awkward, and (6) recorded feedback via Facebook comments was more beneficial.

The validity of digital scoring of the Peer Assessment Rating (PAR) index compared with manual scoring was assessed in the field of orthodontics [39]. Just two operators scored the index on 15 cases at pre- and post-treatment stages using both methods. Measurements were repeated at a one-week interval. There were no significant differences in PAR scores between either methods or raters, and intra- and inter-rater reproducibility was excellent ($\geq$0.95), although both of these findings might have had something to do with the small rater sample size. PAR scoring on digital models showed good validity and reproducibility compared with manual scoring.

5.2. Experimental Paper

In 2021 as well, researchers worked with students in the USA [40], noting that it had been challenging to assess individual contributions to group projects, particularly in online courses. Their participants were in either online or offline classes, with peer feedback on individual contributions to group projects. Students were in groups and each student assessed all members of the group. The findings showed little difference between FtF and

online student perceptions, both perceiving that peer assessment was a good and reliable way to facilitate students' participations and contributions, and a reliable way to assess students' contributions to a group project.

6. Results: Socio-Emotional Outcomes

Five out of the ten studies included in the Jongsma et al. (2022) meta-analysis [22] also reported student perceptions of peer feedback, but one was very weak. Researchers surveyed and interviewed experimental students [41]. Results showed students felt positive about blogs, but were not sure about their confidence in giving and receiving peer feedback. However, they did not feel embarrassed when providing feedback. The use of blogs was easy and time-independent and contributed to the feeling of being a 'real writer'. In another study, students were surveyed about online peer feedback [42]. They said online peer feedback reduced their writing anxiety and gave them more time to think about how to comment on their peers' writing.

Students using Adaptive Comparative Judgment (a form of assessment using comparisons instead of criterion scoring) enjoyed the peer feedback process more, thought the process was easier to follow, and found the peer feedback more helpful than students in the control group using paper-based peer feedback [43]. In 2019, other researchers [44] concluded that students felt giving peer feedback was more helpful than receiving peer feedback, but without any difference between online and offline feedback. Online peer feedback was appreciated because its asynchronous nature gave more time before providing peers with feedback. However, these asynchronous discussions made direct exchanges between students difficult.

Turning to papers in addition to those from the Jongsma et al. (2022) meta-analysis [22] which contained information about socio-emotional issues, in 2004, other researchers [24] used an asynchronous anonymous system and noted that more students were more ready to ask questions and communicate in that context. However, some students gave only positive comments without offering solutions, and this was felt to be unhelpful by about half of the participants. Students also preferred a lengthy grading scale to enable them to be more discriminatory in their feedback. In 2007, other researchers [28] found offline sessions better in that immediate dialogue and the ability to discuss issues in the native language were possible, but worse in that sessions were rushed and participants often did not have enough time. By contrast, online sessions offered more flexibility, resulted in longer peer feedback, and actually resulted in less sense of social pressure, but the time delay presented problems for some.

In another study [38], all the students were actively engaged in the process of sharing comments through both direct and online interactions, and students became more self-sufficient in their learning. Most of the students could identify mistakes in their friends' drafts and make corrections related to those mistakes. Even through Facebook, students actively gave feedback and did it on time. However, training from the instructor was valuable in developing relevant skills. Asynchronous interactions allowed students more time to read and give comments to their peers' writing. The extra time gave the students the chance to read their peers' writing in detail and offer more complete corrections. As a result, online student feedback was more informative than offline feedback. Additionally, some students felt that class was noisy and rushed for time and that online activity enabled them to concentrate better—noisy and rushed classes could result in simpler and more basic feedback. Some students read more of the work of other students than they were required to, since they found it interesting and informative. However, online communication had its difficulties, not the least of which was the absence of non-verbal feedback. The recording of online comments enabled complete re-reading of the feedback as necessary, which was not possible offline.

In terms of the advantages of online review, in general, students seem positive about online peer feedback. In particular, its time-independence was flexible and convenient. If the class was noisy and rushed for time, the online activity helped them to concentrate

better. Online feedback allowed them to check resources and gave them more time to think about and phrase comments before providing feedback, thus resulting in clearer and more elaborated peer feedback. Online peer feedback could thus reduce anxiety surrounding giving peer feedback and result in less social pressure, and especially in an asynchronous anonymous system, more students were ready to ask questions and communicate. Students became more self-sufficient in learning and were more likely to respond on time. Online recording enabled complete re-reading of feedback, which was not possible offline. Online methods also enabled the review of work from multiple peers.

In terms of the advantages of offline, offline sessions were better in that immediate dialogue and the ability to discuss issues in the native language were possible—the time delay presented problems for some. Additionally, offline review included non-verbal feedback, which some students missed in the online mode. However, participating in a synchronous feedback dialogue could be difficult because of the imperative of thinking while talking. Some students still preferred feedback from a teacher, so students did not always trust peer feedback.

7. Discussion

In terms of cognitive outcomes, in general, online peer tutoring and assessment are more effective than offline methods, although some studies found them only equally effective. The substantial contribution of the Jongsma et al. (2022) meta-analysis [22] was critiqued, as were an additional five experimental studies, before and after 2020. However, direct comparisons of online and offline methods were relatively rare. Nonetheless, perhaps this is of relatively little importance if the future trend is to be towards the more effective blended, rather than purely online, learning. It will, of course, remain important for those courses which operate only remotely, such as MOOCs.

In terms of socio-emotional outcomes, online review was generally popular (although it may not be popular at the outset). Its flexibility, convenience, and facilitation of concentration were greatly valued, especially in an asynchronous mode, and led to more elaborated peer feedback, while offline methods were more rushed. Thus, online feedback reduced anxiety and involved less social pressure, so more students were ready to ask questions and communicate. Students became more self-sufficient, recording enabled re-reading of feedback, and students were more likely to respond on time. Online methods also facilitated the review of work from multiple peers. However, offline sessions enabled immediate dialogue, non-verbal feedback, and the ability to discuss issues in the native language, although thinking while talking could be challenging. Some students still preferred feedback from a teacher because they did not trust peer feedback.

There are a number of limitations to the research synthesised here. First, some studies relied entirely on student and/or teacher perceptions. While this is valuable, it is difficult to accept as the only form of data given its subjectivity, and one would wish to see it accompanied by other forms of more objective data. It would be helpful if future studies included objective outcome indicators as a valuable triangulation on student or teacher perceptions. Further, we neither searched for nor found any studies on additional behavioural outcomes from peer learning, such as intent to graduate or class participation, rather than on cognitive and socio-emotional issues.

It is also noteworthy that very few of the papers reviewed here mentioned long-term outcomes. Would students get better at peer feedback with more practice over time? Or would they become bored with it and want to revert to teacher assessment (which would require less of their energy and, at least in one way, be less stressful)? Additionally, one might ask if there was any spontaneous generalization to other subjects and courses. Do students become sufficiently engaged with peer tutoring and peer assessment to begin to do this informally in other courses, where neither is encouraged by the teacher? All of these are questions for future research, which should also seek to conduct a further systematic review and/or meta-analysis at some future date to see how the field has developed.

All of these issues should, of course, be taken up by authors offering guidance on how to conduct peer tutoring, peer assessment, and feedback. Recent examples from 2023 [45–47] are beneficial, although slightly older ones are also likely to be helpful (e.g., [48]).

The first study [45] specifically addresses peer assessment in online courses, having searched technology journals from 2010. Eight principles were proposed based on the literature reviewed: (1) provide training, (2) consider the impact of pair or group formation, (3) consider the pros and cons of anonymity, (4) combine peer grading and peer comments, (5) encourage assessors to address strengths and weaknesses and provide sufficient explanations, (6) use strategies such as scaffolding and monitoring to actively engage assessees, (7) encourage interactions between students, and (8) provide supportive structures. However, the author notes that almost all the papers studied focused on offline peer assessment, which raises questions about how his title can purport to be about online feedback.

An interesting instrument was offered in the second study [46], which reports the varying characteristics of peer assessment designs. A section on context requires details of: subject domain, time/place, setting, requirement, and alignment. A section on instructional design requires details of: purpose, object, product/output, relation to staff assessment, official weight, reward, directionality, degree of interactivity, frequency, group constellation, constellation assessor, constellation assessee, unit of assessment (assessor), unit of assessment (assessee), privacy, contact, matching, format, training, revision, and scope of involvement. A section on outcomes includes: beliefs and perceptions, emotions and motivation, performance, reliability, validity, feedback content, and feedback processing. A section on moderators/mediators includes: gender, age, ability, skills, and culture. Clearly this contains much detail, although again, how much of this is specific to online work is a moot point. Nonetheless, any teacher completing this checklist would become sharply aware of what issues were in danger of being overlooked.

Other researchers offered something broader [47], a scale for assessing students' peer feedback literacy in writing. They noted that previous literature indicated that feedback literacy involved four elements: recognition of the feedback's value and then responding to the feedback by revising, and making judgments (not only of their own work, but also of the received feedback) and managing affect (dealing with their feelings, emotions and attitudes). They also noted that acceptance of feedback varied over time and with experience. In parallel, other researchers [49] found that willingness to accept feedback was lower among students who only experienced peer feedback once, relative to both those students who had no prior experience and those who had more experience (although this study was limited to MOOCs).

The scale [47] had a particular focus on assessing gains made from giving feedback, as well as gains from receiving it. It was developed from the questionnaire responses of 474 Chinese undergraduates, equally balanced between the arts and sciences, recruited by convenience sampling from eight universities, so there may be issues of sample bias and cultural specificity. Thirty items were included, on a six-point Likert scale, coupled with five open-ended items. Four factors, based on 26 items, emerged, accounting for 62% of the total variance: feedback-related knowledge and abilities, cooperative learning ability, appreciation of peer feedback, and willingness to participate. Reliabilities varied between 0.80 and 0.89. A reduced scale of 20 items was then developed.

Follow-up comparisons showed that for two factors (willingness to participate and cooperative learning ability), students with more feedback experience had significantly higher scores than students with less. For feedback knowledge and abilities, only those students with much more peer feedback experience had significantly higher means. For peer feedback appreciation, moderate experience yielded the highest scores. However, the items in the questionnaire were all phrased positively, i.e., there was no alternation of positive and negative statements, which may have led to "yea-saying" (a positive bias), so future users may wish to adapt the questionnaire.

While these practical instruments are undoubtedly of value, the extent to which any of them engage fully with all the issues found in this review of online vs. offline peer assessment effectiveness is an interesting question, and this is equally true of ref. [48].

8. Conclusions

We can answer the research questions thus:

- Which research studies on peer tutoring, assessment, and feedback directly compare the effectiveness of online and offline teaching and learning in the same study?

In addition to the 10 studies in the Jongsma meta-analysis [22] of studies up to 2000, 17 additional studies were identified, 13 before 2000, and 4 after 2000. Overall, there was some evidence that online learning was better than offline learning—studies found that online was better or that online and offline were equal. None found that offline was better.

- Is there evidence of effectiveness, and if so, what proportion of this research is solely dependent on student and teacher perceptions, and how much of it uses other indicators?

A good deal of the evidence used only subjective perceptions, with the other portion using more objective measures or triangulating measures.

There were also questions about whether the Jongsma meta-analysis [22] used entirely appropriate search terms and databases. Other relevant research was found, both before and after this meta-analysis was conducted.

Online peer assisted learning (including peer tutoring and peer assessment) is at least as effective as offline PAL in cognitive terms, and probably modestly more effective in many contexts. From a socio-emotional perspective, online has more advantages than offline PAL, although both have some disadvantages. Students' response to PAL may be affected by cultural values and/or initial conservatism in the first instance, but experience of the benefits of online methods should lead them to a preference for online work. However, given that blended learning seems more effective than purely online learning, it is likely that online PAL will, in most cases, be an element of blended course delivery in the future.

Funding: This research received no external funding.

Data Availability Statement: The papers analysed here may be obtained from the author.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Topping, K.J.; Ehly, S. *Peer-Assisted Learning*; Lawrence Erlbaum: Mahwah, NJ, USA, 1998.
2. Topping, K.J. The Effectiveness of Peer Tutoring in Further and Higher Education: A Typology and Review of the Literature. *High. Educ.* **1996**, *32*, 321–345, (also in *Mentoring and Tutoring by Students*. Edited by S. Goodlad. Kogan Page: London & Stirling VA. 1998). [CrossRef]
3. Leung, K.C. An Updated Meta-Analysis on the Effect of Peer Tutoring on Tutors' Achievement. *Sch. Psychol. Int.* **2019**, *40*, 200–214. [CrossRef]
4. Fernández-Barros, A.; Duran, D.; Viladot, L. Peer Tutoring in Music Education: A Literature Review. *Int. J. Music. Educ.* **2022**, *41*, 129–140. [CrossRef]
5. Bowman-Perrott, L.; Ragan, K.; Boon, R.T.; Burke, M.D. Peer Tutoring Interventions for Students with or At-Risk for Emotional and Behavioral Disorders: A Systematic Review of Reviews. *Behav. Modif.* **2022**, *online first*. [CrossRef]
6. Tibingana-Ahimbisibwe, B.; Willis, S.; Catherall, S.; Butler, F.; Harrison, R. A Systematic Review of Peer-Assisted Learning in Fully Online Higher Education Distance Learning Programmes. *Open Learn.* **2022**, *37*, 251–272. [CrossRef]
7. Topping, K.J. Peer Assessment Between Students in College and University. *Rev. Educ. Res.* **1998**, *68*, 249–276. [CrossRef]
8. 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.* **2016**, *41*, 245–264. [CrossRef]
9. Chang, C.Y.; Lee, D.C.; Tang, K.Y.; Hwang, G.J. Effect Sizes and Research Directions of Peer Assessments: From an Integrated Perspective of Meta-Analysis and Co-Citation Network. *Comput. Educ.* **2021**, *164*, 104123. [CrossRef]
10. Li, H.; Xiong, Y.; Hunter, C.V.; Guo, X.; Tywoniw, R. Does Peer Assessment Promote Student Learning? A Meta-Analysis. *Assess. Eval. High. Educ.* **2020**, *45*, 193–211. [CrossRef]
11. Panadero, E.; Alqassab, M. An Empirical Review of Anonymity Effects in Peer Assessment, Peer Feedback, Peer Review, Peer Evaluation and Peer Grading. *Assess. Eval. High. Educ.* **2019**, *44*, 1253–1278. [CrossRef]

12. Tenório, T.; Bittencourt, I.I.; Isotani, S.; 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]
13. Cui, P.; Zheng, L. A meta-analysis of the peer evaluation effects on learning achievements in blended learning environment. In *Blended Learning: Enhancing Learning Success*; Cheung, S.K.S., Kwok, L., Kubota, K., Lee, L.K., Tokito, J., Eds.; Springer International Publishing: New York, NY, USA, 2018.
14. Huisman, B.; Saab, N.; van den Broek, P.; van Driel, J. The Impact of Formative Peer Feedback on Higher Education Students' Academic Writing: A Meta-Analysis. *Assess. Eval. High. Educ.* **2019**, *44*, 863–880. [CrossRef]
15. 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.* **2020**, *32*, 481–509. [CrossRef]
16. Zheng, L.Q.; Zhang, X.; Cui, P.P. The Role of Technology-Facilitated Peer Assessment and Supporting Strategies: A Meta-Analysis. *Assess. Eval. High. Educ.* **2020**, *45*, 372–386. [CrossRef]
17. Topping, K.J. Digital Hardware for Peer Assessment in K-12 Schools and Universities. *Front. Educ.* **2021**, *6*, 666538. [CrossRef]
18. Topping, K.J. Digital Peer Assessment in School Teacher Education and Development: A Systematic Review. *Res. Pap. Educ.* 2021, *online first*. [CrossRef]
19. Zhan, Y.; Wan, Z.H.; Sun, D. Online Formative Peer Feedback in Chinese Contexts at the Tertiary Level: A Critical Review on Its Design, Impacts and Influencing Factors. *Comput. Educ.* **2022**, *176*, 104341. [CrossRef]
20. Pei, L.S.; Wu, H.B. Does Online Learning Work Better Than Offline Learning in Undergraduate Medical Education? A Systematic Review and Meta-Analysis. *Med. Educ. Online* **2019**, *24*, 1666538. [CrossRef]
21. Topping, K.J.; Douglas, W.; Robertson, D.; Ferguson, N. The Effectiveness of Online and Blended Learning from Schools: A Systematic Review. *Rev. Educ.* **2022**, *10*, e3353. [CrossRef]
22. Jongsma, M.V.; Scholten, D.J.; van Muijlwijk-Koezen, J.E.; Meeter, M. Online Versus Offline Peer Feedback in Higher Education: A Meta-Analysis. *J. Educ. Comput. Res.* **2022**, 1–26. [CrossRef]
23. Jiao, Z.Q.; Yang, Y.J.; Zhang, S.; Xu, J. A Comparative Study About Attitudes Towards the Efficiency, Effectiveness, and Atmosphere of Offline and Online Learning Among Medical Students. *Ann. Transl. Med.* **2002**, *10*, 1270. [CrossRef]
24. Sitthiworachart, J.; Joy, M. Effective Peer Assessment for Learning Computer Programming. *ACM SIGCSE Bull.* **2004**, *36*, 122–126. [CrossRef]
25. Yee, S.; Park, K.S. StudioBRIDGE: Using group, location, and event information to bridge online and offline encounters for co-located learning groups. In Proceedings of the CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Portland, OR, USA, 2–7 April 2005; pp. 551–560. [CrossRef]
26. Guardado, M.; Shi, L. ESL Students' Experiences of Online Peer Feedback. *Comput. Compos.* **2007**, *24*, 443–461. [CrossRef]
27. Cloete, R. Blending Offline and Online Feedback on EAP Writing. *J. Teach. Engl. Specif. Acad. Purp.* **2014**, *2*, 559–571.
28. Ho, M.C.; Savignon, S.J. Face-To-Face and Computer-Mediated Peer Review in EFL Writing. *CALICO J.* **2007**, *24*, 269–290. [CrossRef]
29. Chang, C.F. Peer Review Via Three Modes in An EFL Writing Course. *Comput. Compos.* **2012**, *29*, 63–78. [CrossRef]
30. Miller, M.T.; Olthouse, J. Critical Thinking in Gifted Children's Offline and Online Peer Feedback. *Contemp. Educ. Technol.* **2013**, *4*, 66–80. [CrossRef]
31. Ho, M.C. The Effects of Face-To-Face and Computer-Mediated Peer Review on EFL Writers' Comments and Revisions. *Australas. J. Educ. Technol.* **2015**, *31*, 1–15. [CrossRef]
32. Mellati, M.; Khademi, M. Peer Evaluation in CMC Learning Environment and Writing Skill. *Int. J. Appl. Linguist. Engl. Lit.* **2014**, *3*, 220–228. [CrossRef]
33. Vásquez-Colina, M.D.; Russo, M.R.; Lieberman, M.; Morris, J.D. A Case Study of Using Peer Feedback in Face-To-Face and Distance Learning Classes Among Pre-Service Teachers. *J. Furth. High. Educ.* **2017**, *41*, 504–515. [CrossRef]
34. Usher, M.; Barak, M. Peer Assessment in a Project-Based Engineering Course: Comparing Between On-Campus and Online Learning Environments. *Assess. Eval. High. Educ.* **2018**, *43*, 745–759. [CrossRef]
35. Park, J.Y.; Park, C.W. Effects of Learning Environment Differences (On and Offline Vs. Off Line) on Intermediate Level Learners' Writing Development. *J. Engl. Teach. Through Movies Media* **2018**, *19*, 169–193. [CrossRef]
36. Wetterslev, J.; Jakobsen, J.C.; Gluud, C. Trial Sequential Analysis in systematic reviews with meta-analysis. *BMC Med. Res. Methodol.* **2017**, *17*, 39. [CrossRef] [PubMed]
37. Akçay, A.O.; Güven, U.; Karahan, E. The Views of Pre-Service Elementary Teachers About Online and Traditional Peer Assessment. *Int. J. Assess. Tools Educ.* **2021**, *8*, 409–422. [CrossRef]
38. Astrid, A.; Rukmini, D.; Fitriati, S.W. Syafryadin Experiencing Peer Feedback Activities with Teacher's Intervention Through Face-To-Face and Asynchronous Online Interaction: The Impact on Students' Writing Development and Perceptions. *J. Lang. Educ.* **2021**, *7*, 64–77. [CrossRef]
39. Gera, A.; Gera, S.; Dalstra, M.; Cattaneo, P.M.; Cornelis, M.A. Validity and Reproducibility of the Peer Assessment Rating Index Scored on Digital Models Using a Software Compared with Traditional Manual Scoring. *J. Clin. Med.* **2021**, *10*, 1646. [CrossRef] [PubMed]
40. Lee, S.J.; Kwon, K. Peer Assessment as a Facilitating and Assessment Strategy in Online and Face-To-Face Classes. *Int. J. Online Pedagog. Course Des.* **2021**, *11*, 36–48. [CrossRef]

41. Ciftci, H.; Kocoglu, Z. Effects of peer e-feedback on Turkish EFL students' writing performance. *J. Educ. Comput. Res.* **2012**, *46*, 61–84. [CrossRef]
42. Yang, Y.F. Transforming and Constructing Academic Knowledge Through Online Peer Feedback in Summary Writing. *Comput. Assist. Lang. Learn.* **2016**, *29*, 683–702. [CrossRef]
43. Bartholomew, S.R.; Zhang, L.; Garcia Bravo, E.; Strimel, G.J. A Tool for Formative Assessment and Learning in a Graphics Design Course: Adaptive Comparative Judgement. *Des. J.* **2019**, *22*, 73–95. [CrossRef]
44. Lee, M.; Evans, M. Investigating the Operating Mechanisms of the Sources of L2 Writing Self-Efficacy at The Stages of Giving and Receiving Peer Feedback. *Mod. Lang. J.* **2019**, *103*, 831–847. [CrossRef]
45. Yuan, J.M. Guidelines for Preparing For, Designing, and Implementing Peer Assessment in Online Courses. *Turk. Online J. Educ. Technol.* **2023**, *22*, 115–129.
46. Panadero, E.; Alqassab, M.; Fernández Ruiz, J.; Ocampo, J.C. A Systematic Review on Peer Assessment: Intrapersonal and Interpersonal Factors. *Assess. Eval. High. Educ.* **2023**, *in press*.
47. Dong, Z.; Gao, Y.; Schunn, C.D. Assessing Students' Peer Feedback Literacy in Writing: Scale Development and Validation. *Assess. Eval. High. Educ.* **2023**, *online first*. [CrossRef]
48. Topping, K.J. Using peer assessment to inspire reflection and learning. In *Student Assessment for Educators Series*; MacMillan, J.H., Ed.; Routledge: New York, NY, USA; London, UK, 2018.
49. Kasch, J.; Rosmalen, P.; Löhr, A.; Klemke, R.; Antonaci, A.; Kalz, M. Students' Perceptions of the Peer-Feedback Experience in MOOCs. *Distance Educ.* **2021**, *42*, 145–163. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.