The Effects of a Peer Feedback Intervention on Cognitive Group Awareness and Collaborative Outcomes in Computer-Supported Collaborative Learning

Researcher: L. G. J. Boksebeld

Supervisor: Dr P. Papadopoulos

Second Supervisor: Dr A. M. van Dijk

Master Thesis

March 2022, Enschede

Master Educational Science and Technology

Faculty of Behavioural, Management and Social Sciences

Table of Contents

Acknowledgement

First and foremost, I would like to thank my first supervisor, Dr Pantelis Papadopoulos, for his ongoing support, supervision, feedback, and enthusiasm during the entire process of my thesis. At times, I got the impression that he was even more excited than I was, which was very motivational. Secondly, I would like to thank my second supervisor, Dr Alieke M. van Dijk, for her support, fresh perspective, and extensive feedback on my work. Third, I would like to express my gratitude to the involved teachers and students from the Bachelor of Sport Sciences, for participating in my experiment and allowing me to do the experiment part of my thesis within their organisation. Lastly, I would like to thank my family and friends for being supportive and helping me stay motivated throughout my thesis.

Abstract

Computer-supported collaborative learning (CSCL) is seen as a promising approach in education since it could be effective, efficient, and enjoyable for learning. One requirement of CSCL is cognitive group awareness (CGA), which entails being aware of group members' knowledge and expertise (Janssen & Bodemer, 2013). However, students experience challenges with communication and coordination (Erkens et al., 2005; Janssen et al., 2007). This could lead to a lack of CGA. A potential way to improve CGA could be peer feedback since this has been proven effective in improving communication and coordination (Dominick et al., 1997; Phielix et al., 2011). Increased CGA could also positively affect the outcomes of the collaboration, such as individual achievement and group performance (Janssen & Bodemer, 2013). Therefore, the current study investigated the following research question: "Does peer feedback influence cognitive group awareness in computer-supported collaborative learning, and does (improved) cognitive group awareness influence the collaborative outcomes among students of higher education?". Based on previous research, it was hypothesized that peer feedback will positively influence CGA and that there is a positive correlation between CGA and collaborative outcomes among students of higher education. The current study had a quantitative, quasi-experimental research design. The sample consisted of 45 second-year students ranging in age from 18 to 24 years old (M = 20.22 years, SD = 1.80) of a Higher Education Institute in the Netherlands. An intervention was designed to enhance CGA by improving communication and coordination via peer feedback. CGA was measured via self-and peer assessment of skills and domain knowledge relevant to the learning task, the quality of the peer feedback via a questionnaire, and the collaborative outcomes via the project grades of the students. For the analysis, CGA was split into skill group awareness (skill GA) and domain knowledge group awareness (domain knowledge GA). CGA was also measured through the group agreement on skills and group agreement on domain knowledge. Results showed that peer feedback significantly positively influenced the group agreement on the skills. However, this was not the case for the group agreement on the domain knowledge. The skill GA and domain knowledge GA were also not significantly improved by the peer feedback. An explanation could be that students started to underestimate themselves in the experimental condition after the peer feedback intervention in terms of domain knowledge. The analysis also showed that, in some cases, when group agreement on skills and the skill GA got higher, the grades on the collaborative project also got higher. Based on the current study and previous studies, it is recommended to stimulate social interaction since it could aid in developing a high CGA and improving the collaborative outcomes in CSCL.

Keywords: Cognitive Group Awareness, Peer Feedback, Higher Education, Collaborative Outcomes

Introduction

According to Janssen and Bodemer (2013), computer-supported collaborative learning (CSCL) is seen as a promising approach in education since it could be effective, efficient, and enjoyable. CSCL could be defined as "*the activity of peers interacting with each other for the purpose of learning and with the support of information and communication technologies (ICT)*" (Suthers, 2012, p. 1).

An important aspect of effective collaboration and CSCL is group awareness (Kirschner & Erkens, 2013). Group awareness could be divided into *cognitive group awareness (CGA)*, which entails information about group members' knowledge and expertise, and *social group awareness*, which entails information about group members' contributions to the group process (Janssen & Bodemer, 2013). In both non-computer-supported collaboration and CSCL, CGA of relevant aspects of group members and the group as a whole, such as knowledge and skills, are essential (Bodemer & Dehler, 2011; Bodemer et al., 2018). CGA is essential because it enables students to give direction to the collaboration process and allows students to adjust to the needs of the group, which could lead to effective collaboration (Soller et al., 2005; Bodemer & Dehler, 2011; Fransen et al., 2011). Also, the development of CGA allows for knowledge exchange and knowledge construction (Engelmann et al., 2010; Cai & Gu, 2019). Additionally, CGA could increase cognitive conflicts and, with that, increase discussion, which could lead to better collaborative outcomes (Nickerson, 1999; Buder et al., 2021). Also, CGA could prevent the overestimation of similarities in skills and knowledge between group members (Nickerson, 1999; Shin et al., 2018).

However, in general, students have experienced challenges with communication and coordination (Erkens et al., 2005; Janssen et al., 2007). This could lead to a lack of CGA (Engelmann et al., 2009). A lack of communication and coordination leads to students being unable to externalize their knowledge and skills (Schnaubert & Bodemer, 2019). Students could then overestimate the similarities in knowledge, skills, and opinions and therefore not detect the relevant differences (Nickerson, 1999; Schnaubert & Bodemer, 2019), which results in students being unaware of the knowledge and skills their peers possess (Engelmann et al., 2009). To prevent this from happening more in the future, it is important to research methods that improve communication and coordination in cognitive group awareness in collaboration.

A potential solution to the lack of CGA seems to be peer feedback since this has proven to be effective in improving communication and coordination skills (Dominick et al., 1997; Phielix et al., 2011). Therefore, communication and coordination challenges could potentially be solved by peer feedback. Peer feedback could be defined as *a communicative process in which the peer-assessor and peer-assessee interact in order to gain and make sense of information and use it to improve their work* (Carless, 2015; Jonsson & Panadero, 2018; Carless & Boud, 2018). Peer feedback could potentially improve cognitive group awareness via better communication, which in turn may help

students discuss unshared knowledge (knowledge exchange) and build upon each other's knowledge (knowledge construction) (Engelmann et al., 2010; Engelmann & Hesse, 2010). Therefore, increased cognitive group awareness could positively affect individual achievement and group performance (Janssen & Bodemer, 2013).

The current study aimed to investigate if a peer feedback intervention could improve CGA by enhancing communication skills in CSCL. Also, it was investigated whether there would be a positive correlation between CGA and collaborative outcomes. The self-and peer assessment of skills and knowledge relevant to the learning task were used to measure CGA. A peer feedback intervention was designed that focused on enhancing cognitive group awareness via spoken feedback. This intervention was based on the results of the self-and peer assessment of skills and domain knowledge. The collaborative outcomes were measured in terms of the project grades of the students. The current study was focused on CGA in CSCL among students of higher education because of availability reasons.

Theoretical Framework

Computer-Supported Collaborative Learning

Computer-Supported Collaborative Learning (CSCL) has been defined in different ways. Some studies define CSCL in distinction to cooperative learning (Stahl & Hakkarainen, 2021). In cooperative learning, tasks are divided among group members and worked on (mostly) separately. Whereas in collaborative learning, tasks are worked on jointly (Knorr-Cetina, 2001), and learners seek to understand all the knowledge via meaning sharing and common understanding (Stahl & Hakkarainen, 2021).

Other studies viewed CSCL simply as a form of educational technology, where network devices are used to communicate and possibly combined with an AI application (Stahl & Hakkarainen, 2021). Those studies stated that learning could take place "through" and "around" CSCL technologies (Lehtinen et al., 1999). Through CSCL technologies entails that CSCL technology is used as a medium for synchronous and asynchronous online interaction. Around CSCL technologies entails that learners are able to interact face-to-face in an offline setting, and co-create knowledge and artefacts with the aid of CSCL technologies (Stahl & Hakkarainen, 2021). Synchronous learning in CSCL takes place in real-time and asynchronous learning takes place in non-real-time. (Text) conversations in synchronous learning are more rapid, discussion-like, and have a more social aspect, whereas, in asynchronous learning, text conversations are more formal and focused on the substance (Mabrito, 2006; Lapadat, 2002).

Still, other studies focused on CSCL as an intersubjective way of collaborating, which incorporates working in small groups and focuses on meaning-making (Stahl & Hakkarainen, 2021). Intersubjectivity in collaboration could refer to both of two things. First, it could refer to how two learners inter-relate, where they work together and understand each other from their individual cognition. Second, it could also refer to a type of joint-or group cognition shared by a group and where the individual cognition is transcended, unified, or funded (Stahl, 2016). Meaning-making entails that meaning is built upon and shared via processes of interaction, communication, and coordination (Stahl, 2007).

Yamada et al. (2016) seem to have included aspects of the before-mentioned definitions to describe CSCL. Yamada et al. (2016) stated that, in CSCL, students learn via group interaction and activities, with technology being central in this learning process. For effective learning, ICT could be used for distributed, collaborative, interactive, and constructive learning. One of the key components in CSCL, and collaboration in general, is promoting active social interaction that is needed to share experiences and knowledge, construct new knowledge, and build upon each other's knowledge and experiences (Yamada et al., 2016). The use of social interaction to build/construct new knowledge in CSCL has also been referred to in some studies as transactive knowledge building (Weinberger et al., 2007).

When people collaborate, there are collaborative outcomes, which could be measured on an individual or a group level. The individual collaborative outcomes of CSCL are primarily measured in terms of knowledge gain (Bodemer, 2011; Dehler et al., 2011; Sangin et al., 2011), while group collaborative outcomes are measured in terms of the quality of group products (Janssen et al., 2011; Phielix et al., 2011) or in terms of how the team effectiveness is perceived (Fransen et al., 2011).

The basis of effective and efficient CSCL is that a team or group is more than a few people who try to work together - using technology (Kirschner & Erkens, 2013). There are two underlying prerequisites for effective and efficient CSCL. First, learners must function as teams where they trust each other to do the work, know each other's strengths and weaknesses and use them, and share the same working norms and goals for collaborating (Kirschner & Erkens, 2013). These characteristics are not always there at the start of the collaboration but typically develop over time (Fransen et al., 2013; Kirschner & Erkens, 2013). So, when learners are just simply put together and expected to work as a team, this is where a potential problem could arise. Learners need time and opportunities to become a functioning team that learns (Fransen et al., 2013). Second, learners must have a positive outlook on working in teams in a way that they feel motivated that extra time and effort will pay off. This entails that the efforts to communicate with group members and coordinate activities are repaid with the ease and ability to learn (Kirschner & Erkens, 2013). A potential problem could occur when the costs regarding the invested time and effort exceed the benefits regarding

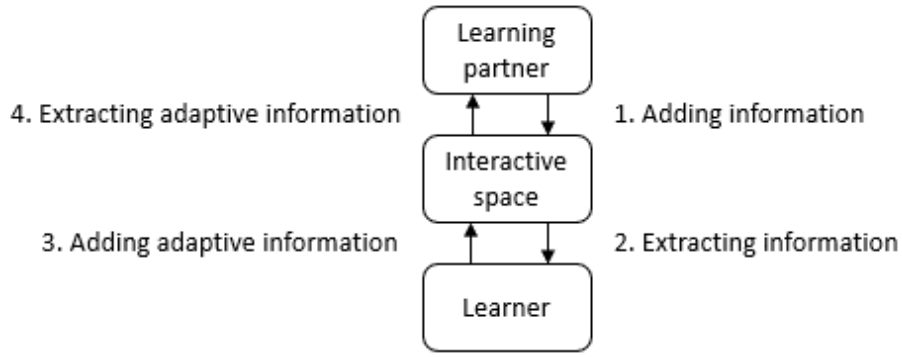
effectiveness and efficiency of learning. If this is the case, the collaboration will not take place (Kirschner et al., 2009).

Cognitive Group Awareness

An essential part of CSCL is cognitive group awareness (CGA) (Bodemer & Dehler, 2011; Bodemer et al., 2018). CGA could be defined as the awareness of the skills and knowledge of group members (Janssen & Bodemer, 2013). Active social interaction and communication will allow knowledge and skills to be exchanged, which will lead to knowledge construction and awareness of each other's knowledge and skills (Yamada et al., 2016). Figure 1 shows how CGA is developed and leads to knowledge construction. For clarity reasons, the model shows one learner and one learning partner, even though more people may be involved. The learning partner adds information on the context and content to the space where the learning partner and learner interact. This information could contain what the learning partner stated on their knowledge and skills but also an elaborated explanation from which the learner could interpret the learning partner's knowledge and skills (Engelmann et al., 2009). The learner is then extracting the added information from the interactive space, which leads to them developing CGA since they have gained information on their partner's knowledge and skills. The learner and learning partner could have also gained new knowledge, which could lead to knowledge construction if they use that knowledge to create new knowledge together (Schmidt, 2002; Engelmann et al., 2009). The process continues as the learner and the partner add and extract information to/from the interactive space. Therefore, the CGA keeps increasing, and knowledge continues to be constructed (Engelmann et al., 2009).

Figure 1

Cyclic model of developing CGA and constructing knowledge in CSCL



Note. Adapted from "Knowledge awareness in CSCL: A psychological perspective," by T. Engelmann, J. Dehler, D. Bodemer, and J. Buder, 2009, *Computers in Human Behavior*, *25*(4), p. 953 (https://doi.org/10.1016/j.chb.2009.04.004). Copyright 2009 by Elsevier Ltd.

There are multiple benefits of CGA. If learners have information about their group members' knowledge and skills, they could use this information to better regulate the interaction in their group (Buder et al., 2021) and regulate the collaborative process (Soller et al., 2005; Bodemer and Dehler, 2011; Fransen et al., 2011). Due to better regulation, learners might engage in germane learning processes and the cognitive load is reduced, which could lead to better performance during a project (Janssen & Bodemer, 2013) and better collaborative outcomes at the end of a project (Nickerson, 1999; Shin et al., 2018). Also, learners need CGA to adjust to the needs of the group (Soller et al., 2005; Bodemer and Dehler, 2011; Fransen et al., 2011).

However, learners also experience difficulties in CSCL that relate to CGA (Engelmann et al., 2009). Social interaction does not automatically take place in CSCL even though technology may support it (Kreijns et al., 2003). Previous studies showed that students, in general, particularly struggle with communication and coordination (Erkens et al., 2005; Janssen et al., 2007). These problems with social interaction could lead to a lack of context information and, more particularly, a lack of information on group members' knowledge and skills (Engelmann et al., 2009). This results in a lack of CGA since students do not have information on – and, therefore, cannot develop awareness of - group members' knowledge and skills (Engelmann et al., 2009). This could be explained by that a lack of such information leads to learners using their own knowledge as a basis for estimating the knowledge of their group members (Nickerson, 1999; Shin et al., 2018). So, by imputing one's own knowledge to group members, the similarity between their own knowledge and group members' knowledge is overestimated (Nickerson, 1999; Engelmann et al., 2009; Shin et al., 2018).

Peer Feedback

A potential solution to the communication challenges could be peer feedback since this could improve communication between group members (Perera et al., 2010; Phielix et al., 2011; Mercader et al., 2020). Peer feedback is the feedback that is provided by peers (Phielix et al., 2010a). So, for example, feedback that is provided from student to student and not feedback that is provided from teacher to student. In previous studies, peer feedback has also been defined as a communicative process in which the peer-assessor and peer-assessee interact to gain and make sense of information and use this information to improve their work (Carless, 2015; Jonsson & Panadero, 2018; Carless & Boud, 2018). Peer feedback can be outcome or process-based, be given by individuals or groups, and be received by individuals or groups (Phielix et al., 2011). Process-based peer feedback is focused on enhancing interpersonal behaviour and group processes. The information peers share is then aimed at improving performance during the project (Phielix et al., 2011). Outcome-based peer feedback is focused on providing information on the outcome of performance (Phielix et al., 2010a).

Peer feedback has multiple benefits. Peer feedback has been known to positively influence performance by giving constructive comments (Gielen et al., 2010) focused on performance criteria, strengths, weaknesses, and tips on improvement (Falchikov, 1996, as cited in Gielen et al., 2010). Also, when students are taught how to give peer feedback, it could improve their communication skills and acceptance of peer feedback (Perera et al., 2010). So, by giving peer feedback, the communication between group members could improve (Perera et al., 2010; Phielix et al., 2011; Mercader et al., 2020). This could then possibly improve CGA since the struggle with communication could be resolved.

The quality of peer feedback could be measured in different ways. One method is to measure the quality in terms of the perceived quality of peer feedback as viewed by the reviewee (Zong et al., 2021). Previous studies showed that the amount of feedback does not necessarily determine the effectiveness, but the way students perceive and accept the feedback is more important (Hattie, 2015). Someone could perceive a comment they received during peer feedback as not understandable, not actionable, incorrect, insufficient, or not persuasive. This person might, therefore, be less likely to act upon this feedback as to opposed when this person would have perceived the feedback more positively (e.g., Nelson & Schunn, 2009; Kaufman & Schunn, 2011; Patchan et al., 2016; Huisman et al., 2018; Wu & Schunn, 2020). Previous studies also showed that when someone views feedback as useful, this could increase their performance (Harks et al., 2014). Also, someone could overvalue peer feedback that is easy to act on instead of looking further into more difficult feedback, which may interfere with the learning process (Chi & Wylie, 2014).

Research Questions and Model

This research aimed to gain insight into the effect of peer feedback on cognitive group awareness (CGA), and the correlation between CGA and collaborative outcomes. Hence, the following main research question was formulated: “Does peer feedback influence cognitive group awareness in computer-supported collaborative learning, and does (improved) cognitive group awareness influence the collaborative outcomes among students of higher education?”. See Figure 2 for the accompanying research model. To be able to do analysis better and draw better conclusions, CGA was split into skill group awareness (skill GA) and domain knowledge group awareness (domain knowledge GA). Additionally, CGA was also measured by how much group members agreed on a single group member's skill-or domain knowledge level. This will be referred to as group agreement on skills and group agreement on domain knowledge. To better answer the main research question, the following hypotheses were formulated;

$H1_{(skill.GA)}$: Peer feedback will positively influence the skill GA, as measured through the difference between the self-and peer assessment scores, among students of higher education;

$H1_{(group.agreement.on.skills)}$: Peer feedback will positively influence the group agreement on skills, as measured through the standard deviation of the average peer assessment scores, among students of higher education;

$H2_{(domain.knowledge.GA)}$: Peer feedback will positively influence the domain knowledge GA, as measured through the difference between the self-and peer assessment scores, among students of higher education;

$H2_{(group.agreement.on.domain.knowledge)}$: Peer feedback will positively influence the group agreement on domain knowledge, as measured through the standard deviation of the average peer assessment scores, among students of higher education;

$H3_{(skill.GA)}$: There will be a positive correlation between skill GA and the collaborative outcomes, measured through project grades, among students of higher education;

$H3_{(group.agreement.on.skills)}$: There will be a positive correlation between group agreement on skills and the collaborative outcomes, measured through project grades, among students of higher education;

$H4_{(domain.knowledge.GA)}$: There will be a positive correlation between domain knowledge GA and the collaborative outcomes, measured through project grades, among students of higher education;

$H4_{(group.agreement.on.domain.knowledge)}$: There will be a positive correlation between group agreement on domain knowledge and the collaborative outcomes, measured through project grades, among students of higher education.

These directional hypotheses were based on what was presented in previous studies. Firstly, the influence of peer feedback on skill GA, group agreement on skills, domain knowledge GA, and group agreement on domain knowledge will be investigated. Previous studies found that, in general, students experienced struggles with communication (Erkens et al., 2005; Janssen et al., 2007), which could lead to less CGA (Engelmann et al., 2009). On the other hand, peer feedback could improve communication between group members (Perera et al., 2010; Phielix et al., 2011; Mercader et al., 2020). This led to the hypotheses (H1 and H2) that peer feedback could improve skill GA, group agreement on skills, domain knowledge GA, and group agreement on domain knowledge. Secondly, the influence of skill GA, group agreement on skills, domain knowledge GA, and group agreement on domain knowledge on collaborative outcomes will be investigated. Previous studies showed that a high CGA led to knowledge construction and engaging in germane learning processes (Engelmann et al., 2010; Janssen & Bodemer, 2013). This could lead to better performance (Janssen & Bodemer, 2013) and better collaborative outcomes (Nickerson, 1999; Shin et al., 2018). This led to the hypotheses (H3 and H4) that there will be a positive correlation between skill GA and collaborative outcomes, between group agreement on skills and collaborative outcomes, between domain knowledge GA and collaborative outcomes, and between group agreement on domain knowledge and collaborative outcomes.

Next to the hypotheses, it was also investigated if other factors influenced the current study results. These other factors that are investigated are the difference in teachers (RQ1), the quality of the peer feedback perceived by the students (RQ2), and the under-or overestimation of students (RQ3). This is not the main focus of the current study but is included to check for interference. This led to the following research questions:

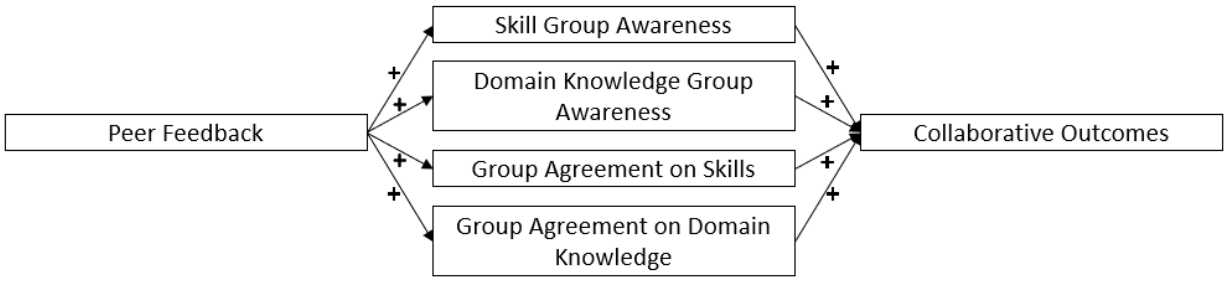
RQ1: “Could the difference in teachers interfere with the results?”;

RQ2: “Could the quality of the peer feedback as perceived by the students interfere with the results?”;

RQ3: “Could the under- or overestimation of students interfere with the results?”.

Figure 2

Research model



Research Design and Methods

Research Design

The current study examined the effects of a peer feedback intervention (independent variable) on the CGA (dependent variable) and examined the correlation between CGA (independent variable) and collaborative outcomes (dependent variable), using a quantitative, quasi-experimental research design. The students in the control condition only did the pre-and post-test. The students in the experimental condition also participated in the experiment. Figure 3, shown in the procedure section, shows an overview of the research design including the activities.

In the theoretical framework, the variables were described using literature. Table 1 gives for each study factor an explanation of how this factor is viewed/reflected in the current study.

Table 1

Description of the study factors

Variable	Description
Computer-supported collaborative learning (CSCL)	The current study incorporated an intersubjective way of collaboration where students worked in small groups on a project. The collaboration took place both around and through CSCL technologies due to the COVID measures at that time. Also, CSCL was both synchronous and asynchronous. When students would meet up (online or live), synchronous collaborative learning would take place. Asynchronous collaborative learning took place when students worked on the project at different times. During the current study, the description of CSCL by Yamada et al. (2016) was leading. The collaborative outcomes were measured using the collaborative group outcomes in terms of the quality of the group products.
Cognitive group awareness (CGA)	CGA was measured by comparing the self-assessment results with the peer assessment results focused on skills and domain knowledge related to the learning task. Metacognitive judgement was used in the self-and peer assessment. Metacognition could be divided into *knowledge of cognition*, which entails the knowledge of one's cognition (thinking about thinking), and *regulation of cognition*, which entails the sequence of actions that are taken by the students to control their learning and thinking (Mahdavi, 2014). So, in the self-and peer assessment, students provided insight into the skills and domain knowledge they thought they and their peers possessed.
Peer feedback	In the current study, students received an explanation on how to give peer feedback before actually giving peer feedback. Students used process-based peer feedback in the intervention, given by individual students in smaller groups. The students gave peer feedback to their group members aimed at developing better CGA. The peer feedback the students provided was based on the self-and peer assessment. The quality of the peer feedback was measured via a questionnaire where students rated the feedback on style, nature, and criteria, which was adapted from the study of Prins et al. (2006). This resulted in scores on how students perceived the quality of the feedback.

Organisational Context

The experiment part of the current study took place at a Higher Education Institute in the Netherlands. The students who participated in the experiment were second-year students enrolled in the Bachelor of Sport Science. At the moment of the experiment, the students were following a course called 'Beweeginterventies op Maat' (Customized sports/exercise interventions). This course

focused on analysing a deprived neighbourhood and developing an intervention that uses exercise or sports to solve a problem in that neighbourhood (Quanjel & Van Boxtel, 2021).

Participants

Before the start of the experiment, approval from the Behavioural, Management, and Social Sciences (BMS) Ethics Committee of the University of Twente was granted. At the start of the experiment, all participants signed an informed consent form.

The sampling method used in the current study was convenience sampling, gathering participants based on their availability and willingness to participate. The Higher Education Institute was contacted first to ask if it were possible to do the experiment there. They suggested doing the experiment within the course “Beweeginterventies op Maat” since this included a collaborative assignment and had space in the schedule. Four teachers of this course were available and willing to participate in the current study. The students in the classes of those teachers were asked if they wanted to participate. The participants were thus in four different classes; two of these classes were randomly assigned to the control condition and two to the experimental condition. For the collaborative assignment, students made groups of three to four. After all the data was collected, students who either missed the pre-test or post-test were filtered out. This resulted in a sample of 45 second-year students (62.20% male) from a Higher Educational Institute in the Netherlands. The students ranged in age from 18 to 24 (M = 20.22 years, SD = 1.80). There were 24 participants in the experimental condition, where 70.80% were male, and the ages ranged from 18 to 24 (M = 20.50 years, SD = 1.77). 21 students were in the control condition, where 52.40% were male, and the ages ranged from 18 to 24 (M = 19.90 years, SD = 1.81).

Instrumentation

In the current study, five instruments were used. One instrument was used to measure CGA, one to enable peer feedback, one to check the quality of the peer feedback, one to ensure objective grading of the presentations of the projects, and one to observe the grading process of the presentations of the projects. The collaborative outcomes were measured by the grades students received from their teacher for their collaborative assignment.

Cognitive Group Awareness Instrument

The idea for the CGA and peer feedback instrument was adapted from the digital tool Radar, used in the research by Phielix et al. (2010b; 2011). In Radar, students used self-reflection and peer assessment on the same criteria and were presented with the results. The current study used the

same concept, but the presentation of the results was done in the peer feedback instrument. Also, the current study used self-assessment instead of self-reflection. In the pre-and post-test, students used the self-and peer assessment to assess the domain knowledge and skills related to the course, see Appendix 1. In the self-and peer assessment, students used metacognitive judgement. So, it was asked what they think they know and are good at and what skills and domain knowledge they think their group members possess. The pre-and post-test were administered via an online questionnaire and consisted of six domain knowledge test questions, six questions on the self-and peer assessment of the domain knowledge, and eight questions on the self-and peer assessment of skills. A five-point Likert Scale was used for the self-and peer assessment questions, with 1 = not well at all and 5 = very well for questions related to skills, and with 1 = very little and 5 = very much for questions related to domain knowledge. An example question for the self-and peer assessment of skills is: "How well do you think you and your groupmates are at analysing and interpreting information from relative sources?". An example of the self- and peer assessment of domain knowledge is: "How much knowledge do you and your groupmates have of the model of Brug?".

Also, a domain knowledge test of six multiple-choice questions was included. The domain knowledge of the students was assessed on the same topics as the self-and peer assessment of the domain knowledge. This was done to check if students' self-assessment on the domain knowledge topics was accurate. Which skills and domain knowledge were assessed, was determined incoherence with the teachers. The topics of the domain knowledge assessment and domain knowledge test differed in the pre-and post-test since new topics were introduced as the course advanced. An example question for the domain knowledge test is: "What is the goal of a SWOT analysis?".

Information on the concepts was provided to prevent misconceptions of the statements on the domain knowledge and skills and improve validity. Even though the concept of the CGA instrument had already been used in previous studies (Phielix et al., 2010b; Phielix et al., 2011), the reliability was improved by having the teachers fill in CGA questionnaires before the start of the experiment. After the teachers filled in the CGA questionnaires, they were asked to provide feedback on the CGA questionnaires regarding, for example, the domain knowledge questions, the answer options, and the formulation. The feedback was implemented before letting the students fill in the pre-and post-test.

Peer Feedback Instrument

The peer feedback instrument was based on Phielix et al. (2010b; 2011). In Phielix et al. (2010b; 2011), the peer feedback was based on the presentation of the results from their digital tool

Radar. The current study used the same concept; students were presented with the results of the pre-test and used these results as a basis for their peer feedback. Additionally, the current study incorporated an explanation on how to give effective feedback and gave pointers on what to discuss and attitude. This explanation was based on the findings presented in the research by Prins et al. (2006) and included three main topics and nine items. The explanation was presented via a PowerPoint presentation, see Appendix 2, and included four pointers as well. An example of a pointer is: “Discuss the statements in which the results of the self-assessment and peer-assessment differ more than .5 points.”. An example of the explanation is: “When giving feedback, it is important to be specific.”.

Quality Check Instrument

The concept for the quality check of the feedback was adapted from the research done by Prins et al. (2006). In the study of Prins et al. (2006), the quality was checked on three main topics, with in total nine items. The topic “criteria” included two items: content and explanations. The topic “nature” included four items: remarks, questions posed, repertoire, and advice. The topic “style” included three items: structure, formulation, and style. The current study used the same items to ensure reliability, but the students rated the feedback instead of an external rater. Also, the instrument used by Prins et al. (2006) was used in a medical context. The questions were, therefore, rephrased. The current study thus measured the quality of the peer feedback perceived by the students. This was done via an online questionnaire that included nine questions (each item was formulated into one question), with a three-point Likert Scale, where 1 = minimally executed and 3 = well executed. An example question is: “The feedback had a clear structure”. The quality check questionnaire is shown in Appendix 3.

Assessment Rubric

The teachers used the assessment rubric to award grades for the final presentations of the students in a more objective manner, see Appendix 4. The final presentations were a mandatory collaborative assignment to finish the course. The grades awarded to the final presentations are seen as the collaborative outcomes in the current study. The assessment rubric was made by all teachers of the course “Beweeginterventies op Maat” together. An example of an assessment criteria is: “The chosen or designed intervention is displayed specifically”. The teachers scored the students on all 18 assessment criteria using a four-point Likert-scale, with 1 = insufficient or missing and 4 = very sufficient and explicit. These 18 criteria were divided into three parts: advice and conclusions, recommendations, and presentation. Advise and conclusions existed of eight criteria and counted for

40% of the final grade of the presentations. Recommendations existed of nine criteria and counted for 50% of the final grade of the presentations. Presentation existed of one item and counted for 10% of the final grade of the presentations. The teachers used an online computer program of the Higher Education Insitute to fill in the assessment rubric. This program also calculated the grades based on what the teachers filled in at each criterion.

Observation Scheme

An observation scheme was made and used to observe the teachers during the rating process, see Appendix 5. This was done to observe whether or not the teachers graded students differently. It was observed how many questions the teachers and students in the audience asked, how many feedback statements were given by the teachers and students in the audience and how many questions each student in the presenting group answered. The observations were done by the researcher of the current study, who used turfing to observe the before mentioned. So, at each presentation, the researcher turfed how many questions the teachers and students in the audience asked, how many feedback statements were given by the teachers and students in the audience, and how many questions each student in the presenting group answered.

Procedure

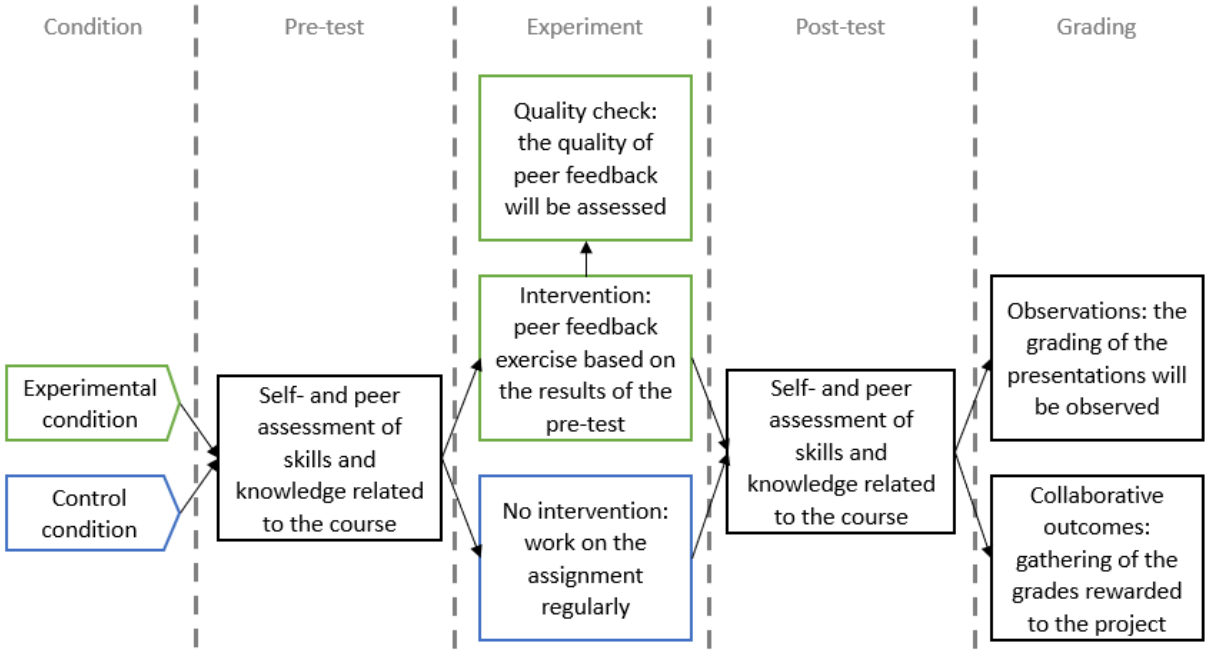
The experiment was done during school hours and in the regular classes of the course that they had the collaborative assignment in (Beweeginterventies op Maat). The first meeting of 50 minutes was in week four of the first quartile of the school year. Groups were already formed at this point, but students were starting on a new assignment. In the first part of the meeting, the participants did the pre-test to measure the level of CGA at the start of the experiment. This part took 15 minutes. The groups in the control condition handed in the pre-test directly after finishing and were not presented with the pre-test results. The groups in the experimental condition continued with the intervention. The second part of the first meeting was only for the groups in the experimental condition. The groups in the control condition just started on the new assignment without the intervention. In the second part of the meeting, the peer feedback intervention was done. In the intervention, the participants first received an explanation on how to give effective feedback and what to focus on, which took 10 minutes. Afterwards, the participants were presented with the results of the pre-test on the same device that the questionnaire was completed. They saw their results on the domain knowledge test and what they have filled in at the skills-and domain knowledge assessment. Then they gave each other peer feedback based on these results from the self-and peer assessment in the pre-test, which took 20 minutes. They discussed parts of the pre-test

that stood out to them. So, parts in which students' opinions differed or were very similar. After the intervention, the students filled in a questionnaire to give their opinion on the quality of the peer feedback to measure if the perceived quality is of influence on the results, which took 5 minutes.

The second meeting of two and a half hours was in week nine of the first quartile. Students from both the experimental- and control condition participated in the post-test during the second meeting. This was the same type of test as the pre-test, but it was done at the end of the quartile to measure, possibly improved, CGA. The post-test took 15 minutes. Also, during the second meeting, all participants finished the assignment by giving a presentation. The presentation took the rest of the time, so two hours and 15 minutes. The teachers rated these presentations using the assessment rubric as they usually would. The researcher observed the grading process of the presentations using the observation scheme. The assessment rubric was used to ensure a more objective way of grading, while the observation scheme was used to check if teachers rated similarly. The grades of the presentations were then gathered and used to measure the collaborative outcomes. Figure 3 shows a model of the research design with all the activities described in the procedure.

Figure 3

Model of the quantitative, quasi-experimental research design



Data Analysis

Computing the variables

To be able to do analysis, the variables needed to be computed. At the end of this paragraph, an overview of the variables with descriptions is shown (Table 2). To be able to perform analysis on

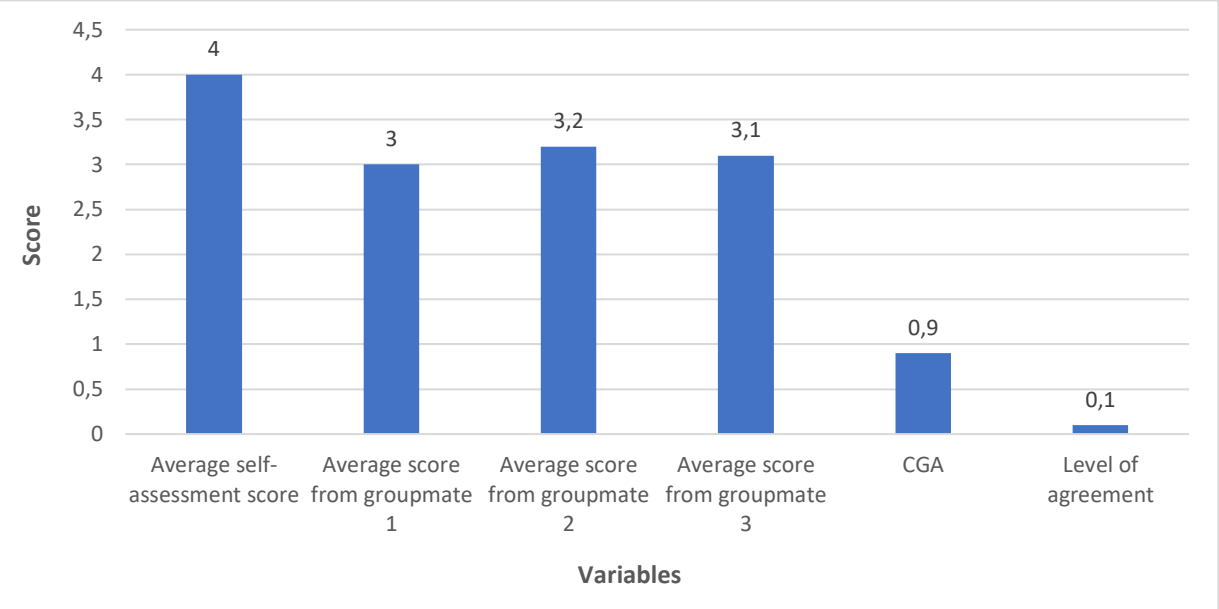
the perceived quality of the peer feedback, the average score was calculated of the questions asked in the quality check questionnaire.

CGA was split into skill GA and domain knowledge GA to perform analysis. In the pre-and post-test, questions were asked about what the students thought their level was regarding a certain skill and what that level was for the others in their group. This was asked for eight different skills. The average score was calculated for both the pre-and post-test for those eight skills. This was done for both the scores students gave themselves and the average scores they received from their groupmates. The same was done for the domain knowledge, except there were six questions. To calculate the skill GA and domain knowledge GA, the average score from themselves was subtracted from the average score they received from their groupmates, see Appendix 6 for an example calculation. Those scores were then transformed into absolute scores, so the scores for the pre- and post-test could be compared better. This number, thus, shows the absolute difference between what the students think of themselves and what their groupmates think of them. If this number is close to zero, this entails that they are in full agreement and there is a high CGA. These variables will be referred to as 'skill GA' and 'domain knowledge GA'. One thing that needed to be considered for the domain knowledge scores is that the topics were different in the pre-and post-test since the course continued over the period of the experiment.

Another way to measure CGA is to calculate the level of agreement between the groupmates. It is important to also measure the level of agreement since the average score of the groupmates could be close to the average score of themselves, but this does not necessarily mean that the groupmates are in agreement with each other. Also, when the average scores are not close to each other, this does not necessarily mean that there is low CGA. It could mean that the groupmates of a student are in high agreement but that the student over-or underestimated him- or herself. To calculate the level of agreement between the groupmates, the standard deviation was calculated for the average score the groupmates gave to one groupmate. Figure 4 gives a visual image of how the average self-assessment scores, average peer-assessment scores, CGA, and level of agreement relate. Again, to be able to do analysis, the level of agreement was calculated for the skills and domain knowledge separately. When the distance between the scores of the groupmates (standard deviation) is close to zero, this means that there is a high level of agreement. The level of agreement is, namely, measured through the standard deviation, and when the standard deviation is close to zero, this means that all groupmates gave a similar average score. These variables will be referred to as 'group agreement on skills' or 'group agreement on domain knowledge'.

Figure 4

Visualisation of self-assessment, peer-assessment, CGA, and level of agreement



The CGA questionnaire asked questions regarding self-assessment of domain knowledge and a domain knowledge test. To see if the students' self-assessment was accurate, the difference was calculated between the score on the domain knowledge test and the self-assessment score. This was done by calculating the percentage of both variables and then subtracting the percentage of the self-assessment score from the percentage of the domain knowledge test score. The percentage of the self-assessment score was calculated by dividing the self-assessment score by the scale of questions, which was five. The percentage of the score on the domain knowledge questions was calculated by dividing the score on the domain knowledge test by the scale of questions, which was six. This variable will be referred to as ‘difference’ or ‘accuracy’. When the score for the difference is close to zero, this means that the student’s self-assessment score was similar to their domain knowledge test score. The student was then accurate in their self-assessment.

The collaborative outcomes were measured via the grades awarded to the project. Grades were collected from the project done during the experiment (advise for the neighbourhood, referred to as grade of advice). These grades were given by the four teachers of the four classes that participated in the current study. So, there were no calculations needed to be able to do analysis with this variable.

Table 2

Study variables with descriptions

Variable (scale)	Description
Average skill self-assessment (1-5)	The average score of what the student scored him-/herself on the skill self-assessment questions
Average domain knowledge self-assessment (1-5)	The average score of what the student scored him-/herself on the domain knowledge self-assessment questions
Average skill peer assessment (1-5)	The average score of what the groupmates scored the student on the skill peer-assessment questions
Average domain knowledge peer assessment (1-5)	The average score of what the groupmates scored the student on the domain knowledge peer-assessment questions
Domain knowledge score (0-6)	The number of domain knowledge questions the student answered right on the domain knowledge test
Skill GA (0-4)	The absolute difference between the average skill self-assessment score and the average skill peer-assessment score
Domain knowledge GA (0-4)	The absolute difference between the average domain knowledge self-assessment score and the average domain knowledge peer-assessment score
Group agreement on skills (0-4)	The difference in the average skill assessment score from groupmate 1, groupmate 2, and (if applicable) groupmate 3, measured through the standard deviation
Group agreement on domain knowledge (0-4)	The difference in the average domain knowledge assessment score from groupmate 1, groupmate 2, and (if applicable) groupmate 3, measured through the standard deviation
Grade of advice (1-10)	The grade the students got for the collaborative assignment that was done during the period of the experiment
Difference (-1-1)	The difference between the calculated percentage of the average domain knowledge self-assessment and the calculated percentage of the domain knowledge score
Perceived quality of the peer feedback (1-3)	The average score of how the student scored the feedback on the criteria
Questions teachers (0-no max)	The observed number of questions the teachers asked during the grading process of the presentations
Questions students (0-no max)	The observed number of questions the students in the audience asked during the grading process of the presentations
Feedback teachers (0-no max)	The observed number of feedback comments the teachers gave during the grading process of the presentations
Feedback students (0-no max)	The observed number of feedback comments the students in the audience gave during the grading process of the presentations
Answered questions (0-no max)	The observed number of questions each student in the presenting group answered after they gave their presentation

Statistical tests

An experiment was used to determine if there was a statistically significant effect of peer feedback on CGA, and if there was a statistically significant correlation between CGA and collaborative outcomes. Quantitative analysis with descriptive- and inferential statistics was performed in IBM SPSS version 25. After all the variables were computed, the descriptive statistics of those variables were calculated. Afterwards, all assumptions for normal distribution and equal variance were tested. No assumptions were violated. Next, ANCOVA's were performed to investigate if there is a significant difference between the two conditions while considering other variables. Also, multiple Spearman's rho bivariate correlation tests were conducted to test for correlation between CGA and the collaborative outcomes. Lastly, Pearson's bivariate correlation tests and t-tests were conducted to check for potential interfering variables. Table 3 shows how the study variables were operationalised to address the hypotheses (H) and research questions (RQ).

For all statistical tests, a *p*-value of .05 was considered significant. For the ANCOVA tests, partial eta squared (η_p^2) was used to measure the effect size. For t-tests, Cohen's D (*d*) was used to measure the effect size. If the η_p^2-value was higher than .14 or the *d*-value was higher than .80, the effect size was considered high.

Table 3

Study variables operationalization

H/RQ	Operationalization
$H1_{(skill.GA)}$	ANCOVA with skill GA (post) as the dependent variable, skill GA (pre) as the covariate, and the condition as the fixed factor.
$H1_{(group.agreement.on.skills)}$	ANCOVA with group agreement on skills (post) as the dependent variable, group agreement on skills (pre) as the covariate, and the condition as the fixed factor.
$H2_{(domain.knowledge.GA)}$	ANCOVA with domain knowledge GA (post) as the dependent variable, domain knowledge GA (pre) as the covariate, and the condition as the fixed factor.
$H2_{(group.agreement.on.domain.knowledge)}$	ANCOVA with group agreement on domain knowledge (post) as the dependent variable, group agreement on domain knowledge (pre) as the covariate, and the condition as the fixed factor.
$H3_{(skill.GA)}$	Spearman's rho bivariate correlation test between grade of advice and skill GA (post).
$H3_{(group.agreement.on.skills)}$	Spearman's rho bivariate correlation test between grade of advice and group agreement on skills (post).
$H4_{(domain.knowledge.GA)}$	Spearman's rho bivariate correlation test between grade of advice and domain knowledge GA (post).
$H4_{(group.agreement.on.domain.knowledge)}$	Spearman's rho bivariate correlation test between grade of advice and group agreement on domain knowledge (post).
$RQ1_{(Different.teachers)}$	Independent samples t-tests with questions teachers, questions students, feedback teachers, feedback students, and answered questions as the dependent variables and the condition as the grouping variable.
$RQ2_{(Perceived.quality.of.feedback)}$	Pearson's bivariate correlation tests between: The perceived quality of the feedback and skill GA (pre and post); The perceived quality of the feedback and domain knowledge GA (pre and post); The perceived quality of the feedback and group agreement on skills (pre and post); The between the perceived quality of the feedback and group agreement on domain knowledge (pre and post).
$RQ3_{(Under\text{-}or.overestimation)}$	ANCOVA with average peer assessment (post) as the dependent variable, average self-assessment (post) as the covariate, and the condition as the fixed factor. Independent samples t-tests with the difference (pre and post) as the dependent variables and the condition as the grouping variable. Paired samples t-tests for both the experimental- and control condition, with the difference in the pre-and post-test as the pair.

Results

First, the results for the skills group awareness (skill GA) and the group agreement on skills are reported. This is followed by the results for the domain knowledge group awareness (domain knowledge GA) and the group agreement on domain knowledge. Lastly, the results regarding the collaborative outcomes are reported.

Skill Group Awareness and Group Agreement on Skills

First, ANCOVA's were conducted to determine whether peer feedback could improve skill GA and/or the group agreement on skills. First, an ANCOVA was conducted where the skill GA of the post-test was the dependent variable, the skill GA of the pre-test was the covariate, and the fixed factor was the condition. The results showed no significant difference between the conditions, with $F(1, 42) = 2.18$, $p = .147$. Second, an ANCOVA was performed where the group agreement on skills of the post-test was the dependent variable, the group agreement on skills of the pre-test was the covariate, and the condition was the fixed factor. The results showed a significant difference, with $F(1, 40) = 4.35$, $p = .043$, $\eta_p^2 = 0.098$, where the group agreement on skills was significantly better in the experimental condition than in the control condition while considering the group agreement on skills during the pre-test. Table 4 shows the descriptive statistics for the mentioned variables for the experimental-and control condition separately.

Table 4

Descriptive statistics of the skills for the experimental-and control condition.

		Experimental (n = 24)		Control (n = 21)	
		M	*SD*	*M*	*SD*
Skill GA	Pre-test	.42	(.31)	.37	(.31)
	Post-test	.35	(.23)	.49	(.42)
Group agreement on skills	Pre-test	.32[a]	(.19)	.38	(.23)
	Post-test*	.29	(.18)	.46	(.33)

Note. [a] n = 22

*$p < .05$

Lastly, Pearson's correlation tests were performed on the correlations between the perceived quality of the peer feedback and the skill GA, and between the perceived quality of the peer feedback and the group agreement on skills, during the pre-and post-test. This was done to determine whether there were any statistically significant correlations between the perceived quality

of the peer feedback and the before-mentioned variables that needed to be taken into consideration when drawing conclusions. Results showed a moderate positive correlation between the perceived quality of the peer feedback and the skill GA of the post-test, which was statistically significant (r = .43, p = .037, n = 24). No statistically significant correlations were found between the perceived quality of the peer feedback and the skill GA of the pre-test (r = -.00, p = .991, n = 24), between the perceived quality of the pee feedback and the group agreement on skills of the pre-test (r = .12, p = .589, n = 22), and between the perceived quality of the peer feedback and the group agreement on skills of the post-test (r = .04, p = .864, n = 24).

Domain Knowledge Group Awareness and Group Agreement on Domain Knowledge

For the domain knowledge GA, the same tests were performed as was done for the skill GA. However, there were also some other tests performed since domain knowledge was not only tested via self-and peer assessment but also through a domain knowledge test. To begin, multiple ANCOVA's were conducted to determine whether peer feedback could improve domain knowledge GA and/or the group agreement on domain knowledge while considering other variables. First, an ANCOVA was conducted with the domain knowledge GA of the post-test (dependent variable), the domain knowledge GA of the pre-test (covariate), and the condition (fixed factor). No main effect was observed, with F(1, 42) = .34, p = .562. Second, an ANCOVA was performed with the group agreement on domain knowledge of the post-test (dependent variable), the group agreement on domain knowledge of the pre-test (covariate), and the condition (fixed factor). Again, no main effect was observed, with F(1, 42) = .54, p = .466. See Table 6 for the descriptive statistics.

Table 6 also shows the descriptive statistics of the average domain knowledge self-assessment scores and the average domain knowledge peer-assessment scores. Mainly in the average peer-assessment scores of the post-test, there seemed to be differences between the conditions, especially if the average self-assessment scores of the post-test are taken into consideration. So, an ANCOVA was conducted with peer-assessment of domain knowledge of the post-test as the dependent variable, the self-assessment of domain knowledge of the post-test as the covariate, and the condition as the fixed factor. The results showed that the conditions were comparable, with F(1, 42) = 2.64, p = .112.

As mentioned before, tests were performed on the average domain knowledge self-assessment and the assessment of domain knowledge via a domain knowledge test. The descriptive statistics showed that the mean scores of the control condition seemed accurate in their domain knowledge self-assessment in both the pre-and post-test, as these scores are close to zero (Table 5). This entails that there was close to no difference between the average domain knowledge self-

assessment and the domain knowledge test score. The experimental condition also seemed accurate in the pre-test, but the difference in the post-test increased, and the accuracy thus decreased. An independent t-test was performed to determine if there was a significant difference between the two conditions regarding the difference in the self-assessment scores and domain knowledge test scores. Results showed that the conditions were comparable, with pre-test: $t(43) = 1.15$, $p = .255$ and post-test: $t(43) = 1.90$, $p = .064$. A paired t-test was performed to determine if there was a significant difference between the difference in the pre- and post-test. Results showed a significant difference in the experimental condition, with $t(23) = 2.22$, $p = .037$, $d = .70$, where the accuracy decreased significantly. In the control condition no significant difference was found, with $t(20) = 1.61$, $p = .123$.

Table 5

Descriptive statistics of the domain knowledge for the experimental-and control condition.

		Experimental (n = 24)		Control (n = 21)	
		M	*SD*	*M*	*SD*
Domain knowledge GA	Pre-test	.31	(.13)	.32	(.19)
	Post-test	.39	(.22)	.34	(.17)
Group agreement on domain knowledge	Pre-test	.33	(.26)	.33	(.19)
	Post-test	.43	(.34)	.38	(.28)
Average domain knowledge self-assessment	Pre-test	3.24	(.32)	3.30	(.46)
	Post-test	3.09	(.42)	3.07	(.41)
Average domain knowledge peer-assessment	Pre-test	3.45	(.27)	3.35	(.35)
	Post-test	3.32	(.38)	3.16	(.24)
Difference	Pre-test	-.02*	(.14)	.04	(.19)
	Post-test	-.14*	(.20)	-.03	(.18)

*$p < .05$

Lastly, Pearson's correlation tests were performed on the correlations between the perceived quality of the peer feedback and the domain knowledge GA and between the perceived quality of the peer feedback and the group agreement on domain knowledge during the pre-and post-test. This was done to determine whether there were any statistically significant correlations between the perceived quality of the peer feedback and the before-mentioned variables that needed to be taken into consideration when drawing conclusions. Results showed no statistically significant correlations between the perceived quality of the peer feedback and the domain knowledge GA of the pre-test ($r = -.12$, $p = .580$, $n = 24$), between the perceived quality of the peer feedback and the

domain knowledge GA of the post-test (r = .28, p = .191, n = 24), between the perceived quality of the peer feedback and the group agreement on domain knowledge of the pre-test (r = .05, p = .833, n = 24), and between the perceived quality of the peer feedback the group agreement on domain knowledge of the post-test (r = -.10, p = .653, n = 24).

Collaborative Outcomes: Grades

First, it was determined whether teachers rated similarly since the four classes had four different teachers. Multiple independent t-tests were performed to do so, with the following variables: questions teacher, questions students, feedback teacher, feedback students, and answered questions. Results showed no significant difference between the conditions for questions students (t(43) = .62, p = .540), feedback teachers (t(43) = .82, p = .415), feedback students (t(43) = 1.83, p = .057), and answered questions (t(43) = .98, p = .331). Results showed a significant difference between the conditions for questions teachers, with t(43) = 2.78, p = .008, d = .83. The teachers in the experimental condition asked significantly more questions than the teachers in the control condition. Table 6 shows the descriptive statistics of the mentioned variables.

Table 6

Descriptive statistics of the grading process for the experimental-and control condition.

		Experimental (n = 24)		Control (n = 21)	
		M	*SD*	*M*	*SD*
Number of questions asked	Teacher*	4.17	(1.88)	2.52	(2.09)
	Students	1.88	(1.90)	2.19	(1.47)
Number of feedback comments given	Teacher	1.29	(1.20)	.95	(1.56)
	Students	0	(0)	.14	(.36)
Number of answered questions		2.75	(1.87)	2.24	(1.58)

*p < .05

Since teachers rated independently and results showed that the grading process differed for teachers, Spearman's rho bivariate correlation tests were conducted for the four classes separately. Spearman's rho bivariate correlation tests were conducted between the grade of advice and the skill GA (post), between the grade of advice and the group agreement on skills (post), between the grade of advice and the domain knowledge GA (post), and between the grade of advice and the group

agreement on domain knowledge (post). See Table 7 for the descriptive statistics, where classes 1 and 2 belong to the experimental condition, and classes 3 and 4 belong to the control condition.

Table 7

Descriptive statistics of group awareness and group agreement for the classes

	Class 1 (n = 13)		Class 2 (n = 11)		Class 3 (n = 6)		Class 4 (n = 15)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Skill GA (post)	.31	(.23)	.39	(.23)	.44	(.28)	.51*	(.47)
Group agreement on skills (post)	.23	(.15)	.36	(.20)	.35	(.24)	.50*	(.36)
Domain knowledge GA (post)	.39	(.30)	.48	(.39)	.33	(.35)	.39	(.26)
Group agreement on domain knowledge (post)	.30	(.18)	.49	(.23)	.35	(.25)	.34	(.14)
Grade of advice	5.64	(.61)	7.20	(1.95)	7.60	(.66)	6.41	(1.41)

*$p < .05$

For class 1, there were no statistically significant correlations between the grade of advice and the skill GA (post) (r_s = -.42, p = .153, n = 13), between the grade of advice and the group agreement on skills (post) (r_s = -.41, p = .170, n = 13), between the grade of advice and the domain knowledge GA (post) (r_s = .17, p = .585, n = 13), and between the grade of advice and the group agreement on domain knowledge (post) (r_s = -.52, p = .071, n = 13).

For class 2, there were also no statistically significant correlations between the grade of advice and the skill GA (post) (r_s = -.55, p = .080, n = 11), between the grade of advice and the group agreement on skills (post) (r_s = -.165, p = .627, n = 11), between the grade of advice and the domain knowledge GA (post) (r_s = -.029, p = .933, n = 11), and between the grade of advice and the group agreement on domain knowledge (post) (r_s = .40, p = .226, n = 11).

For class 3, there were again no statistically significant correlations between the grade of advice and the skill GA (post) (r_s = -.29, p = .573, n = 6), between the grade of advice and the group agreement on skills (post) (r_s = -.40, p = .437, n = 6), between the grade of advice and the domain knowledge GA (post) (r_s = -.24, p = .573, n = 6), and between the grade of advice and the group agreement on domain knowledge (post) (r_s = -.503, p = .310, n = 6).

For class 4, there was a moderate negative correlation between the grade of advice and the skill GA (post), which was statistically significant (r_s = -.52, p = .049, n = 15). So, as the scores for the skills GA decreased (which means the skill GA got better), the grades of advice increased. There was

also a strong negative correlation between the grade of advice and the group agreement on skills (post), which was statistically significant (r_s = -.61, p = .015, n = 15). So, as the scores for the group agreement on skills decreased (which means the group agreement on skills improved), the grades of advice increased. On the other hand, there were no statistically significant correlations between the grade of advice and the domain knowledge GA (r_s = -.34, p = .217, n = 15), and between the grade of advice and the group agreement on domain knowledge (r_s = .15, p = .606, n = 15).

Discussion

The main research question of the current study was: "Does peer feedback influence cognitive group awareness in computer-supported collaborative learning, and does (improved) cognitive group awareness influence the collaborative outcomes among students of higher education?". Cognitive group awareness (CGA) was divided into skill group awareness (skill GA) and domain knowledge group awareness (domain knowledge GA) for analysis. Additionally, CGA was also measured through the group agreement, which was also divided into the group agreement on skills and group agreement on domain knowledge. The results showed that students who engaged in the peer feedback session had significantly better group agreement on skills at the end of the experiment. Also, the analysis showed that, in some cases, when the group agreement on skills improved, the collaborative outcomes (measured through the presentation grades) improved as well. The same goes for the skill GA. In some cases, when the skill GA improved, the grades improved as well. On the other hand, results showed no significant difference in the group agreement on domain knowledge, skill GA, or domain knowledge GA between students who participated in the peer feedback session and students who did not. Additionally, no statistically significant correlation was found between the grades and the domain knowledge GA, and between the grades and the group agreement on domain knowledge.

The Influence of Peer Feedback on CGA and Collaborative Outcomes

Skill Group Awareness and Group Agreement on Skills

When CGA was measured through the group agreement on skills, the group agreement in the groups from the experimental condition was significantly better than the agreement in the groups from the control condition when considering the group agreement on skills during the pre-test. So, students who participated in the peer feedback session had a significantly higher level of agreement than those who did not, considering the level of agreement before the peer feedback. With these results, the hypothesis ($H1_{(group.agreement.on.skills)}$) that peer feedback will positively influence the group agreement on skills, is accepted. These results were also in line with what the literature suggested.

Namely, social interaction and communication are needed for students to have and improve CGA (Engelmann et al., 2009; Yamada et al., 2016). However, students are known to have problems with social interaction and communication (Kreijns et al., 2003; Erkens et al., 2005; Jansen et al., 2007), which could interfere with the development of CGA (Engelmann et al., 2009). Previous studies stated that peer feedback could improve communication skills and communication between group members (Perera et al., 2010; Phielix et al., 2011; Mercader et al., 2020), which could potentially solve the problems in social interaction and communication and allow for CGA to develop. With regards to CGA being measured through the group agreement on skills, peer feedback had that desired effect.

There was also a statistically significant positive correlation between the perceived quality of the peer feedback and the skill group awareness (skill GA) during the post-test. This shows that the perceived quality of peer feedback could have interfered with the results regarding skill GA, which answers RQ2: “Could the quality of the peer feedback as perceived by the students interfere with the results?”. When the perceived quality of the peer feedback got higher, the skill GA in the post-test also got higher. This correlation could, however, be explained by previous studies. Previous studies stated that when somebody perceived the feedback as high-quality, they were more likely to act upon and accept the feedback (e.g., Nelson & Schunn, 2009; Kaufman & Schunn, 2011; Patchan et al., 2016; Huisman et al., 2018; Wu & Schunn, 2020). So, students were more accepting of their group members' views of them. This could result in high CGA as students now accept and are aware of how group members view them.

However, when CGA was measured through the difference between the average self-assessment scores and average peer-assessment scores, there were no significant differences found between the conditions for the skill GA when the scores of the pre-test on skill GA were taken into consideration. This means that the students who participated in the peer feedback session did not have better skill GA than those who did not participate in the peer feedback session while considering the students’ skill GA at the start. Considering these results, the hypothesis ($H1_{(skill.GA)}$) that peer feedback will positively influence the skill GA, is rejected. This was not in line with the line of reasoning that the literature suggested. As mentioned above, previous studies found that students experienced struggles with communication (Erkens et al., 2005; Janssen et al., 2007), which could lead to less CGA (Engelmann et al., 2009). On the other hand, peer feedback could improve communication between group members (Perera et al., 2010; Phielix et al., 2011; Mercader et al., 2020) and potentially (indirectly) improve CGA. The difference in findings could be explained by the differences in the types of peer feedback used and that the peer feedback was, therefore, less effective in improving communication in the current study. In the studies of Perera et al. (2010), Phielix et al. (2011), and Mercader et al. (2020), peer feedback did improve communication. In these

studies, written feedback was used, whereas spoken feedback was used in the current study. Previous studies showed that even though spoken feedback has benefits, such as being less laborious and students becoming less anxious (Attali & Powers, 2010), there are also some drawbacks opposed to written feedback (Buckley, 2012). Firstly, spoken feedback is mostly instant feedback. The feedback is then given right after or only a short time after the assessment (Buckley, 2012). However, feedback is better retained if given a day or more after, which is mostly the case with written feedback (Kippel, 1975). Secondly, Buckley (2012) also found that students actually prefer written feedback since they could review it later, hold the giver accountable, and not forget it.

Another explanation could be that previous studies show that a large amount of feedback at once is ineffective (Bitchener & Knoch, 2009). When students receive large amounts of feedback, they tend to feel overwhelmed (Freeman & Roger, 2016). It is not certain that the students in the current study gave a large amount of feedback, but they were given extensive time to give feedback on the whole self-and peer assessment, so it is very likely. This may have resulted in a lot of - but ineffective - feedback and students feeling overwhelmed.

There were also differences found in the participants, context, and goals of the feedback between the consulted studies and the current study. The study of Phielix et al. (2011) was done in a high-school setting with fifteen-and sixteen-year-olds with a similar goal as the current study, namely to improve cognitive group awareness. The study of Perera et al. (2010) was done in a medical context with Medical Science students and with the goal of improving communication skills regarding empathy, addressing concerns, and interview style. The study of Mercader et al. (2020) was done within the Faculty of Education and with mostly women. On the other hand, the current study was done with Sport Science students, where the majority were men, and the goal was to improve CGA. Previous studies state that the use, process, and effectiveness of feedback could vary for different individuals and contexts (Boud & Falchikov, 2007; Purchase & Hamer, 2018). This could explain the contraries in findings since there were differences in the participants, contexts, and goals between the current study and the consulted studies. In conclusion, all of the explanations mentioned above could be part of why peer feedback did not lead to improved CGA in the current study.

Domain Knowledge Group Awareness and Group Agreement on Domain Knowledge

In the post-test, topics were more difficult than in the pre-test since the course advanced. Results showed that the domain knowledge group awareness (domain knowledge GA) and group agreement on domain knowledge decreased in both conditions as topics got more difficult. This resulted in no significant differences between the conditions. Therefore, the hypothesis ($H2_{(domain.knowledge.GA)}$) that peer feedback will positively influence the domain knowledge GA and the

hypothesis ($H2_{(group.agreement.on.domain.knowledge)}$) that peer feedback will positively influence the group agreement on domain knowledge, are rejected. This was not in line with what the literature suggested. Communication is needed to develop CGA (Engelmann et al., 2009; Yamada et al., 2016). Since peer feedback has been known to improve communication (Perera et al., 2010; Phielix et al., 2011; Mercader et al., 2020), it was expected that it would also improve CGA. The difference in findings could be explained by the same factors mentioned in the previous paragraph. So, the spoken feedback could not have been retained properly (Kippel, 1975). Also, students could not review it, hold the giver accountable, and students could forget the feedback (Buckley, 2012). The latter could also occur because of the large amount of feedback that was given (Bitchener & Knoch, 2009). With large amounts of feedback, students could feel overwhelmed (Freeman & Roger, 2016). Additionally, the differences in participants, contexts, and goals between the consulted studies and the current study could explain the different findings since the use, process, and effectiveness of feedback could vary for different individuals and contexts (Boud & Falchikov, 2007; Purchase & Hamer, 2018).

Another explanation could be that the students might have started to underestimate themselves since results showed that the domain knowledge GA and group agreement on domain knowledge decreased in both conditions when topics became more difficult. However, the descriptive statistics showed that the scores decreased more in the experimental condition. Although this was not a significant difference, it suggested that students might have underestimated themselves more in the experimental condition with respect to the scores they received from their peers. Analysis regarding the average peer-assessment scores showed no main effect, but further analysis that was done with the difference in the self-assessment scores and the domain knowledge test score did show a significant difference. Students in the experimental condition underestimated themselves significantly more after the peer feedback than before the peer feedback. Their self-assessment scores were thus lower than their domain knowledge test score. In the control condition, no significant difference was found between the start and the end of the experiment regarding the difference in self-assessment and domain knowledge scores. This underestimation in the experimental condition could explain why no significant differences between the conditions were found. If this underestimation did not take place in the experimental condition, they could have, arguably, significantly improved their domain knowledge GA and group agreement on the domain knowledge. This also shows that underestimation could have interfered with the results of the current study, which answers RQ3: “Could the under-or overestimation of students interfere with the results?”. Nevertheless, this underestimation could be explained by literature. Previous studies showed that students who had to judge themselves multiple times became under-confident and started to underestimate themselves (Koriat et al., 2002). This phenomenon is called under-confidence with practice (UWP) (Koriat et al., 2002). UWP could have also occurred in the current

study since students had to judge themselves in the self-assessment and again during the peer feedback session.

Collaborative Outcomes: Grades

The four classes that participated in the current study had four different teachers who all rated independently, which resulted in a significant difference in the grading process. In the experimental condition, teachers asked significantly more questions than the teachers in the control condition. This also shows that the presence of four different teachers in the four classes did interfere with the results, which answers RQ1: "Could the difference in teachers interfere with the results?". Tests were performed for all four classes separately to be on the safe side. Regarding the hypothesis ($H3_{(skill.GA)}$) that there will be a positive correlation between skill GA and the collaborative outcomes and the hypothesis ($H3_{(group.agreement.on.skills)}$) that there will be a positive correlation between group agreement on skills and the collaborative outcomes, results were inconclusive. In class 4, results showed a statistically significant negative correlation between the grades and the skill GA in the post-test, and between the grades and the group agreement on skills in the post-test. This entails that when the scores on the group agreement on skills and the scores on the skill GA got lower, which means that the group agreement on skills and the skill GA got better, the grades got higher. These results were in line with what was suggested by the literature. Studies by Nickerson (1999) and Shin et al. (2018) stated that CGA could lead to better collaborative outcomes. Janssen and Bodemer (2013) explained that by having a high CGA, the cognitive load could be reduced, and germane learning processes could take place. This could result in better performance (Janssen & Bodemer, 2013). The results that, in class 4, grades got higher as the skill GA and the group agreement on skills improved, confirmed what the literature stated. However, in classes 1, 2, and 3, there were no statistically significant correlations between the grades and the skill GA, and between the grades and the group agreement on skills. Additionally, in all four classes, there were no statistically significant correlations between the grades and domain knowledge GA, and between the grades and group agreement on domain knowledge. With these results, the hypothesis ($H4_{(domain.knowledge.GA)}$) that there will be a positive correlation between domain knowledge GA and the collaborative outcomes and the hypothesis ($H4_{(group.agreement.on.domain.knowledge)}$) that there will be a positive correlation between group agreement on domain knowledge and the collaborative outcomes, are rejected. This was not in line with what was suggested in the literature since it was expected that better CGA would lead to better collaborative outcomes (Nickerson, 1999; Janssen & Bodemer, 2013; Shin et al., 2018). The difference in findings could be explained by the small sample size now that analysis is done for the four classes separately. With a small sample size, there is a greater chance for no logical or expected

results in general (Hackshaw, 2008). In correlation particularly, a larger sample size considerably improves the probability of detecting a correlation (Juslin & Olssen, 2005). Some studies even state that the sample size should be close to at least 250 for stable estimates in correlation (Schönbrodt & Perugini, 2018). One factor that is also of influence in correlations are outliers. Previous studies showed that after outliers were removed, the correlations were more accurate, and the magnitude improved (Osborne & Overbay, 2004). However, potential outliers were not removed in the current study and could have still influenced the correlations. Potential outliers were not removed because it was not certain if they would be outliers due to the small sample sizes. Sample sizes ranged from only six to fifteen in the classes. This resulted in a minimum of two project groups and a maximum of four project groups per class.

Another explanation could be that the grades were given to the project groups, while the CGA and group agreement scores were on an individual level. So, it could be that one member of the project group had low CGA and group agreement but did get a good grade on the project since this was rewarded to the whole project group. To clarify, an example is given. Group member one has low skill GA, whereas group members two and three have high skill GA. All three worked on the project together and received a relatively good grade, for example, a seven. In the correlation analysis, there are now two students who have high skill GA and a good grade but also one student with low skill GA and a good grade. This could have influenced the outcomes of the correlation tests, especially considering that there were only a few project groups per class and thus only a few different grades that were analysed.

Theoretical and Practical Implications

Theoretical Implications

The current study showed that peer feedback could improve group agreement on skills. Additionally, the current study showed that, in some cases, when the group agreement on the skills or the skill GA improved, the grades improved as well. These results confirmed the theory of the current study that peer feedback could improve (a form of) CGA and that (in some cases) better CGA could lead to better collaborative outcomes. This theory was based on different studies done by Nickerson (1999), Engelmann et al. (2009), Perera et al. (2010), Phielix et al. (2011), Janssen and Bodemer (2013), Yamada et al. (2016), Shin et al. (2018), and Mercader et al. (2020). These studies were done in different contexts than the current study. So, the current study also showed that part of the theory is also applicable to students in the sports sector.

On the contrary, the current study also showed that peer feedback did not improve domain knowledge GA and group agreement on domain knowledge. However, the setup and the type of

feedback in the current study differed from the consulted studies. Studies that investigated a similar theory regarding CGA used self-reflection, peer assessment, and written peer feedback (Phielix et al., 2010b; Phielix et al., 2011). On the other hand, the current study used a combination of self-assessment, peer assessment, and spoken peer feedback. The main difference between these studies (Phielix et al., 2010b; Phielix et al., 2011) and the current study is the type of peer feedback used. So, it could be that the type of feedback has different effects on CGA. To rectify this outcome, research should be done on how different types of peer feedback (e.g., written versus spoken, little versus large, or short-term versus long-term) influence the development of CGA.

Practical Implications

Regarding the effect of peer feedback on CGA, results showed that it could improve the group agreement on the skills. The current study also showed that, in some cases, higher skills GA and group agreement on skills leads to higher grades. As mentioned earlier, a high CGA leads to a reduced cognitive load and allows for germane learning processes to occur, which leads to better performance (Janssen & Bodemer, 2013). So, since it has been widely established that social interaction is critical in developing a high CGA (Schmidt, 2002; Engelmann et al., 2009; Yamada et al., 2016), teachers should try to stimulate that in CSCL. One method to stimulate social interaction is via peer feedback (Perera et al., 2010; Phielix et al., 2011; Mercader et al., 2020). The current study showed that this could significantly improve group agreement on skills, which also had a statistically significant positive correlation with the grades. Even so, there are also other methods that teachers could use to stimulate social interaction, such as self-prompted communication guidelines or visualisation followed by discussion. Self-prompted communication guidelines help initiate social interaction, leading to an increased discussion on conversational/relevant topics and decreased discussion on irrelevant topics (Hughes et al., 2000). A visualisation could be made of students' performance or opinions on statements (Sangin et al., 2011; Gijlers & De Jong, 2009). This visualisation is then fed back to the students (Sangin et al., 2011; Gijlers & De Jong, 2009), which will lead to an elaborate discussion (Janssen & Bodemer, 2013). Using these methods could thus stimulate social interaction and help students develop better CGA, which could result in higher grades according to the results of the current study.

Limitations and Future Research

Limitations

There are some limitations of the current study that need to be considered when looking at the results. Firstly, the sample size (n = 45) of the current study is relatively small. This could limit the

generalisation of the results since taking the results of a small sample and projecting them to a large population could be risky and difficult (Tipton et al., 2017). With a small sample, there is a risk that results might not be a real effect but are a product of chance. So, in statistical tests, small sample sizes could fail to produce logical, expected, or reliable results (Hackshaw, 2008). However, small sample sizes also have benefits. It allows for theories to be tested faster and avoids spending too many resources, such as financial means (Hackshaw, 2008), which was also the case in the current study.

The second limitation is that there was no way to conduct an inter-rater reliability analysis to determine whether the four teachers graded similarly. Since the teachers graded independently and the results showed a significant difference in the grading process, it was very likely that the four teachers did not grade in the same manner. So, analysis regarding the grades was done for the four classes separately, which resulted in even smaller sample sizes. If inter-rater reliability could have been established, analysis regarding the grades could have been done with the whole sample instead of the four classes separately. This could have potentially shown different effects but, more importantly, resulted in more reliable and logical results (Hackshaw, 2008).

The last limitation is that the experiment had to be done in a short time period and had to fit the course schedule. This resulted in that there was only time for one peer feedback session in which the students had to discuss all parts of the CGA questionnaire. Previous studies showed that feedback is less effective when given in large amounts at once (Bitchener & Knoch, 2009). So, it could be that having only one feedback session due to limited time may have caused the feedback to be less effective.

Future Research

Future studies could investigate if different types of peer feedback have a different effect on CGA. Previous studies used written peer feedback and got different results than the current study, which used spoken peer feedback (e.g., Perera et al., 2010; Phielix et al., 2011; Mercader et al., 2020). Another difference that could be investigated is online versus offline peer feedback. In the consulted studies, the peer feedback was also online, whereas the peer feedback in the current study was offline. The study of Phielix et al., 2011 also used a peer feedback tool in their intervention. The current study used self-and peer assessment as a basis for peer feedback but did not use a specific tool. These differences suggest that different types of peer feedback, such as written versus spoken, online versus offline, and tool versus no tool, could potentially affect CGA differently. It would be interesting to find out if this is indeed the case.

Another suggestion for future research would be to investigate the effect of peer feedback on CGA when the students are trained in giving and receiving peer feedback. Previous studies have found that training students in giving and receiving peer feedback significantly improved the quality of the peer feedback and led to more active discussions (Zhu, 1995). Also, previous studies stated that, after training, students were also significantly better at receiving feedback in the sense that they incorporated more feedback (Min, 2006). It would be interesting to find out if training students in peer feedback could make a significant difference instead of not training them regarding developing CGA.

The last suggestion would be to do a longitudinal-experimental study. The current study had a quasi-experimental design in which the experiment had to be done within five weeks and had to fit the course schedule. This resulted in only one peer feedback session where students had to give much feedback to one another, which could have been less effective (Bitchener & Knoch, 2009). A study with a longitudinal-experimental design could do multiple smaller sessions of peer feedback. Students would give smaller amounts of feedback in these sessions, which would probably be more effective (Bitchener & Knoch, 2009). It would be interesting to find out if this way of giving peer feedback would improve CGA significantly.

Conclusions and Recommendations

To conclude, in the current study, when cognitive group awareness (CGA) was measured through the group agreement on skills, students in the experimental condition had significantly higher group agreement on skills than students in the control condition. However, when CGA was measured through the difference was measured in the average self-and peer assessment scores on the skills (skill GA), no significant difference was found between the conditions. Another finding was that when the perceived quality of the peer feedback got higher, the skill GA also got higher.

Regarding the domain knowledge GA and the group agreement on knowledge, there were no significant differences between the two conditions. However, students in the experimental condition started to underestimate themselves after the peer feedback intervention. After the peer feedback intervention, the difference between their self-assessment and domain knowledge scores was significantly bigger, with the self-assessment score being lower than the domain knowledge score.

Regarding the grades and thus the collaborative outcomes, the results were inconclusive. Tests did show that there was, in some cases, a statistically significant positive correlation between the grades and the skill GA and between the grades and the group agreement on skills. So, when the skill GA and group agreement on skills improved, grades improved as well. Nevertheless, this was not the case in all four classes. Also, there were no statistically significant correlations between the

grades and the domain knowledge GA and the grades and the group agreement on domain knowledge.

There are a few recommendations. Firstly, based on the current study and previous studies, it would be recommended for teachers (or other people in a leadership position) to stimulate social interaction when working in groups or teams to improve CGA and get better collaborative outcomes (Schmidt, 2002; Engelmann et al., 2009; Janssen & Bodemer, 2013; Yamada et al., 2016). Secondly and lastly, it would be recommended to do further research into the following three suggestions; 1) the effect of different types of peer feedback on CGA; 2) the effect of training students in peer feedback on CGA; 3) a longitudinal-experimental study with multiple smaller peer feedback sessions.

References

Attali, Y., & Powers, D. (2010). Immediate feedback and opportunity to revise answers to open-ended questions. *Educational and Psychological Measurement*, *70*(1), 22-35. https://doi.org/10.1177/0013164409332231

Bitchener, J., & Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT journal*, *63*(3), 204-211. https://doi.org/10.1093/elt/ccn043

Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior, 27*(3), 1079-1086. https://doi.org/10.1016/j.chb.2010.05.016

Bodemer, D., & Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, *27*(3), 1043-1045. https://doi.org/10.1016/j.chb.2010.07.014

Bodemer, D., Janssen, J., & Schnaubert, L. (2018). Group awareness tools for computer-supported collaborative learning. In *International handbook of the learning sciences* (pp. 351-358). New York: Routledge/Taylor & Francis. http://dx.doi.org/10.4324/9781315617572-34

Boud, D., & Falchikov, N. (2007). Introduction: Assessment for the longer term. In *Rethinking assessment in higher education*, 13-23. London: Routledge. https://doi.org/10.4324/9780203964309-7

Buckley, P. (2012). Can the effectiveness of different forms of feedback be measured? Retention and student preference for written and verbal feedback in level 4 bioscience students. *Journal of Biological Education*, *46*(4), 242-246. https://doi.org/10.1080/00219266.2012.702676

Buder, J., Bodemer, D., & Ogata, H. (2021). Group awareness. *International Handbook of Computer-Supported Collaborative Learning*, 295-313. https://link.springer.com/chapter/10.1007/978-3-030-65291-3_16

Cai, H., & Gu, X. (2019). Supporting collaborative learning using a diagram-based visible thinking tool based on cognitive load theory. *British Journal of Educational Technology*, *50*(5), 2329-2345. https://doi.org/10.1111/bjet.12818

Carless, D. (2015). *Excellence in university assessment: Learning from award-winning practice* (1st ed.). Routledge. https://doi.org/10.4324/9781315740621

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315-1325. https://doi.org/10.1080/02602938.2018.1463354

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist*, *49*(4), 219-243. https://doi.org/10.1080/00461520.2014.965823

Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior, 27*(3), 1068-1078. https://doi.org/10.1016/j.chb.2010.05.018

Dominick, P. G., Reilly, R. R., & McGourty, J. W. (1997). The effects of peer feedback on team member behavior. *Group & Organization Management, 22*(4), 508-520. https://doi.org/10.1177/1059601197224006

Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, *25*(4), 949-960. https://doi.org/10.1016/j.chb.2009.04.004

Engelmann, T., & Hesse, F. W. (2010). How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning*, *5*(3), 299-319. https://doi.org/10.1007/s11412-010-9089-1

Engelmann, T., Tergan, S. O., & Hesse, F. W. (2010). Evoking knowledge and information awareness for enhancing computer-supported collaborative problem solving. *The Journal of Experimental Education*, *78*(2), 268-290. https://doi.org/10.1080/00220970903292850

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior*, *21*(3), 463-486. https://doi.org/10.1016/j.chb.2004.10.038

Falchikov, N. (1996) Improving learning through critical peer feedback and reflection. *Higher Education Research and Development, 19,* 214–218.

Fransen, J., Kirschner, P. A., & Erkens, G. (2011). Mediating team effectiveness in the context of collaborative learning: The importance of team and task awareness. *Computers in Human Behavior, 27*(3), 1103–1113. https://doi.org/10.1016/j.chb.2010.05.017

Fransen, J., Weinberger, A., & Kirschner, P. A. (2013). Team effectiveness and team development in CSCL. *Educational Psychologist, 48*(1), 9–24. https://doi.org/10.1080/00461520.2012.747947

Freeman, R., & Roger, L. (2016). *Planning and Implementing Assessment* (1st ed.). Routledge. https://doi.org/10.4324/9781315041858

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and instruction*, *20*(4), 304-315. https://doi.org/10.1016/j.learninstruc.2009.08.007

Gijlers, H., & de Jong, T. (2009). Sharing and confronting propositions in collaborative inquiry learning. *Cognition and Instruction, 27*(3), 239-268. https://doi.org/10.1080/07370000903014352

Hackshaw, A. (2008). Small studies: strengths and limitations. *European Respiratory Journal, 32*(5), 1141-1143. https://doi.org/10.1183/09031936.00136408

Harks, B., Rakoczy, K., Hattie, J., Besser, M., & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educational Psychology*, *34*(3), 269-290. https://doi.org/10.1080/01443410.2013.785384

Hattie, J. (2015). The applicability of visible learning to higher education. *Scholarship of teaching and learning in psychology*, *1*(1), 79. https://doi.org/10.1037/stl0000021

Hughes, C., Rung, L. L., Wehmeyer, M. L., Agran, M., Copeland, S. R., & Hwang, B. (2000). Self-prompted communication book use to increase social interaction among high school students. *Journal of the Association for Persons with Severe handicaps*, *25*(3), 153-166. https://doi.org/10.2511/rpsd.25.3.153

Huisman, B., Saab, N., Van Driel, J., & Van Den Broek, P. (2018). Peer feedback on academic writing: undergraduate students' peer feedback role, peer feedback perceptions and essay performance. *Assessment & Evaluation in Higher Education*, *43*(6), 955-968. https://doi.org/10.1080/02602938.2018.1424318

Janssen, J., & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist*, *48*(1), 40-55. https://doi.org/10.1080/00461520.2012.749153

Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in human behavior*, *23*(3), 1105-1125. https://doi.org/10.1016/j.chb.2006.10.005

Janssen, J., Erkens, G., & Kirschner, P. A. (2011). Group awareness tools: It's what you do with it that matters. *Computers in human behavior, 27*(3), 1046-1058. https://doi.org/10.1016/j.chb.2010.06.002

Jonsson, A., & Panadero, E. (2018). Facilitating Students' Active Engagement with Feedback. In A. Lipnevich & J. Smith (Eds.), *The Cambridge Handbook of Instructional Feedback* (pp. 531-553). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316832134.026

Juslin, P., & Olsson, H. (2005). Capacity Limitations and the Detection of Correlations: Comment on Kareev (2000). *Psychological Review, 112*(1), 256–267. https://doi.org/10.1037/0033-295X.112.1.256

Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science*, *39*(3), 387-406. https://doi.org/10.1007/s11251-010-9133-6

Kippel, G. M. (1975). Information feedback, need achievement and retention. *The Journal of Educational Research*, *68*(7), 256-261. https://doi.org/10.1080/00220671.1975.10884766

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational psychology review*, *21*(1), 31-42. https://doi.org/10.1007/s10648-008-9095-2

Kirschner, P. A., & Erkens, G. (2013). Toward a framework for CSCL research. *Educational Psychologist*, *48*(1), 1-8. https://doi.org/10.1080/00461520.2012.750227

Knorr-Cetina, K. (2001). Objectual practices. In K.-C. T. Schatzki & K. E. Von Savigny (Eds.), *The practice turn in contemporary theory* (pp. 175–188). London: Routledge. http://dx.doi.org/10.4324/9780203977453-22

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General, 131*(2), 147–162. https://doi.org/10.1037/0096-3445.131.2.147

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in human behavior*, *19*(3), 335-353. https://doi.org/10.1016/S0747-5632(02)00057-2

Lapadat, J. C. (2002). Written interaction: A key component in online learning. *Journal of computer-mediated communication*, *7*(4), JCMC742. https://doi.org/10.1111/j.1083-6101.2002.tb00158.x

Lehtinen, E., Hakkarainen, K., Lipponen, L., Rahikainen, M., & Muukkonen, H. (1999). Computer supported collaborative learning: A review. *The JHGI Giesbers reports on education*, *10*, 1999. https://www.researchgate.net/profile/Hanni-Muukkonen/publication/250788384_Computer_Supported_Collaborative_Learning_A_Review/links/0c96051f22f00d0694000000/Computer-Supported-Collaborative-Learning-A-Review.pdf

Mabrito, M. (2006). A Study of Synchronous Versus Asynchronous Collaboration in an Online Business Writing Class. *The American Journal of Distance Education, 20(*2), 93-107. https://doi.org/10.1207/s15389286ajde2002_4

Mahdavi, M. (2014). An overview: Metacognition in education. *International Journal of Multidisciplinary and current research, 2*(6), 529-535. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1085.650&rep=rep1&type=pdf

Mercader, C., Ion, G., & Díaz-Vicario, A. (2020). Factors influencing students' peer feedback uptake: instructional design matters. *Assessment & Evaluation in Higher Education, 45*(8), 1169-1180. https://doi.org/10.1080/02602938.2020.1726283

Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of second language writing*, *15*(2), 118-141. https://doi.org/10.1016/j.jslw.2006.01.003

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science, 37*(4), 375–401. https://doi.org/10.1007/s11251-008-9053-x

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological bulletin*, *125*(6), 737. https://doi.org/10.1037/0033-2909.125.6.737

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research, and Evaluation*, *9*(1), 6. https://doi.org/10.7275/qf69-7k43

Patchan, M. M., Schunn, C. D., & Correnti, R. J. (2016). The nature of feedback: How peer feedback features affect students' implementation rate and quality of revisions. *Journal of Educational Psychology, 108*(8), 1098–1120. https://doi.org/10.1037/edu0000103

Perera, J., Mohamadou, G., & Kaur, S. (2010). The use of objective structured self-assessment and peer-feedback (OSSP) for learning communication skills: evaluation using a controlled trial. *Advances in Health Sciences Education*, *15*(2), 185-193. https://doi.org/10.1007/s10459-009-9191-1

Phielix, C., Prins, F. J., & Kirschner, P. A., (2010a). Assessment of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior, 26*(2), 151-161. https://doi.org/10.1016/j.chb.2009.10.011

Phielix, C., Prins, F., & Kirschner, P. (2010b). Group Awareness of Social and Cognitive Behavior in a CSCL Environment. In Gomez, K., Lyons, L., & Radinsky, J. (Eds.), *Proceedings of the 9th International Conference of the Learning Sciences* (ICLS 2010) - Volume 1 (pp. 230-237). Chicago IL: International Society of the Learning Sciences. https://repository.isls.org//handle/1/2676

Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior*, *27*(3), 1087-1102. https://doi.org/10.1016/j.chb.2010.06.024

Prins, F. J., Sluijsmans, D. M., & Kirschner, P. A. (2006). Feedback for general practitioners in training: Quality, styles, and preferences. *Advances in Health Sciences Education, 11*(3), 289. https://doi.org/10.1007/s10459-005-3250-z

Purchase, H., & Hamer, J. (2018). Peer-review in practice: eight years of Aropä. *Assessment & Evaluation in Higher Education*, *43*(7), 1146-1165. https://doi.org/10.1080/02602938.2018.1435776

Quanjel, I., & Van Boxtel, F. (2021). *Modulehandleidng Beweeginterventies op Maat*. Nijmegen: HAN University of Applied Sciences.

Sangin, M., Molinari, G., Nüssli, M. A., & Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior, 27*(3), 1059-1067. https://doi.org/10.1016/j.chb.2010.05.032

Schmidt, K. (2002). The problem with awareness': Introductory remarks on awareness in CSCW'. *Computer Supported Cooperative Work (CSCW)*, *11*(3), 285-298. https://doi.org/10.1023/A:1021272909573

Schnaubert, L., & Bodemer, D. (2019). Providing different types of group awareness information to guide collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, *14*(1), 7-51. https://doi.org/10.1007/s11412-018-9293-y

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize?. *Journal of Research in Personality*, *47*(5), 609-612. https://doi.org/10.1016/j.jrp.2013.05.009

Shin, Y., Kim, D., & Jung, J. (2018). The effects of representation tool (visible-annotation) types to support knowledge building in computer-supported collaborative learning. *Journal of Educational Technology & Society*, *21*(2), 98-110. http://www.jstor.org/stable/26388383

Soller, A., Martínez, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of Artificial Intelligence in Education*, *15*(4), 261-290. https://content.iospress.com/articles/international-journal-of-artificial-intelligence-in-education/jai15-4-02

Stahl, G. (2007). Meaning Making in CSCL: Conditions and Preconditions for Cognitive Processes by Groups. In Chinn, C. A., Erkens, G., & Puntambekar, S. (Eds.), *The Computer Supported Collaborative Learning (CSCL) Conference 2007*, Volume 8, Part 2 (pp. 651-660). New Brunswick, NJ, USA: International Society of the Learning Sciences. https://repository.isls.org//handle/1/3424

Stahl, G. (2016). From intersubjectivity to group cognition. *Computer Supported Cooperative Work (CSCW)*, *25*(4), 355-384. https://doi.org/10.1007/s10606-016-9243-z

Stahl, G., & Hakkarainen, K. (2021). Theories of CSCL. *International handbook of computer-supported collaborative learning*, 23-43. https://link.springer.com/chapter/10.1007/978-3-030-65291-3_2

Suthers, D. D., & Seel, N. M. (2012). Computer-supported collaborative learning. *Encyclopedia of the sciences of learning*, 719-722. https://doi.org/10.1007/978-1-4419-1428-6_389

Tipton, E., Hallberg, K., Hedges, L. V., & Chan, W. (2017). Implications of small samples for generalization: Adjustments and rules of thumb. *Evaluation review*, *41*(5), 472-505. https://doi.org/10.1177/0193841X16655665

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and instruction*, *17*(4), 416-426. https://doi.org/10.1016/j.learninstruc.2007.03.007

Wu, Y., & Schunn, C. D. (2020). From feedback to revisions: Effects of feedback features and perceptions. *Contemporary Educational Psychology*, *60*, 101826. https://doi.org/10.1016/j.cedpsych.2019.101826

Yamada, M., Goda, Y., Yasunami, S., Matsukawa, H., & Hata, K. (2016). A computer-supported collaborative learning design for quality interaction. *Ieee Multimedia, 23*(1), 48–59. https://doi.org/10.1109/MMUL.2015.95

Zhu, W. (1995). Effects of training for peer response on students' comments and interaction. *Written communication*, *12*(4), 492-528. https://doi.org/10.1177/0741088395012004004

Zong, Z., Schunn, C. D., & Wang, Y. (2021). What aspects of online peer feedback robustly predict growth in students' task performance?. *Computers in Human Behavior*, *124*, 106924. https://doi.org/10.1016/j.chb.2021.106924

Appendices

Appendix 1a: CGA Questionnaire Pre-Test

Algemene vragen

Hieronder worden een aantal algemene vragen gesteld.

Wat is je leeftijd?

Wat is je geslacht?

- Man
- Vrouw
- Zeg ik liever niet

In welke klas zit je?

Domein Kennis

Hieronder worden een paar vragen gesteld om je kennis te testen. Onthoud dat dit niet meetelt voor je cijfer en een momentopname is.

1. Sport in de wijk: Wat is belangrijk bij sport in de wijk?
   - Houd het laagdrempelig en vrijblijvend, en breng het aanbod naar de mensen toe.
   - Houd het laagdrempelig en zorg voor verplichting, en breng het aanbod naar de mensen toe.
   - Houd het laagdrempelig en vrijblijvend, laat de mensen naar jou toe komen.
2. Buurtsportcoach: Wat is/zijn de belangrijkste taken van de buurtsportcoach?
   - Gezond, vitaal en fit worden zijn belangrijk. Sport wordt voornamelijk ingezet als doel.
   - Maatschappelijke problemen zoals eenzaamheid en lage sociale cohesie zijn belangrijk. Sport wordt voornamelijk ingezet als middel.
   - Sport wordt zowel als doel gezien en als middel ingezet.
3. Achterstandswijken: Wat zijn typische achterstandsdoelgroepen die je tegenkomt in achterstandswijken en welke maatschappelijke problemen komen veelal voor in achterstandswijken?
   - Probleemjongeren en langdurig werklozen | criminaliteit en sociaal isolement van ouderen
   - Kinderen/jongeren met overgewicht en actieve jonge senioren (50 - 65 jaar) | sociaal isolement van ouderen en hangjongerenproblematiek
   - Allochtone jongens (12 - 18 jaar) en asielzoekers (18 - 40 jaar) of vluchtelingen | sociaal isolement van ouderen en veel laagopgeleiden

4. Wijkscans: Wat is belangrijke informatie dat moet worden opgenomen in de interne en externe analyse van de wijkscan?

- Informatie over de leefbaarheid in de wijk (intern) en informatie over de problematiek in de wijk (extern)
- Citaten uit het interviews (intern) en DESTEP (extern)
- Landelijke literatuurgegevens over het beweeggedrag (intern) en landelijk beleid ten aanzien van het probleem en de doelgroep (extern)

5. Sociale infrastructuur in de wijk: Bij de sociale infrastructuur kun je denken aan scholen, buurthuizen en kerken, maar ook maatschappelijke participatie en andere activiteiten horen daarbij. De sociale infrastructuur is ook van belang bij het opstellen van de WAP en WUP. Wat is het verschil tussen de WAP en WUP?

- De WAP is het wijkuitvoeringsplan en is gericht op een periode van 5 tot 10 jaar. De WUP is het wijkactieplan en is gericht op een periode van 1 tot 2 jaar.
- De WAP is het wijkactieplan en is gericht op een periode van 5 tot 10 jaar. De WUP is het wijkuitvoeringsplan en is gericht op een periode van 1 tot 2 jaar
- De WAP is het wijkactieplan en de WUP is het wijkuitvoeringsplan. Beide zijn ze gericht op een periode van 5 tot 10 jaar.

6. SWOT-analyse: Wat is het doel van een SWOT-analyse?

- Aan de hand van de SWOT-analyse kun je de interne en externe analyse doen.
- Met de SWOT-matrix krijg je een overzicht van de uitkomsten van de interne en externe analyse.
- De SWOT-analyse is een vervolg op de confrontatiematrix.

Kennis

Vul eerst jouw naam in en vul daarna de namen van je groepsgenoten in op alfabetische volgorde. Geef vervolgens voor de stellingen aan in hoeverre deze van toepassing zijn op jou en je groepsgenoten.

Naam:

Naam groepsgenoot 1:

Naam groepsgenoot 2:

Eventueel naam groepsgenoot 3 (mocht je geen 3e groepsgenoot hebben, laat dit dan leeg):

1. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben over wat belangrijk is bij sport in de wijk?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

2. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben over wat de functie van buurtsportcoach inhoudt?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

3. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben van de kenmerken van achterstandswijken?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

4. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben van wat een wijkscan inhoudt?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0

Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

5. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben van wat onderdeel is van de sociale infrastructuur en wat de WAP en WUP inhouden?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

6. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben van het doel/nut van een SWOT-analyse?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

Vaardigheden

Herhaal hier eerst jouw naam in en vul daarna de namen van je groepsgenoten weer in op alfabetische volgorde. Zo hoef je niet terug te gaan naar de vorige sectie als je de volgorde vergeet. Geef vervolgens voor de stellingen aan in hoeverre deze van toepassing zijn op jou en je groepsgenoten.

Naam:

Naam groepsgenoot 1:

Naam groepsgenoot 2:

Naam groepsgenoot 3 (mocht je geen 3e groepsgenoot hebben, laat dit dan leeg):

1. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het analyseren en interpreten van informatie uit relevante bronnen?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

2. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het maken van verbindingen tussen de resultaten uit onderzoek en de ontwikkeling van het sport- en beweegaanbod?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

3. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het ontwerpen en ontwikkelen van het sport- en beweegaanbod op basis van onderzoek naar de wensen, mogelijkheden en beperkingen van de doelgroep?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

4. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het vertalen van het beleid van de organisatie naar concrete sport- en beweegprogramma's?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

5. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het doen van gestructureerd onderzoek naar de kwaliteit en effectiviteit van het sport- en beweegaanbod?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

6. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het, op een verantwoorde wijze, begeleiden, coachen en adviseren van doelgroepen binnen het sport- en beweegaanbod?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

7. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het afstemmen van de planning, organisatie en begeleiding van sport- en beweegprogramma's met relevante partijen?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

8. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het samenwerken in een multidisciplinair verband?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

Appendix 1b: CGA Questionnaire Post-Test

Post-test

De post-test werkt volgens hetzelfde principe als de pre-test. Eerst worden vragen gesteld over je eigen kennis, deze tellen wederom niet mee voor je cijfer. Daarna worden vragen gesteld over de kennis in jouw groep van het wijkadvies, gevolgd door vragen over de vaardigheden in jouw groep.

Domein Kennis

Hieronder worden een paar vragen gesteld om je kennis te testen. Onthoud dat dit niet meetelt voor je cijfer en een momentopname is.

1. Model van Brug: Het model van brug bevat verschillende stappen. Welke optie geeft de goede volgorde aan van de stappen?
   - Analyse van gedrag, Analyse van gezondheidsproblemen, Analyse van determinanten van gedrag, Interventie-implementatie en -disseminatie, Interventieontwikkeling.
   - Analyse van determinanten van gedrag, Analyse van gedrag, Analyse van gezondheidsproblemen, Interventieontwikkeling, Interventie-implementatie en -disseminatie.
   - Analyse van gezondheidsproblemen, Analyse van gedrag, Analyse van gedragsdeterminanten, Interventieontwikkeling, Interventie-implementatie en -disseminatie.
2. Gedragsdeterminanten: Wat zijn persoonlijke gedragsdeterminanten?
   - Gedragsdeterminanten zijn factoren die het gedrag van iemand kunnen beïnvloeden, zoals attitude.
   - Gedragsdeterminanten zijn factoren die de wijk kunnen beïnvloeden, zoals externe omgeving

- Gedragsdeterminanten zijn factoren die de openbare ruimte kunnen beïnvloeden, barrières.

3. Bewegen in de openbare ruimte: Wat valt NIET onder bewegen in de openbare ruimte?
   - Bewegen in een aangelegd plein met buitenfitnessapparaten
   - Mountainbiken in het bos
   - Freerunning in de gymzaal

4. Judo- en karate aanpak: Wat is het verschil tussen de judo- en karate aanpak?
   - De judo aanpak is de harde aanpak en is meer confronterend, bij de karate-aanpak maak je gebruik van de eigenschappen van de straatjongeren en is meer gelijkwaardig.
   - De aanpakken zijn beide hard, er is geen verschil
   - De judo aanpak maakt gebruik van de eigenschappen van straatjongeren en is meer gelijkwaardig, de karate-aanpak is de harde aanpak en is meer confronterend.

5. Maatschappelijk verantwoord ondernemen (MVO): MVO maakt niet gebruik van de klassieke benadering, maar van de triple-P-benadering. Waar staan de P's voor?
   - People, Planet en Profit
   - Planet, Plants en Participation
   - Profit, Planet en Participation

6. Participatieladder: Wat is het hoogste niveau van participatie op de participatieladder van Pretty?
   - Participatie via informatie
   - Zelf-mobilisatie
   - Interactieve participatie

Kennis

Geef hieronder eerst jouw naam en daarna de namen van je groepsgenoten in alfabetische volgorde. Geef vervolgens voor de volgende stellingen aan in hoeverre deze van toepassing zijn op jou en je groepsgenoten.

Mijn naam:

Naam groepsgenoot 1:

Naam groepsgenoot 2:

Naam groepsgenoot 3 (mocht je geen derde groepsgenoot hebben, laat dit dan leeg):

1. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben over het model van Brug?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

2. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben over gedragsdeterminanten?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

3. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben over bewegen in de openbare ruimte?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

4. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben over de judo- en karate aanpak?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

5. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben over maatschappelijk verantwoord ondernemen?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

6. Hoeveel kennis denk jij dat jij en je groepsgenoten hebben over de participatieladder?

	Erg weinig	Minder dan gemiddeld	Gemiddeld	Meer dan gemiddeld	Erg veel
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

Vaardigheden

Herhaal hier eerst jouw naam in en vul daarna de namen van je groepsgenoten weer in op alfabetische volgorde. Zo hoef je niet terug te gaan naar de vorige sectie als je de volgorde vergeet. Geef vervolgens voor de stellingen aan in hoeverre deze van toepassing zijn op jou en je groepsgenoten.

Naam:

Naam groepsgenoot 1:

Naam groepsgenoot 2:

Naam groepsgenoot 3 (mocht je geen 3e groepsgenoot hebben, laat dit dan leeg):

1. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het analyseren en interpreten van informatie uit relevante bronnen?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

2. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het maken van verbindingen tussen de resultaten uit onderzoek en de ontwikkeling van het sport- en beweegaanbod?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

3. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het ontwerpen en ontwikkelen van het sport- en beweegaanbod op basis van onderzoek naar de wensen, mogelijkheden en beperkingen van de doelgroep?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

4. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het vertalen van het beleid van de organisatie naar concrete sport- en beweegprogramma's?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

5. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het doen van gestructureerd onderzoek naar de kwaliteit en effectiviteit van het sport- en beweegaanbod?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

6. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het, op een verantwoorde wijze, begeleiden, coachen en adviseren van doelgroepen binnen het sport- en beweegaanbod?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

7. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het afstemmen van de planning, organisatie en begeleiding van sport- en beweegprogramma's met relevante partijen?

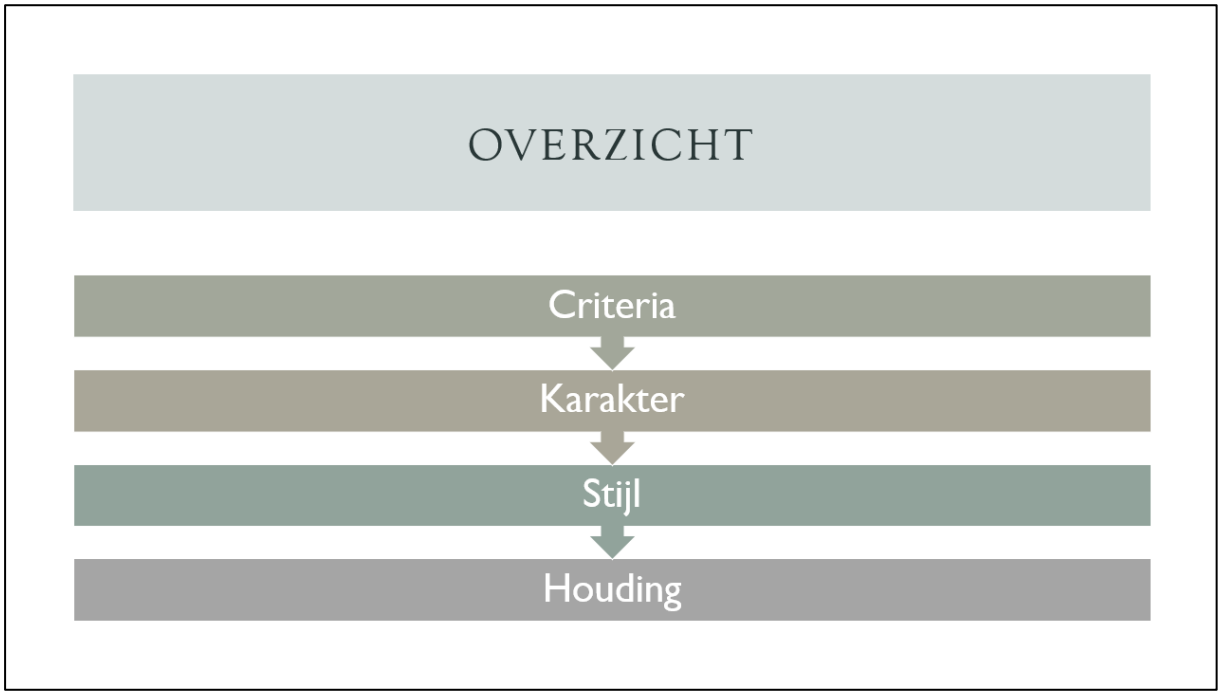
	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

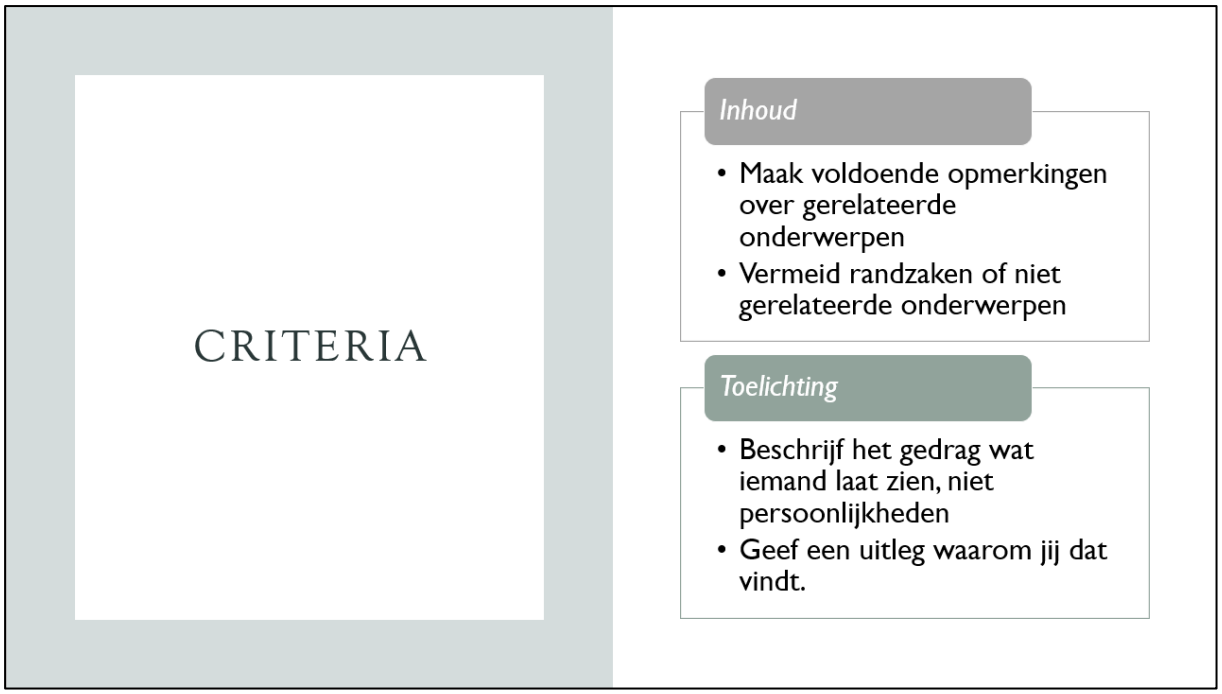
8. Hoe goed denk jij dat jij en jouw groepsgenoten zijn in het samenwerken in een multidisciplinair verband?

	Helemaal niet goed	Onder gemiddeld	Gemiddeld	Boven gemiddeld	Heel goed
Ik	0	0	0	0	0
Groepsgenoot 1	0	0	0	0	0

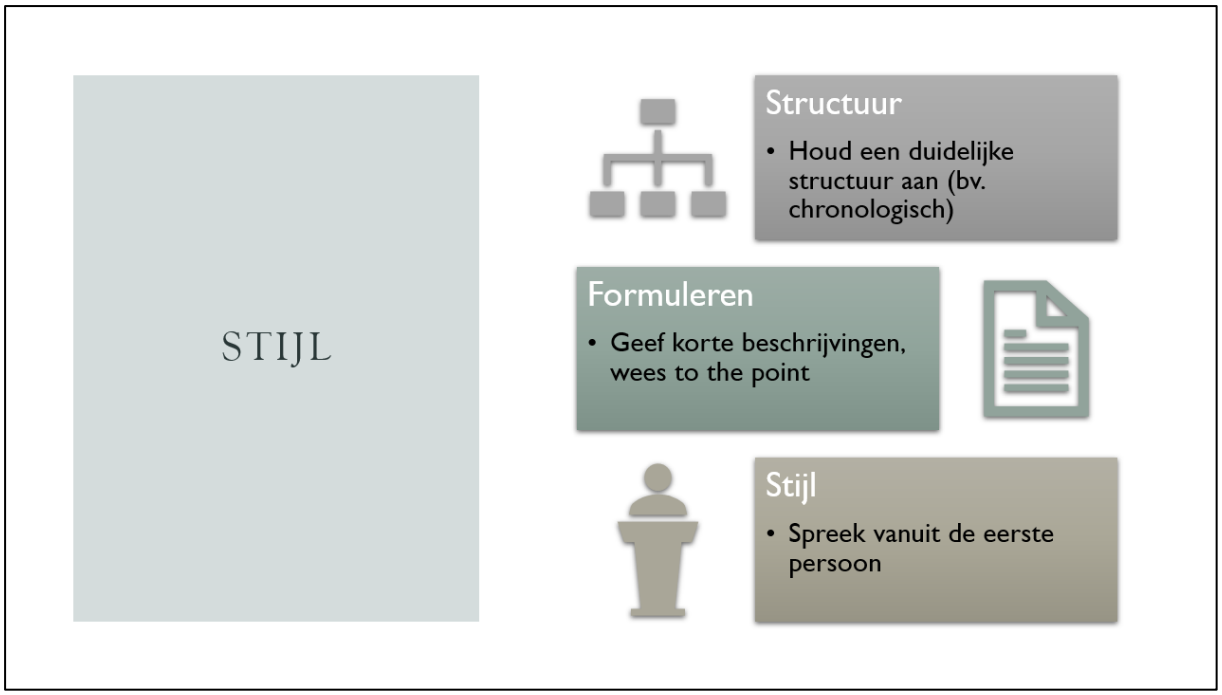
Groepsgenoot 2	0	0	0	0	0
Groepsgenoot 3	0	0	0	0	0

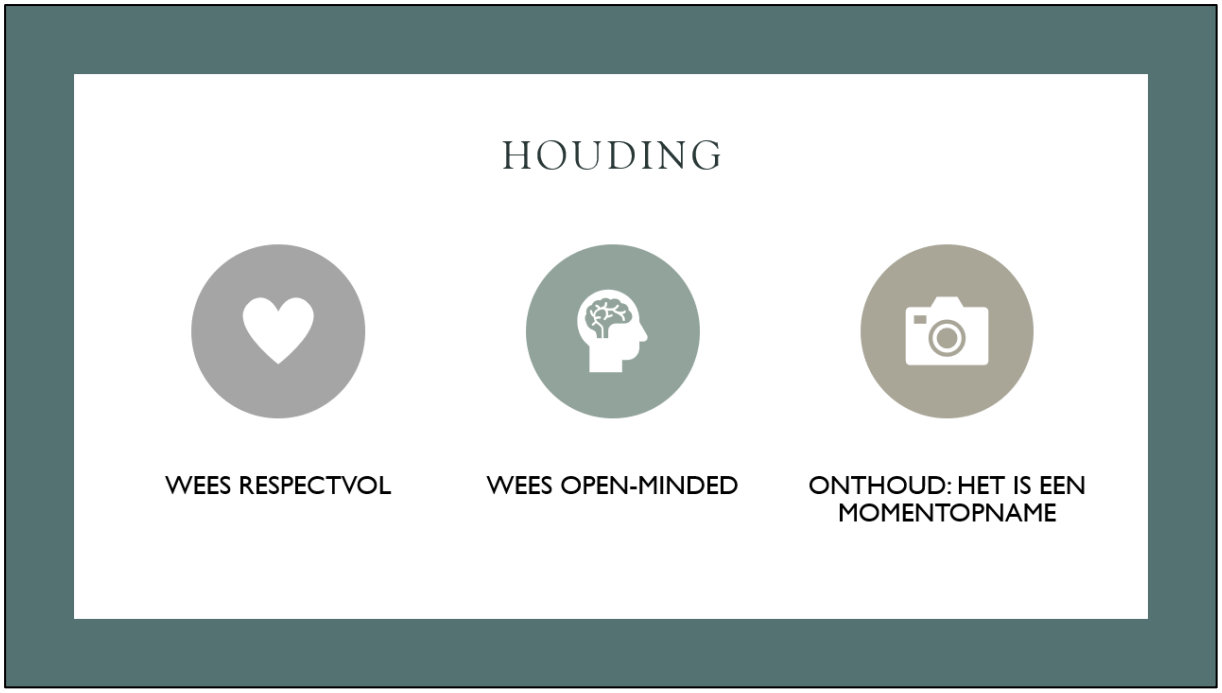
Appendix 2: Peer Feedback Presentation

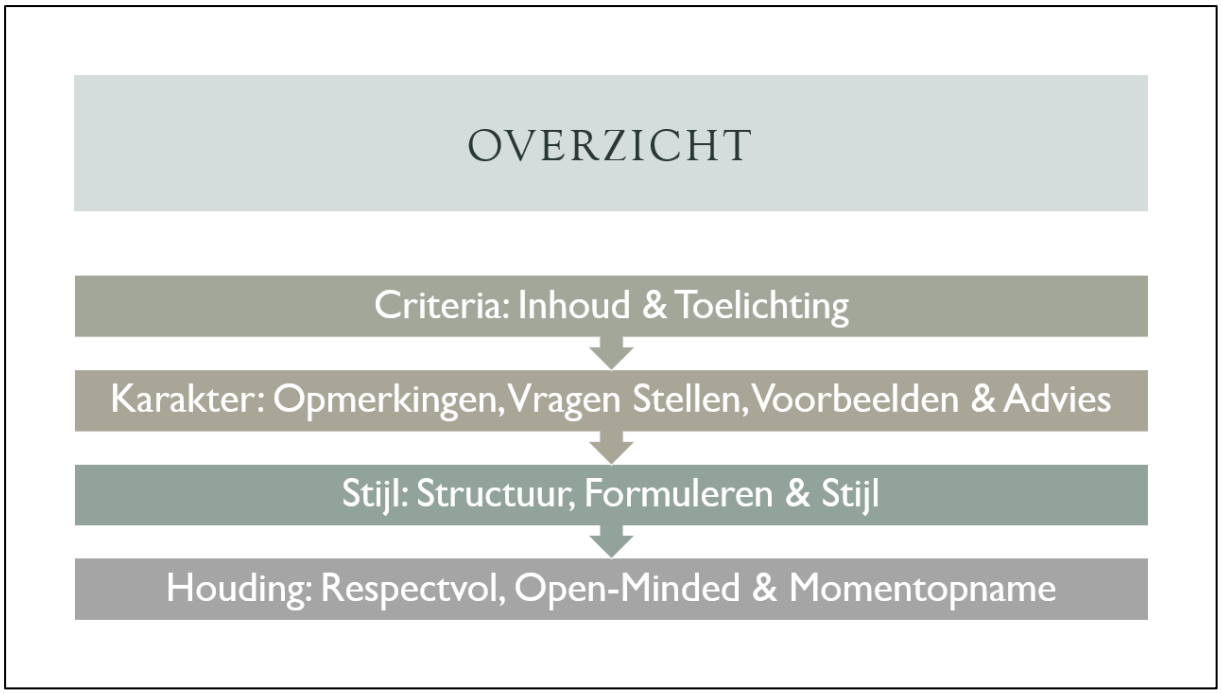



CRITERIA
Inhoud
• Maak voldoende opmerkingen over gerelateerde onderwerpen
• Vermeid randzaken of niet gerelateerde onderwerpen
Toelichting
• Beschrijf het gedrag wat iemand laat zien, niet persoonlijkheden
• Geef een uitleg waarom jij dat vindt.


KARAKTER
Opmerkingen
• Zorg voor een goede balans tussen positieve en negatieve punten
Vragen stellen
• Stel vragen om een reflectie op te roepen
Voorbeelden
• Gebruik voorbeelden vanuit je eigen ervaring (externe voorbeelden)
Advies
• Geef goede en duidelijke suggesties, ofwel constructief advies


STIJL
Structuur
• Houd een duidelijke structuur aan (bv. chronologisch)
Formuleren
• Geef korte beschrijvingen, wees to the point
Stijl
• Spreek vanuit de eerste persoon


HOUDING
WEES RESPECTVOL
WEES OPEN-MINDED
ONTHOUD: HET IS EEN MOMENTOPNAME



Appendix 3: Peer Feedback Quality Check

Peer Feedback Quality Check

Zojuist hebben jullie elkaar peer feedback gegeven. In deze vragenlijst zullen vragen gesteld worden over de kwaliteit van de peer feedback op verschillende vlakken. Deze antwoorden kan alleen de onderzoeker inzien.

Wat is je naam?

In welke klas zit je?

Criteria

De volgende vragen gaan over de algemene criteria van de peer feedback. Een 1 staat voor een minimale uitvoering; het wordt niet of nauwelijks toegepast. 2 staat voor een gemiddelde uitvoering; het wordt soms of deels toegepast. 3 staat voor een goede uitvoering; het wordt vrijwel altijd toegepast.

1. Inhoud - Er werden voldoende opmerkingen gemaakt over gerelateerde onderwerpen (onderwerpen die voorbijkwamen in de pre-test).

Minimaal uitgevoerd	1	2	3	Goed uitgevoerd

2. Toelichting - Het gedrag werd beschreven en er werd een uitleg gegeven.

Minimaal uitgevoerd	1	2	3	Goed uitgevoerd

Karakter en aard

De volgende vragen gaan over het karakter en de aard van de peer feedback. Een 1 staat voor een minimale uitvoering; het wordt niet of nauwelijks toegepast. 2 staat voor een gemiddelde uitvoering; het wordt soms of deels toegepast. 3 staat voor een goede uitvoering; het wordt vrijwel altijd toegepast. Deze nummering is lijdend, tenzij anders wordt aangegeven.

3. Opmerking - Er was een goede balans tussen positieve en negatieve opmerkingen. 1 = negatief overheerst, 2 = positief overheerst, 3 = balans

Negatief overheerst	1	2	3	Balans

4. Gestelde vragen - Er werden voldoende vragen gesteld die een reflectie opriepen.

Minimaal uitgevoerd	1	2	3	Goed uitgevoerd

5. Gebruik van voorbeelden - Er werden voldoende externe voorbeelden (bv. eigen ervaringen) gebruikt.

Minimaal uitgevoerd	1	2	3	Goed uitgevoerd

6. Advies - Er werden goede en duidelijke suggesties gegeven, ofwel er werd constructief advies gegeven.

Minimaal uitgevoerd	1	2	3	Goed uitgevoerd

Stijl

De volgende vragen gaan over de stijl van de peer feedback. Een 1 staat voor een minimale uitvoering; het wordt niet of nauwelijks toegepast. 2 staat voor een gemiddelde uitvoering; het

wordt soms of deels toegepast. 3 staat voor een goede uitvoering; het wordt vrijwel altijd toegepast. Deze nummering is lijdend, tenzij anders wordt aangegeven.

7. Structuur - De feedback werd in een duidelijke structuur gegeven (bv. chronologische volgorde).

Minimaal uitgevoerd	1	2	3	Goed uitgevoerd

8. Formuleren - Er werden korte beschrijvingen gegeven. 1 = alleen steekwoorden, 2 = deels steekwoorden en deels korte beschrijvingen, 3 = korte beschrijvingen.

Alleen steekwoorden	1	2	3	Korte beschrijvingen

9. Stijl - Er werd vanuit de eerste persoon gesproken.

Minimaal uitgevoerd	1	2	3	Goed uitgevoerd

Appendix 4: Assessment Rubric

Voorwaardelijke criteria		**(Presentatie)Wijkadvies**
De presentatie van het wijkadvies begint met een beknopte samenvatting van de interne en externe analyse van de gekozen wijk en de HAP wordt benoemd		
De presentatie van het wijkadvies bevat zowel het advies als de aanbevelingen en de verplichte bijlagen.		
De bronvermeldingen en de literatuurlijst zijn conform de APA-normen opgenomen in de wijkscan en in de presentatie		

De tekst in de wijkscan en de taal binnen de presentatie is vrij van een taalalarm		
Een eventuele herkansing wordt aangeleverd in een rode kleur tekst.		
	Advies en conclusies	
Maakt de verbinding tussen de resultaten uit onderzoek en de ontwikkeling van de sport- en beweegaanbod	Bevat een verantwoording van het wel of niet gebruiken van bestaande succesvolle interventies in aansluiting op de gekozen HAP	
	De gekozen interventie of zelfontworpen interventie is concreet weergegeven	
	Het is helder op welke gedragsdeterminant wordt ingestoken om tot gedragsverandering te komen binnen de doelgroep in relatie tot het gesignaleerde probleem.	
	Binnen het advies is aandacht voor het bevorderen van een beweegvriendelijke omgeving binnen de wijk.	
	Er wordt gepresenteerd hoe MVO tot stand is gekomen binnen het wijkadvies	
	Het advies is gebaseerd op de bevindingen in de literatuur en op de bevindingen in het veld (interviews en observaties).	
Ontwerpt en ontwikkelt het sport- en beweegaanbod op basis van onderzoek naar de wensen, mogelijkheden en beperkingen van de doelgroep;	De vraagstelling van het onderzoek wordt beantwoord	
	Het advies is realistisch en creatief	
	Aanbevelingen	
	Bevat een stappenplan met concrete stappen die de gemeente/wijkpartners moet ondernemen om het hoofdprobleem op te lossen	
Stemt de planning, organisatie en begeleiding van sport- en beweegprogramma's af met relevante partijen;	De aanbevelingen zijn voorzien van een specifieke vertaling naar de lokale situatie	
	De aanbevelingen zijn direct implementeerbaar	
	Bevat een raming van de kosten	
	Bevat een uitleg van de wijze waarop actief burgerschap tot stand komt	
	Binnen de aanbevelingen wordt aandacht besteed aan het bevorderen van de leefbaarheid en/ of veiligheid in de wijk	

	Bevat een toelichting van de wijze waarop de samenwerkingspartners vanuit de netwerkanalyse betrokken worden bij het plan en wie wanneer wat moet doen.	
	Het is duidelijk hoe de interventie geëvalueerd gaat worden	
Doet gestructureerd onderzoek naar de kwaliteit en de effectiviteit van het sport- en beweegaanbod en rapporteert hierover;	Foto's van woonomgeving en sportfaciliteiten in de wijk voorzien van onderschrift en met duidelijke link naar de vraagstelling	**Opnemen in de presentatie**

Wijkadvies (BLOK1/3)						
Onderdeel/ Punten	**1**	**2**	**3**	**4**	**Weging in %**	**Score**
	Ontbreekt of onjuist	Delen onjuist/onvolledig, of uitwerking summier	Juist en redelijk volledig	Juist en uitgebreide uitwerking/ veel diepgang		
Advies en conclusies					40	0,0
Aanbevelingen					50	0,0
	De presentatie is saai en spreekt onvoldoende tot de verbeelding.	De presentatie bevat boeiende en minder boeiende onderdelen. De presentatie bevat onvoldoende overtuigingskracht.	De presentatie bevat boeiende onderdelen. Er wordt gebruik gemaakt van beeldmateriaal, bronvermelding en het taalgebruik is in orde.	De presentatie is boeiend van begin tot eind overtuigend, kort en krachtig en bevat relevant beeldmateriaal, passend taalgebruik en juiste bronvermelding		
Presentatie					10	0,0
Totale score wijkadvies						**0,0**
Dit onderdeel telt voor 25% mee binnen Beweeginterventies						

Appendix 5: Observation Scheme

Teacher:

Group:

Questions	
Teacher	Students

Feedback	
Teacher	Students

Student #	Asked Questions	Answers Given	Notes
1			
2			
3			
4			

Appendix 6: Example Calculation

	Skill 1	Skill 2	Skill 3	Skill 4	Skill 5	Skill 6	Skill 7	Skill 8
Me	3	2	4	5	4	2	4	5
Groupmate 1	3	3	4	4	4	3	3	4
Groupmate 2	2	3	3	5	3	4	2	5
Groupmate 3	4	3	5	5	3	2	4	4

Average score from myself: (3+2+4+5+4+2+4+5)/8 = 3.625

Average score from groupmates: (((3+3+4+4+4+3+3+4)/8) + ((2+3+3+5+3+4+2+5)/8) + ((4+3+5+5+3+2+4+4)/8)) / 3 = 3.542

Skill GA: Average score from groupmates – average score from myself = 3.542 - 3.652 = -.083

Then the scores were turned into absolute scores, so the skill GA in this example is .083.