A Study on Teacher Characteristics and the Use of CSCL Dashboards

Date: June 8, 2020
Laura Gerrits - 5728738
Supervisor: Dr. Anouschka van Leeuwen
Word count: 7995

Abstract

CSCL dashboards are the best way to present student interactions during collaborative learning. Monitoring student interactions on CSCL dashboards is important for teachers to provide appropriate support. Previous research suggests that two teacher characteristics could relate to the use of CSCL dashboards: data literacy and experience with collaboration. Literature has shown that there might be a relationship between data literacy and the use of CSCL dashboards because high data literacy ensures proper ‘data evaluation’ and ‘data application’. Moreover, there are indications that experience with collaboration might moderate the relationship between data literacy and the use of CSCL dashboards. However, these teacher characteristics remain mostly unexamined. This study aims to investigate the relationship between data literacy and experience of collaboration on the use of CSCL dashboard. Forty student teachers and teachers participated in the study in which they completed questionnaires and were presented with eight dashboard situations. Two multiple regression analyses and four moderation analyses were performed to analyse the data. However, no significant relationships were found. Follow-up research could investigate other factors that underlie the use of CSCL dashboards, such as teachers' attitude towards computers.

Keywords: CSCL dashboards, teacher dashboards, collaborative learning, data literacy, experience with collaboration

In recent years, there has been a growing interest in Computer-Supported Collaborative Learning (CSCL) (Stahl, Koschmann, & Suthers, 2006). CSCL refers to "the activity of peers interacting with each other for the purpose of learning and with the support of ICT" (Suthers, 2012, p. 1). During CSCL, it is essential that teachers have an accurate representation of their students, so they can better decide which help the students need to stimulate effective collaboration (Kaendler, Wiedmann, Rummel, & Spada, 2015). The interactions between students can be monitored through CSCL dashboards, where teachers are provided with information about their collaborating students (Van Leeuwen, Rummel & Van Gog, 2019).

Monitoring CSCL dashboards requires teachers to use different skills and competencies to use them properly (Wasson & Hansen, 2015). CSCL dashboards usually consist of data widgets, and a skill that relates to working with and interpreting data is 'data literacy'. Data literacy refers to collecting, managing, evaluating and applying data (Ridsdale et al., 2015). It could lead to better use of CSCL dashboards because teachers with high levels of data literacy can better interpret the data displayed on the CSCL dashboards. Furthermore, experience with collaboration could improve the way teachers monitor student interactions on CSCL dashboards. Teachers who possess more experience with collaboration tend to have better-developed knowledge structures about good collaboration (Peterson & Comeaux, 1987). As a result, they have a better understanding of what to look for while monitoring and assessing the classroom (Van den Bogert, van Bruggen, Kostons, & Jochems, 2014). In terms of CSCL dashboards, this implies that teachers have a better understanding of what data they should look for.

To the best of our knowledge, the relationship between data literacy and how teachers use CSCL dashboards has not been studied thoroughly. More insight into this relationship can be obtained by examining it in a controlled setting. Furthermore, the relationship between

experience with collaboration and CSCL dashboards remains mostly unexamined. The results may be relevant if it appears that data literacy and experience with collaboration have a substantial impact on the use of CSCL dashboards. These results would imply that schools need to be more aware of those skills and experiences when applying CSCL, and facilitate opportunities for teachers to learn more about it.

Collaborative learning

A method teachers often use in classrooms is collaborative learning (Fawcett & Garton, 2005). Collaborative learning entails reaching a mutual understanding with others, in which group members are involved in the same goals and tasks (Järvela et al., 2015). Research by Moshman and Geil (1998) and Underwood, Underwood, and Wood (2000) has shown that children who work on a task together achieve a higher learning performance than children who perform the task alone. Besides improving the learning performance, it is useful to prepare children for the labour market, in which they often have to apply their collaborative skills when working in teams (De Lisi & Golbeck, 1999).

Nonetheless, good collaboration is not self-evident and several problems can arise during collaborative learning: dominance by one of the students, getting stuck on an assignment, trial-and-error behaviour, off-task behaviour, ineffective discussion and alternating tasks. If one of these six cases occur in a group, the group might face a problem (Van Leeuwen et al., 2019). First, collaboration can be disrupted by dominance of one of the students. When a student takes a dominant role in the group, other members of the group cannot provide input, and therefore, no discussion can take place (Schmitz & Winskel, 2008).

Second, disruption of collaborative learning can be caused by the number of attempts it takes to complete a task. According to Fawcett and Garton (2005), collaborating students search

for logical coherence and make deliberate attempts. Many attempts per assignment may indicate that students get stuck on the assignment.

Third, trial-and-error behaviour is ineffective for collaboration. The number of attempts to complete a task, combined with the speed of their attempt, can indicate trial-and-error behaviour (Van Leeuwen et al., 2019). This behaviour, where students do not discuss the task and do not gain an understanding of the material together, would interfere with identifying a students' true ability (Oh, 2004; Van Leeuwen et al., 2019).

Fourth, Schmitz and Wiskel (2008) suggest that too many off-task conversations would disturb collaboration. During these conversations, students discuss, for example, their personal matters, which are not conducive to the learning performance (Arvaja, Häkkinen, Rasku-Puttonen, & Eteläpelto, 2002).

Fifth, ineffective discussions do not contribute to collaboration between students. These discussions occur when students express their opinions without considering each other's ideas. This form of discourse is considered one of the least valuable forms since it does not promote critical problem solving (Mercer, 1996; Schmitz & Winskel, 2008).

Lastly, collaborative learning is not promoted when students alternate tasks. For the learning process, it is essential that students have the opportunity to compare their understanding with others, to integrate the different perspectives (Kruger, 1993).

CSCL and CSCL dashboards

With the arrival of computers, a new phenomenon has emerged: Computer-Supported Collaborative Learning (CSCL) (Zurita & Nussbaum, 2004). Suthers (2012) defines this concept as the activity of students who collaborate with the support of information and communication technologies (ICT). During CSCL, it is essential that teachers have an accurate representation of

the student interactions to provide the right support (Kaendler et al., 2015). However, teachers tend to have difficulties seeing the collaboration processes between students, and they typically only look at the final product of the groups' activity (Martinez Maldonado, Kay, Yacef, & Schwendimann, 2012).

Therefore, the progress of students and the interactions between students can be best presented by teacher dashboards. Teacher dashboards display student information and help teachers to follow their progress (Verbert et al., 2014). Teacher dashboards focused on collaborative learning are called CSCL dashboards. Van Leeuwen et al. (2019) define CSCL dashboards as: "visual displays that provide teachers with information about their collaborating students to aid teachers in monitoring their students' progress in CSCL settings" (p. 262). CSCL dashboards are helpful to teachers because they ensure that the collaboration processes between students are made visible; these are usually hard to see without a dashboard. However, the use of a CSCL dashboard proves to be difficult in practice, because there is not much time for teachers to look at the dashboard, due to the dynamic nature of the classroom during collaborative learning (van Leeuwen & Rummel, 2020). Therefore, teachers must be able to use a CSCL dashboard properly. Otherwise, the dashboard is no longer an aid, but an obstacle. Several studies suggested that two teacher characteristics could relate to monitoring student interactions on CSCL dashboards. These teacher characteristics are data literacy and experience with collaboration, which will be elaborated upon in the sections below.

Data literacy

CSCL dashboards provide large amounts of data through data widgets that are helpful for teachers (Charleer, Klerkx, & Duval, 2014). Data widgets can display information in the form of tables, graphs, lists, cross tabs and more (Pourshahid, 2015). Therefore, monitoring CSCL

dashboards and, thus, dealing with data widgets requires new skills and competencies (Mandinach & Gummer, 2016; Wasson, Hansen, & Netteland, 2016). Since there are large amounts of data and different sources, learning to read and interpret this data from the data widgets is essential for teachers. Therefore, researchers believe that this should be promoted (Ridsdale et al., 2015). The ability to read and interpret data is best described by the term data literacy. According to Chinien and Boutin (2011), data literacy has been recognized as one of the most beneficial skills for the 21st century as it contributes to a valuable knowledge-based economy. There even are authors who believe that data literacy should merge with the 21st century skill 'information literacy' (Prado & Marzal, 2013; Ridsdale et al., 2015). Furthermore, Mandinach (2012) states that data literacy is needed to use data effectively and to transform data into actionable instructional strategies.

Data literacy is a complex concept since its definition overlaps with other terms such as 'data information literacy', 'information literacy', 'data management literacy', and 'assessment literacy'. The difficulty with the concept is that the term is used inconsistently and unsystematic in the literature (Ridsdale et al. 2015). For example, Prado and Marzal (2013) explain the term as accessing, interpreting, managing, handling and ethically using data and Johnson (2012) as the ability to process, sort and filter large amounts of information. Ridsdale et al. (2015) analysed and summarized all existing definitions of data literacy to "the ability to collect, manage, evaluate, and apply data, in a critical manner" (p. 11). However, the CSCL dashboard itself already has a significant role in collecting and managing data.

Therefore, in this study, data literacy is defined as the ability to evaluate and apply data in a critical manner. The two knowledge areas in this study are data evaluation and data application (see Appendix A). The competencies, knowledge, and tasks described by Ridsdale et al. (2015)

that apply to CSCL dashboards are included in this study. In the present study, data evaluation is the ability to interpret data, identify problems using data, visualize data, and make decisions based on data. Data application involves critical thinking about data and one's attitude towards data.

Teacher experience with collaboration

Besides data literacy, experience with collaboration can relate to the way teachers monitor their students on CSCL dashboards. Teachers with more experience in collaboration tend to have better-developed knowledge structures about good collaboration (Lee & Tsai, 2011; Peterson & Comeaux, 1987). Knowledge structures are defined as "the knowledge of how concepts in a domain are interrelated" (Ley, 2019, p. 333). Teachers who have better-developed knowledge structures can process more visual information, distribute their attention better and identify the relevant cues in the class (Van den Bogert et al., 2014). Thus, teachers with experience in collaboration are better able to notice the six problems that might occur in a group during collaboration, such as dominance by one of the students. Furthermore, they tend to have a better understanding of what to look for when observing and assessing the class, such as student discussions or the exchange of learning resources between students (Lee & Tsai, 2011). Given these points, it can be assumed that teachers with experience in collaboration have a better understanding of what data they should look for when working with CSCL dashboards, which results in a more effective use of CSCL dashboards.

Dealing with CSCL dashboards

Now it has been discussed which teacher characteristics may relate to the use of CSCL dashboards, it is essential to specify more precisely how teachers deal with CSCL dashboards. Van Es and Sherin (2002, 2008) refer in their studies to the 'learning to notice framework', which

contains noticing phases for teachers. The first phase is the ability to identify what is important in a particular situation and decide whether it needs further attention. Van Leeuwen et al. (2019) have named this phase *detection*. Detection in the sense of CSCL dashboards can be useful for detecting groups that might face a problem during collaboration.

The second phase *interpretation* concerns "making connections between specific events and broader principles of teaching and learning" (Van Es & Sherin, 2002, p. 245). In the interpretation phase, three dimensions can be distinguished when taking data literacy into account. Van Es and Sherin (2002) describe that the teacher's adopted attitude is vital for the support they can offer students. The interpretative attitude seems to be the most beneficial, where teachers are supposed to link learning principles to what they see and do not consider the situation as an isolated event. Furthermore, van Leeuwen et al. (2019) and van Es and Sherin (2008) recognize the importance of specificity in interpretations: an interpretation can be less or more specific. The final interpretation dimension is the use of data widgets and linking them to each other. The amount of data can be processed more thoroughly when there is high data literacy, which could lead to a richer interpretation (Ridsdale et al., 2015).

Present study

Monitoring CSCL dashboards requires different skills and competencies for teachers to use them properly. This study investigates teacher characteristics in relation to the use of CSCL dashboards. Two teacher characteristics were discussed in the literature study: data literacy and experience with collaboration.

CSCL dashboards consist of data widgets in different forms (e.g., tables, graphs and lists) (Pourshahid, 2015). As mentioned, teachers with high data literacy possess higher levels of 'data evaluation' and 'data application' (Ridsdale et al., 2015). Therefore, it can be assumed that

teachers with high levels of data literacy can better understand data widgets on the CSCL dashboards, which could lead to better detection of groups that might face a problem and a richer interpretation of CSCL dashboards. Data literacy could, therefore, predict the use of CSCL dashboards to a certain extent.

Furthermore, teachers with more experience in collaboration have better-developed knowledge structures about good collaboration (Lee & Tsai, 2011). Therefore, teachers with experience in collaboration probably have a better understanding of what to look for when observing and assessing the class during collaboration, which can result in a more effective use of CSCL dashboards (Lee & Tsai, 2011; Peterson & Comeaux, 1987; Van den Bogert et al., 2014).

Based on this literature, it can be hypothesized that the relationship between data literacy and how teachers use CSCL dashboards is moderated by experience with collaboration. A teacher without experience in collaboration but with high data literacy, may identify problematic groups and interpret CSCL dashboards to a certain extent, but it can be presumed that this will be ineffective. The reason for this could be that teachers without experience in collaboration might not know which data are relevant and which data they should look for in CSCL dashboards as their knowledge structures about collaboration are less developed (Peterson & Comeaux, 1987). So, depending on whether the experience with collaboration is high or low, the relationship between data literacy and the use of CSCL dashboards could change: it could be strengthened when teachers' ideas about collaboration are more specific, and diminished when they have less specific ideas about collaboration. This implies that experience with collaboration can be a moderator.

Because little research has been done on both concepts in relation to CSCL dashboards, this study aims to investigate these relationships. Therefore, three research questions are

formulated. First, what is the relationship between teachers' data literacy and the detection of problematic groups of CSCL dashboards? Second, what is the relationship between teachers' data literacy and their interpretation of CSCL dashboards? Third, is the relationship between data literacy and how teachers use CSCL dashboards moderated by teacher experience with collaboration?

The considerations in the literature review lead to the hypothesis that it is expected that the relationship between data literacy and the detection of problematic groups is positive. Furthermore, it is expected that the relationship between data literacy and the interpretation of CSCL dashboards is positive. Finally, it is expected that teacher experience with collaboration will moderate the relationship between data literacy and how teachers use CSCL dashboards (see Figure 1).

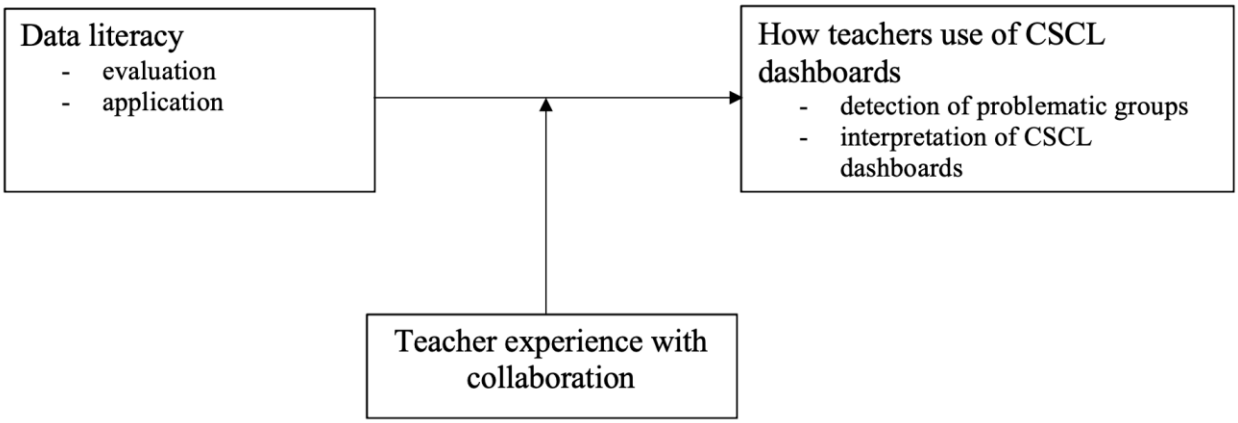


Figure 1. Conceptual Model

Method

Research design

An explanatory quantitative survey design was chosen to answer the research questions. This design is used when relationships, based on theoretical grounds, are analysed and tested. Hypotheses were drawn up through literature. In this study, only one condition was studied.

Participants

Before acquiring participants, the number of participants to reduce a type II error was determined through a power analysis with G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). A medium effect size was chosen because both data literacy and experience with collaboration were relatively unknown subjects. With a medium effect size of 0.5 and a significance level of $\alpha = .05$, two-tailed, the ideal sample size would be ≥ 29 participants. The research population consisted of PABO students (teacher program at higher vocational education), ALPO students (teacher program at university level combined with the study Educational Sciences), and teachers with a completed degree. All participants teach or have recently taught at Dutch schools. This study used the self-selected sampling method, which means participants with certain characteristics can participate voluntarily. The acquisition of participants was carried out by contacting schools and teachers.

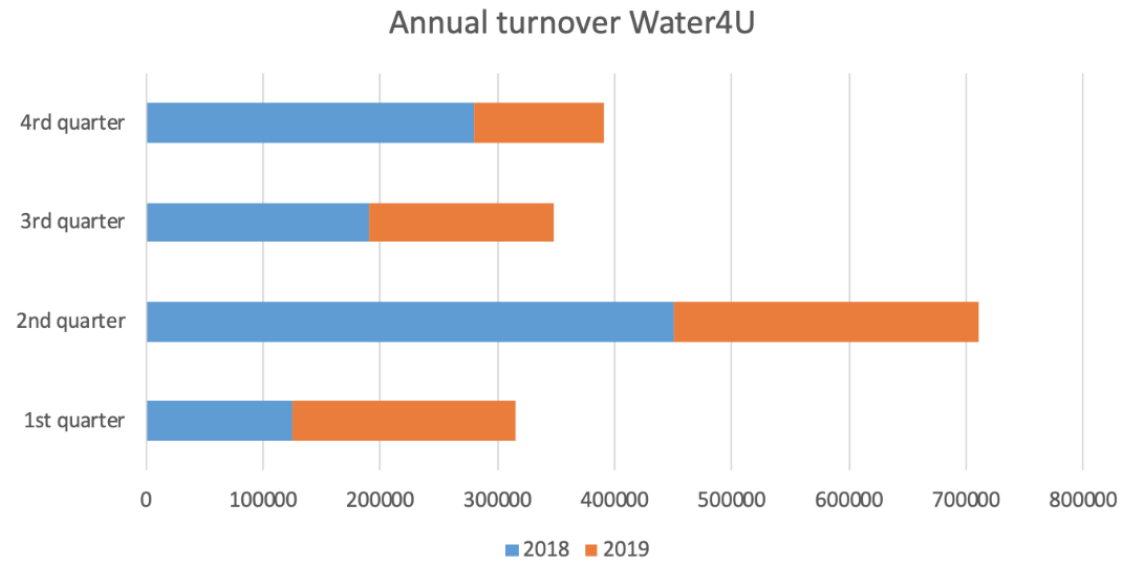
In total, 40 participants took part in this study, of which 7 men and 33 women. The age ranged from 18 to 65 years ($M = 29.47$, $SD = 11.12$). In this study, 3 PABO students, 5 ALPO students and 32 teachers with a diploma participated. Among the participants, 53.7% had more than 15 lessons of experience with collaboration and 53.7% had previous experience with teacher dashboards.

Instrumentation

Questionnaire with background information. In this questionnaire, consisting of ten items, participants were asked about their age, gender, what grade they teach, experience with fractions, experience with collaboration, and experience with dashboards to report the descriptive statistics (see Appendix B).

Questionnaire data literacy. The data literacy of the participants was tested with a questionnaire, consisting of fourteen items (see Appendix C). Since there was no existing questionnaire about data literacy, it was decided to construct a questionnaire based on the conceptual framework of Ridsdale et al. (2015). The adapted knowledge areas were data evaluation and data application (see Appendix A).

***Data evaluation*.** In the knowledge area data evaluation, the competencies were measured with a test. For each question, the participants were presented with a figure and then received a question about it. The questionnaire from Van den Bosch, Espin, Chung, and Saab (2017) about graph literacy was used as inspiration. The questions were multiple-choice, with three answer options (see Figure 2). In total, the knowledge area data evaluation consisted of eight items. The data evaluation score was calculated using the following formula: (9 * number of correct answers / number of items) + 1, so the score could range from 1 to 10.



What can be concluded from this figure?

A. Water4U performs as well in 2018 as in 2019
B. In 2019 there was more turnover in the second quarter than in 2018
C. Water4U generated more turnover in 2018 than in 2019

Figure 2. An Example Question from the Knowledge Area Data Evaluation

Data application. In the knowledge area data application, the competencies were measured with statements on a five-point Likert scale (1=completely disagree, 5=completely agree). The six statements were based on the adjusted conceptual framework of Ridgdale et al. (2015), see Appendix A. Data application consisted of critical thinking (e.g., "I am aware of the risks involved with large amounts of data, such as a data breach or privacy-sensitive datasets") and data culture (e.g., "I like working at an organization where data are used to make informed decisions"). Both competencies consisted of three items (statements 1-3 representing critical thinking, and 4-6 representing data culture). A high score on a statement indicated high application of data literacy. The average score of all statements was calculated for each participant, so the score could range from 1 to 5.

Reliability. After drawing up a first version of the questionnaire, a pilot was conducted among 51 participants, ranging from 19 to 77 years (M = 30.33, SD = 14.15). The reliability of

data evaluation was ‘poor’ (.41), and data application was ‘questionable’ (.60) (Field, 2013). Furthermore, the factor analysis of the statements did not show two factors. An improvement plan for an adjusted questionnaire with further inspection of the items was written (see Appendix I). The second pilot among 30 participants from 19 to 52 years (M = 25.83, SD = 8.57) gave a good Kuder Richardson of .66 for data evaluation and a Cronbach's alpha of .54 for data application. The factor analysis indicated the two intended factors. Before the questionnaire was used in the study, statements 2 and 3 were adjusted because the corrected item-total correlation was too low.

The definitive Kuder Richardson of data evaluation was .40, which gave reason to delete items. After removing items 1, 6 and 8, the Kuder Richardson became .53. This score was not decisive, because the questionnaire consisted of less than twenty items (Pallant, 2020). It was more important to look at the adjusted p and RIR-values (Universiteit Utrecht, z.d.). Closer inspection of the items showed that the RIR-value combined with the adjusted p-value of all remaining items was sufficient (see Appendix J). Data application showed a Cronbach's alpha of .58. After removing items 1, 2, and 3, Cronbach's alpha became .69. Again, the corrected item-total correlation was taken into account since data application consisted of less than twenty items (Pallant, 2011). These values were all greater than 0.3, which can be considered reliable (Ferketich, 1991). The factor analysis showed that the three statements loaded on one construct, which can be interpreted as the competency data culture (Ridsdale et al., 2015).

Questionnaire experience with collaboration. This questionnaire, used to measure teachers' experiences with collaboration, was based on the Lee and Tsai (2011) scale. This scale was originally an instrument on student perceptions of experience with collaboration, self-regulated learning, and information seeking in a traditional classroom setting and internet-based

learning. The validity of the instrument was good, and the reliability was satisfactory (between .84 and .91). It was decided to use the questions on student perceptions of experience with collaboration in a traditional classroom setting, and convert these student perceptions into teacher perceptions (see Appendix D). A five-point Likert scale with seven items (1=never, 5=very frequent) measured the ideas teachers had about collaborative learning (e.g., "I think it is important that students discuss problems that arise during learning with their classmates"). A high score on this scale indicated that teachers had specific ideas about collaboration since they knew what to look for when observing and assessing the class during collaborative learning, due to their developed knowledge structures about good collaboration. The average score of all items was calculated for each participant, so the score could range from 1 to 5.

Reliability. The reliability of the adjusted questionnaire from Lee and Tsai (2011) about experience with collaboration was measured using Cronbach's alpha and was .57, which can be considered questionable. Therefore, items were deleted. After removing items 3 and 7, Cronbach's alpha was .61. The corrected item-total correlation was taken into account since the questionnaire consisted of less than ten items (Pallant, 2011). Those correlations were all greater than 0.3 which can be considered reliable (Ferketich, 1991).

Dashboard trials. Dashboard trials were used to test whether data literacy and teacher experience with collaboration is related to how teachers use CSCL dashboards (Appendix E). The dashboard trials were adapted from the study by Van Leeuwen et al. (2019), in which they presented five dyads with three types of dashboards. The three types of dashboards were different in terms of support: mirroring, alerting, and advising. In this study, only the mirroring dashboard was used, where the results were displayed without support or advice being given to the participant, which made it possible to detect whether data literacy could make a difference. The

information on the dashboards was the result of dyads working together on assignments about fractions. Eight different dashboards were displayed to the participant with a description of the situation before each dashboard trail. The participant was given the task to look at the different dyads on the dashboard. The dashboard displayed six data widgets per dyad (see Figure 3). In the classroom, teachers do not have time to study a dashboard extensively due to other tasks that they have to do while teaching (van Leeuwen & Rummel, 2020). Therefore, participants were given 50 seconds per dashboard to imitate the classroom situation.

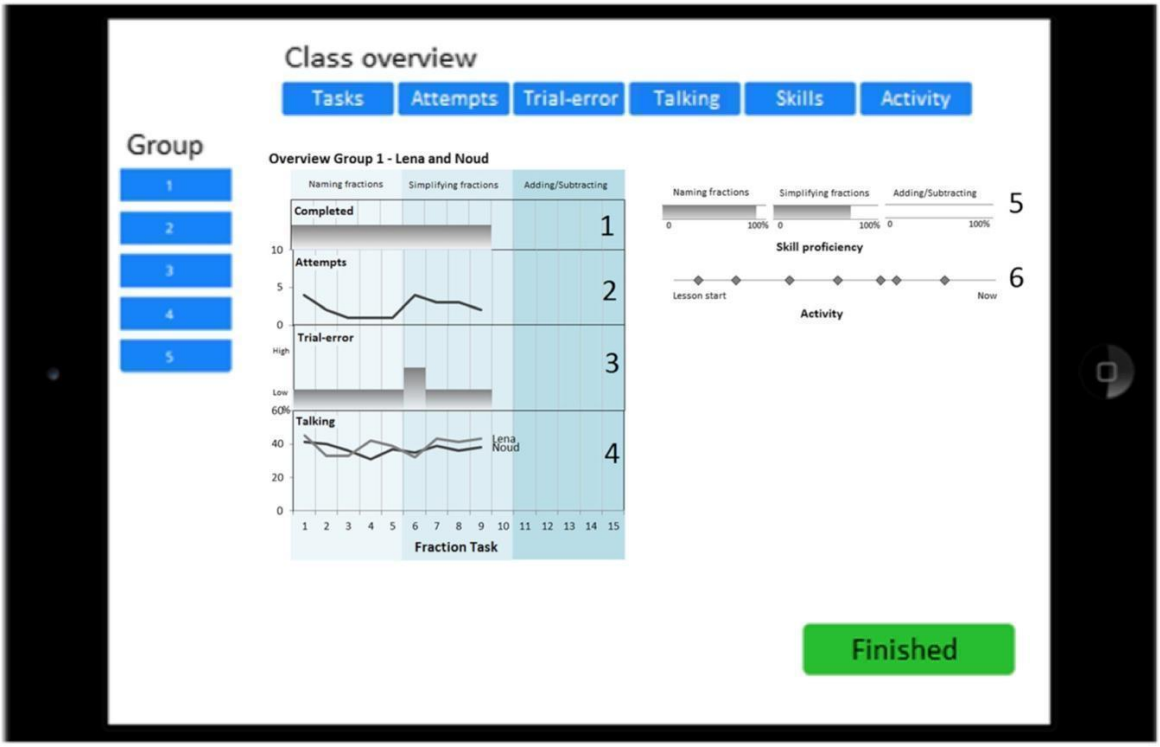


Figure 3. An Example Overview of a Dashboard Trial With Data Widgets (1 = completed assignments, 2 = attempts per assignment, 3 = chance of trial-and-error behaviour, 4 = amount of talk, 5 = skill proficiency, and 6 = activity timeline). Adapted from "What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations?" by van Leeuwen, A., et al., 2019, *International Journal of Computer-Supported Collaborative Learning*, 14(3), pp. 272-273

If participants clicked on ‘Finished’ or if their time of 50 seconds had expired, they received questions about detection and interpretation.

Detection. The participants first received one question about *detection*: “In which group do you think there was something going on in this situation?” In all dashboard trials, there was only one correct answer. As described in the literature review, there are six examples of problems that can arise in a group during collaboration. In each dashboard, there was one dyad where such a problem arose (see Appendix F). The score of all detection questions was calculated using the following formula: (9*number of correct answers/8)+1, so the score could range from 1 to 10.

Interpretation. Next, participants were asked about the *interpretation* of CSCL dashboards: “Explain what you think was the problem as specific as possible.” This open question was quantified with a coding scheme (see Appendix G). The coding scheme consisted of three dimensions: specificity, attitude, and use of data widgets. Based on research by Van Leeuwen et al. (2019), naming many specific elements results in a high specificity score. If the situation was described and linked to a learning principle, this resulted in a high attitude score. Finally, if more than two data widgets were used and associated with each other, this led to a high score on the use of data widgets. A weight was assigned to the dimensions (specificity=1, attitude=2.5, data widgets=1.66) since each dimension had a different number of points that could be allocated. As a result, the total score on interpretation was calculated with a grade: (9*(specificity+attitude+data widgets)/15)+1. This means that the range of this score can be between 1 and 10. Two assessors rated 25% of the comments on the interpretation question. There was a 75.56% agreement between the assessors, which, according to Cicchetti and Sparrow (1981), can be considered as excellent. Therefore, it was decided to have the rest of the answers coded by a single assessor.

Questionnaire about the study. Participants received a questionnaire at the end of the study, based on the study by Van Leeuwen et al. (2019), about how they experienced this study (see Appendix H). One item was added to the questionnaire: "I had the idea that I had the right skills to read the dashboard." In total, the questionnaire consisted of six items. There were four statements on a five-point Likert scale (1=completely disagree, 5=completely agree), e.g., "the investigation procedure was clear", and two open questions "can you indicate which strategy you generally used to deal with situations on the dashboard?" and "do you have any other comments about the study?"

Procedure

The investigation was supposed to take place on location. However, due to the COVID-19 outbreak, the study took place in Gorilla, a digital environment for experiments and surveys. First, all participants gave active informed consent with a form that took the ethical guidelines into account, and where the purpose, procedure, anonymity and confidentiality of the study were explained (see Appendix K). Second, they filled out the questionnaires in the correct order: background information, data literacy, and teacher experience with collaboration. Afterwards, participants were presented with eight different dashboard trials as described in the instrumentation section. After each dashboard trial, participants had to answer the detection and interpretation questions. Finally, the investigation ended with a questionnaire about the study. On average, the study took thirty minutes per participant.

Data analysis

A multiple regression analysis was performed with the score of the problematic groups as dependent variable, and data evaluation and data application as independent variables to answer the first research question. The same analysis was performed for the second research question.

Here, the similar independent variables were used, but the dependent variable consisted of the quantified interpretation score of CSCL dashboards. For the third research question, four moderation analyses were performed to determine if experience with collaboration was a moderator.

All other variables measured in this investigation were used for descriptive statistics. The analyses were all done in SPSS, where the PROCESS macro from Hayes (2017) was used for the moderation analyses.

Results

Before the analysis of the data started, several adjustments were made to the data set. The study was started by 40 participants, of which 8 participants who did not complete the study. Furthermore, one participant was removed because this participant completed the questionnaires within five minutes, which indicated that he or she may not have had serious intentions. Another participant was removed from the dataset because the study was completed on a mobile device. The dashboards cannot be optimally displayed on a mobile device, making the results unreliable. Lastly, a limit was established for former teachers. Former teachers who had not taught for more than five years were removed from the data (n = 1). As a result, the data of 29 participants were analysed.

The data was checked on univariate outliers. Boxplots concerning all continuous variables that were used to test the hypotheses were generated and indicated three outliers. Due to the sample size, it was chosen to transform the outliers to the highest/lowest non-outlier. This way, all 29 participants were retained. Table 1 shows the descriptive statistics.

Table 1

Descriptive Statistics

	n	Min	Max	*M*	*SD*
Experience with teaching fractions	29	1	5	3.76	1.41
Lessons collaborative learning	29	1	5	3.69	1.51
Data evaluation score	29	6.40	10.00	9.01	1.41
Data application score	29	1.33	4.33	3.07	.72
Experience with collaboration	29	3.20	4.80	3.99	.39
Correct detection of problematic groups (grade)	29	5.50	10.00	7.91	1.15
Interpretation of dashboards grade	29	3.14	6.54	4.84	.91
Time questionnaires in minutes	29	7.81	61.82	18.05	11.66

The table shows the experience of teaching with fractions (M = 3.76, SD = 1.41) and the lessons with collaborative learning (M = 3.69, SD = 1.51). The data evaluation score (M = 9.01, SD = 1.51) was established by first adding the correct answers from the test followed by calculating the grade using the formula (see Method). The data application score (M = 3.07, SD = .72) was established by the added scores from the statements (on a scale from 1 to 5), divided by the number of statements. The same applies to experience with collaboration score (M = 3.99, SD = .39). The correctly detected problem groups grade (M = 7.91, SD = 1.15) was established by

adding the number of correct answers followed by calculating the score using the formula (see Method). The interpretation score (M = 4.84, SD = .91) was achieved by quantifying the interpretation for each interpreted dashboard, followed by calculating the score with the formula (See Appendix G), then adding these scores for all dashboards, and finally dividing it by the number of dashboards that the participant had interpreted. Furthermore, the average time participants needed to complete the questionnaires was 18.05 minutes (SD = 11.66).

At the end of the study, participants responded to four statements. An average of 4.17 (SD = .85) on a scale from 1 to 5 was found for the clarity of procedure, 3.48 (SD = .83) for having enough time to estimate the dashboards situations, 3.45 (SD = 1.09) for wanting a similar dashboard in their classroom and 3.83 (SD = .76) for having the right skills to use the dashboard.

Teachers' data literacy and the detection of problematic groups in CSCL dashboards

A multiple regression analysis was performed to investigate the relationship between data literacy and problem group detection. Before the results of the analysis could be interpreted, the assumptions had to be tested. First, the box plots and the stem-and-leaf indicated that data evaluation was not normally distributed. However, because the study was conducted among ± 30 participants, it can be assumed that all variables were normally distributed according to the Central Limit Theorem (Chakrapani, 2011; Field, 2013). Second, Mahalanobis distance gave no reason to remove multivariate outliers, as it did not exceed 13.82. Third, the data met the assumption of normality, linearity, and homoscedasticity of residuals, according to the standardized residuals against the standardized predicted values plot. Finally, the relatively high tolerance for both predictors demonstrated that multicollinearity did not interfere with the ability to interpret the results of the multiple regression analysis (Field, 2013).

The results of the regression indicated that data evaluation and data application did not significantly predict the detection of problematic groups in CSCL dashboards, F (2, 26) = .65, p = .53, R^2 = .05. Table 2 shows the unstandardised (*B)* and standardised (β) regression coefficients for each predictor in the regression model.

Table 2

Unstandardised (B) and Standardised (β) Regression Coefficients, and Squared Semi-Partial Correlations (sr^2) for Each Predictor in a Regression Model Predicting the Detection of Problematic Groups in CSCL Dashboards

Variable	*B* [95% CI]	β	sr^2	*p*
Data evaluation	.13 [-0.21, 0.46]	.16	0.02	.44
Data application	-.33 [-0.99, 0.33]	-.20	0.04	.32

Note. N = 29. CI = confidence interval.

Teachers' data literacy and their interpretation of CSCL dashboards

The relationship between teachers' data literacy and their interpretation of CSCL dashboards was also investigated with a multiple regression analysis. Again, the assumptions for the multiple regression analysis were tested, in the same way as the first hypothesis. All assumptions were supported. The results of the regression indicated that data evaluation and data application did not significantly predict the interpretation of CSCL dashboards, F (2, 26) = .58, p = .57, R^2 = .04. Table 3 shows the unstandardised (*B)* and standardised (β) regression coefficients for each predictor in the regression model.

Table 3

Unstandardised (B) and Standardised (β) Regression Coefficients, and Squared Semi-Partial Correlations (sr^2) for Each Predictor in a Regression Model Predicting the Interpretation of CSCL dashboards

Variable	*B* [95% CI]	β	sr^2	*p*
Data evaluation	-.07 [-0.34, 0.19]	-.11	0.01	.58
Data application	-.18 [-0.71, 0.34]	-.14	0.02	.48

Note. $N = 29$. CI = confidence interval.

The relationship between data literacy, experience with collaboration and the use of CSCL dashboards

The multiple regression analyses showed that the independent variables have no significant relationship with the use of the CSCL dashboard. However, it is still possible to add a moderator to the analysis to see if it changes the relationship between data literacy and the use of CSCL dashboards (Field, 2013). To examine if the relationship between data literacy and how teachers use CSCL dashboards was moderated by teacher experience with collaboration, four moderator analyses are performed.

First, a moderator analysis was performed with the detection of problematic groups as the outcome variable and data evaluation as predictor variable. The moderator variable used, in this and subsequent analysis, was experience with collaboration. No significant interaction effect was found, $b = -.33$, 95% CI [-1.19, .53], $t = -.79$, $p = .44$. Second, a moderator analysis was performed with detection of problematic groups as outcome variable and data application as predictor variable. Again, no significant interaction effect was found, $b = 1.16$, 95% CI [-0.45,

2.77], t = 1.48, p = .15. Third, another moderator analysis was performed with the interpretation of CSCL dashboards as the outcome variable and data evaluation as predictor variable. No significant interaction effect was found, b = .55, 95% CI [-0.11, 1.22], t = 1.71, p = .10. Finally, a moderator analysis was performed with interpretation of CSCL dashboards as outcome variable and data application as predictor variable. No significant interaction effect was found, b = .47, 95% CI [-0.88, 1.82], t = .71, p = .48. Since no significant interaction effects were found, Bonferroni method to correct a type I error was not necessary (Armstrong, 2014).

Discussion

During CSCL, it is essential that teachers have an accurate representation of their students, so they can better decide which help the students need to stimulate effective collaboration (Kaendler et al., 2015). The interactions between students can be monitored through CSCL dashboards, where teachers are provided with information about their collaborating students. Previous research suggested two teacher characteristics that could relate to the use of CSCL dashboards: data literacy and experience with collaboration. However, this remains largely unexamined. Therefore, this paper investigated the relationship between data literacy and the use of CSCL dashboards. Furthermore, it investigated whether experience with collaboration can moderate the relationship between data literacy and CSCL dashboards. This was done by conducting questionnaires and presenting dashboard situations among forty participants. The findings of the study will be summarized and discussed below. Additionally, the limitations of this study, and recommendations for future research will be discussed.

The first hypothesis of this study assumed that the relationship between data literacy and problem group detection was positive. However, the results showed no significant relationship between these two constructs. This was inconsistent with previous investigations, which

suggested that teachers with more data literacy would better understand data widgets (Ridsdale et al., 2015), and this would enable them to better identify problem groups (Charleer et al., 2014; Van Es & Sherin, 2002). The results have not been able to confirm this hypothesis. Research into the relationship between data literacy and the detection of problems, or the use of CSCL dashboards in general, is scarce. To the best of our knowledge, there are no studies that can confirm the results of this study. A possible reason for this result could be that the layout of the dashboard was already an aid in detecting problem groups. Van Leeuwen et al. (2019), who used the same design, indicated that the usability of the dashboard was high, due to low cognitive load, high confidence and clarity of the procedure. One of the features of a design with high usability is that it ensures that a goal is reached quickly and easily (Dumas & Redish, 1999), in this case detecting problem groups. Because of this high usability and low cognitive load to identify problem groups, it may be that data literacy in combination with the dashboard used in this study could not show a significant difference.

Another reason that could explain the result may be that the participants had too much time to recognize the problem group. Participants were given 50 seconds to view the dashboard trial. It could be argued that if they were given less time, only teachers with high levels of data literacy could recognize problem groups since data literacy leads to a more effective use of data (Ridsdale et al., 2015), and thus a faster recognition of problem groups. On the other hand, in a classroom where teachers do not have time to study the dashboard extensively, 50 seconds is a realistic time to look at the dashboard. Thus it may also be that data literacy, as the results confirm, is just not necessary to detect problem groups.

The second hypothesis assumed that there would be a positive relationship between data literacy and the interpretation of CSCL dashboards. The results showed no significant

relationship, which was contradictory to previous research of Charleer et al. (2014) and Ridsdale et al. (2015). There are several possible explanations for this result. First, the participants received the interpretation question after viewing the CSCL dashboard, which required them to recall the dashboard situation. Thus, not only data literacy could be important, but also remembering the situation when interpreting the CSCL dashboard. Perhaps their data literacy was high, but recalling this data was more difficult, resulting in a less rich interpretation. According to Nielsen (2005), it is not beneficial for users when they have to remember information to reproduce it later on. So, if the participant had the option to take a quick look at the dashboard while answering the interpretation question, this might have led to different results.

Another explanation for the result might be the coding scheme that was too strict; few of the participants got high results. The low average results of the interpretation questions might have been caused by the coding scheme, as only a few high scores have been distributed. For example, naming five specific elements might have been too difficult for the participants; only three percent of the comments scored the maximum of three points. In the ‘use of data widgets’ part of the coding scheme, participants did not manage score high either.

Regarding the third and last hypothesis, it was expected that experience with collaboration would moderate the relationship between data literacy and how teachers use CSCL dashboards. The relationship between data literacy and the use of CSCL dashboards was assumed to be strengthened when teachers had more specific ideas about collaboration, and to be diminished when their ideas about collaboration were less specific. Nevertheless, the four moderation analyses showed no significant relationships. Therefore, it can be concluded that experience with collaboration does not moderate the relationship between data literacy and how teachers use

CSCL dashboards in this study. These results were not in line with the findings of Peterson and Comeaux (1987) and Van den Bogert et al. (2014).

The reason for this result may have something to do with the translation of teacher experience with collaboration to the dashboard situation. Mantyjarvi, Himberg, and Huuskonen (2003) describe that when using technological applications, there is less context information for the user. In this study, the teachers were unable to see their students in real-life, nor could they contact them when using the dashboard. Some participants indicated that it was hard to only use the dashboard without seeing the children. So, because there was less context information for teachers when using a CSCL dashboard and it felt more unnatural, it may be that their ideas about collaboration were harder to translate into the dashboard. To rule out this presumption, research should investigate the use of CSCL dashboards in the real class scenarios, where the teachers have the option to actually observe the collaborating students and at the same time receive dashboard information about their interactions. In this way, teachers have more context information and it may be that the role of experience in collaboration turns out differently.

The results of this study suggest that experience with collaboration might not moderate the relationship with data literacy and the use of CSCL dashboards as much as initially hypothesized. This could imply that teachers without experience with collaboration can use CSCL dashboards just as well as teachers with experience in collaboration when they have high levels of data literacy, which makes the use of CSCL dashboards more accessible for every teacher.

Limitations

The scope of this study was limited in terms of the reliability of the data questionnaire. The results of the questionnaire could have been influenced by the decision to pilot the

questionnaire on everyone older than 18. Because the target group in this study slightly differs from this, the results of the reliability are somewhat surprising. The data literacy scores, especially for the knowledge area data evaluation, were relatively high. As a result, a ceiling effect has arisen for this area. When a ceiling effect occurs, the scores of the participants reach the high end of the instrument. This ceiling effect, in turn, influences the reliability of the instrument; if the best possible scores can no longer be distinguished from each other, a lack of variance emerges (Michalos, 2014). A ceiling effect occurs more often with similar instruments, such as the graph literacy scale by van den Bosch et al. (2017).

The relatively high scores from the target sample compared to the pilot sample may be explained by the fact that the questionnaire was too easy for the target group. According to Cowie and Cooper (2017) and Van den Bosch et al. (2017), the development of data literacy probably requires mathematical and statistical insight. The average teacher may have more knowledge about mathematics and statistics because they were selected for this. Since 2006, students at PABO and ALPO have been given a test (The Wiscat-PABO) that is part of the binding recommendation on further study at the end of the first year of the program (Straetmans & Eggen, 2005). Taking into account the average age of the participants in this study (M = 29.47), a large proportion had experienced this binding recommendation during their study. This standard could be a reason that the data literacy of the target sample was greater than the pilot sample and thus ensured that through the ceiling effect, the reliability turned out different.

Additionally, a few items were removed from the questionnaire following the reliability analysis. In the knowledge area data application, this resulted in the loss of the critical thinking competency which concerned awareness of high-level issues related to data, and critical thinking

when dealing with data. Consequently, the content validity of the questionnaire can be called into question.

There are several reasons why care should be taken in generalizing the results. First, this study used a specific CSCL dashboard designed by Van Leeuwen et al. (2019). Results may have been different when another design was used. Second, the content of the dashboards used deals with a specific math topic: fractions. Finally, the CSCL dashboard only presented dyads; other group sizes were not tested.

Recommendations for future research

Future research should be undertaken to investigate other teacher characteristics that underlie the use of CSCL dashboards. One of those teacher characteristics could be the attitude towards computers. There are indications that a positive attitude towards computers leads to better technological skills, making it easier to use the CSCL dashboard (Bame, Dugger, de Vries, & McBee, 1993). An instrument that can investigate this presumption is the Computer Attitude Scale, where a high score on the scale implies a positive attitude towards computer use (Teo, 2008). More scientific knowledge in such skills or attitudes could lead to more awareness among school leaders and trainers to stimulate them, which would result in a more effective use of CSCL dashboards.

Subsequently, it would be interesting if future research can investigate the field of data literacy among teachers, as Mandinach (2012) calls this *pedagogical data literacy*. Pedagogical data literacy entails the combination of pedagogical content knowledge and how the data is used to impact classroom practices in order to affect change in student learning. It may be useful for future research to include pedagogical context knowledge in the definition of data literacy. Afterwards, one can develop a reliable and validated questionnaire measuring the data literacy of

teachers more difficult questions to make a stronger distinction between teachers with high and low levels of data literacy.

The last suggestion for future research would be to investigate more dyads, preferably equivalent to a class, in CSCL dashboards. In this study, five dyads were examined on the CSCL dashboard; however, ten to fifteen dyads are more representative for a classroom setting. Therefore, future studies could investigate if the relationship between data literacy and the use of CSCL dashboards changes with more dyads. Data literacy may take on a more prominent role as teachers have to monitor more data widgets, and, thus, more data widgets.

As this is one of the first studies that investigates the relationship between teacher data literacy and the use of CSCL dashboards, this study has contributed to defining specific data literacy skills required for the use of a CSCL dashboard. An instrument has been developed that measures data literacy. Besides, this study showed that there was no significant relationship found between data literacy and the use of CSCL dashboards. Furthermore, this study has shown that experience with collaboration did not moderate the relationship between data literacy and the use of CSCL dashboards. Additionally, it contributed to the development of a questionnaire that measures experience with collaboration based on the ideas that teachers have about collaborative learning. This questionnaire complements the work that is found so far.

The results of this study could mean that data literacy has a less strong relationship with the use of CSCL dashboards than expected. However, the skill remains important for the 21st century as it contributes to a valuable knowledge-based economy (Chinien & Boutin, 2011). For this reason, and because education is becoming increasingly digital, educators and schools have to be more committed to the skill. Besides, teachers might not need any experience with collaboration to use a CSCL dashboard, which makes the use more accessible for every teacher.

As mentioned earlier, more research should be conducted into these and other characteristics that relate to the use of CSCL dashboards. Thereafter, guidelines can be made up about which aspects should be taught and stimulated in teacher training.

References

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, *34*(5), 502–508. https://doi.org/10.1111/opo.12131

Arvaja, M., Häkkinen, P., Rasku-Puttonen, H., & Eteläpelto, A. (2002). Social Processes and Knowledge Building During Small Group Interaction in a School Science Project. *Scandinavian Journal of Educational Research*, *46*(2), 161–179. https://doi.org/10.1080/00313830220142182

Bame, E. A., Dugger, W. E., de Vries, M., & McBee, J. (1993). Pupils' attitudes toward technology—PATT-USA. *The Journal of Technology Studies*, *19*(1), 40-48. Retrieved from www.jstor.org/stable/43603608

Chakrapani, C. (2011, April). Statistical reasoning vs. magical thinking, shamanism as statistical knowledge: is a sample size of 30 all you need. *Vue Magazine*. Retrieved from http://www.chuckchakrapani.com/articles/pdf/0411chakrapani.pdf

Charleer, S., Klerkx, J., & Duval, E. (2014). Learning Dashboards. *Journal of Learning Analytics*, *1*(3), 199–202. https://doi.org/10.18608/jla.2014.13.22

Chinien, C., & Boutin, F. (2011). *Defining essential digital skills in the canadian workplace*. Retrieved from http://en.copian.ca/library/research/digi_es_can_workplace/digi_es_can_workplace.pdf

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American journal of mental deficiency*. Retrieved from https://psycnet.apa.org/record/1982-00095-001

Cowie, B., & Cooper, B. (2017). Exploring the challenge of developing student teacher data

literacy. *Assessment in Education: Principles, Policy & Practice, 24*(2), 147–163. https://doi.org/10.1080/0969594x.2016.1225668

De Lisi, H., & Golbeck, A. L. (2014). Implications of Piagetian Theory for Peer Learning. In A. M. O'Donnell & A. King (Red.), *Cognitive Developmental Bases of Pear Learning* (pp. 3–38). Retrieved December 10, 2019 https://books.google.nl/books?hl=nl&lr=&id=7_ZRAwAAQBAJ&oi=fnd&pg=PA3&dq=De+Lisi,+R.,+%26+Golbeck,+S.+L.+(1999).+Implications+of+Piagetian+theory+for+peer+learning.&ots=fnTV7_OzF2&sig=qYl9yKh0z0yHdHKMkPToAgSKpb8#v=onepage&q&f=false

Dumas, J. S., & Redish, J. (1999). *A practical guide to usability testing*. Intellect books. Retrieved from http://www.jedbrubaker.com/wp-content/uploads/2013/03/Dumas-99.pdf

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

Fawcett, L. M., & Garton, A. F. (2005). The effect of peer collaboration on children's problem-solving ability. *British Journal of Educational Psychology*, *75*(2), 157–169. https://doi.org/10.1348/000709904x23411

Ferketich, S. (1991). Focus on psychometrics. Aspects of item analysis. *Research in Nursing & Health*, *14*(2), 165–168. https://doi.org/10.1002/nur.4770140211

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th edition). Thousand Oaks, Canada: SAGE Publications.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A*

regression-based approach. Guilford publications.

Järvelä, S., Kirschner, P.A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, … Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: Designing for CSCL regulation tools. *Educational Technology Research and Development, 63*, 125-142. doi:10.1007/s11423-014-9358-1

Johnson, C. A. (2012). *The Information Diet: A Case for Conscious Consumption* (1st edition). California, United States: O'Reilly Media.

Kaendler, C., Wiedmann, M., Rummel, N., & Spada, H. (2015). Teacher Competencies for the Implementation of Collaborative Learning in the Classroom: a Framework and Research Review. *Educational Psychology Review*, *27*(3), 505–536. https://doi.org/10.1007/s10648-014-9288-9

Kruger, A. C. (1993). Peer collaboration: conflict, cooperation, or both? *Social Development*, *2*(3), 165–182. https://doi.org/10.1111/j.1467-9507.1993.tb00012.x

Lee, S. W. Y., & Tsai, C.C. (2011). Students' perceptions of collaboration, self-regulated learning, and information seeking in the context of Internet-based learning and traditional learning. *Computers in Human Behavior*, *27*(2), 905–914. https://doi.org/10.1016/j.chb.2010.11.016

Ley, T. (2019). Knowledge structures for integrating working and learning: A reflection on a decade of learning technology research for workplace learning. *British Journal of Educational Technology*, *51*(2), 331–346. https://doi.org/10.1111/bjet.12835

Mandinach, E. B. (2012). A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice. *Educational Psychologist, 47*(2), 71–85. doi:10.1080/00461520.2012.667064

Mandinach, E. B., & Gummer, E. S. (2016). *Data Literacy for Educators*. Amsterdam, Nederland: Amsterdam University Press.

Mantyjarvi, J., Himberg, J., & Huuskonen, P. (2003, March). Collaborative context recognition for handheld devices. In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003 (PerCom 2003).* (pp. 161-168). IEEE.

Martinez Maldonado, R., Kay, J., Yacef, K., & Schwendimann, B. (2012). An Interactive Teacher's Dashboard for Monitoring Groups in a Multi-tabletop Learning Environment. *Intelligent Tutoring Systems*, 482–492. https://doi.org/10.1007/978-3-642-30950-2_62

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction*, *6*(4), 359–377. https://doi.org/10.1016/s0959-4752(96)00021-7

Michalos, A. C. (2014). *Encyclopedia of quality of life and well-being research*. Dordrecht: Springer Netherlands.

Moshman, M., Geil, D. (1998). Collaborative Reasoning: Evidence for Collective Rationality. *Thinking & Reasoning*, *4*(3), 231–248. https://doi.org/10.1080/135467898394148

Nielsen, J. (2005). *Ten Usability Heuristics*. Retrieved from https://pdfs.semanticscholar.org/5f03/b251093aee730ab9772db2e1a8a7eb8522cb.pdf

Oh, H. J. (2004, April). Reasoning test takers' guessing strategy and their understanding of formula scoring. In *annual meeting of the American Educational Research Association (AERA)*.

Paes, E. C., & ten Cate, O. (2009). Meerkeuzevragen met drie, vier of vijf alternatieven: wat is beter? *Tijdschrift voor Medisch Onderwijs*, *28*(3), 124–129. https://doi.org/10.1007/bf03081770

Pallant, J. (2020). *SPSS Survival Manual*. New York, United States: McGraw-Hill Education.

Peterson, P. L., & Comeaux, M. A. (1987). Teachers' schemata for classroom events: The mental scaffolding of teachers' thinking during classroom instruction. *Teaching and Teacher Education*, *3*(4), 319–331. https://doi.org/10.1016/0742-051x(87)90024-2

Pourshahid, A. (2015). *U.S. Patent Application No. 14/176,180*.

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., … Wuetherick, B. (2015). *Strategies and best practices for data literacy education: Knowledge synthesis Report.* https://doi.org/10.13140/RG.2.1.1922.5044

Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, *63*(2), 123-134.

Schmitz, M. J., & Winskel, H. (2008). Towards effective partnerships in a collaborative problem-solving task. *British Journal of Educational Psychology*, *78*(4), 581–596. https://doi.org/10.1348/000709908x281619

Stahl, G., Koschmann, T. & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (ed.), *Cambridge handbook of the learning sciences* (pp. 409-426), Cambridge University Press.

Straetmans, G. J. J. M., & Eggen, T. J. H. M. (2005). Afrekenen op rekenen: over de rekenvaardigheid van pabo-studenten en de toetsing daarvan. *Tijdschrift voor hoger onderwijs*, *23*(3), 123-139. Retrieved from https://www.researchgate.net/profile/Theo_Eggen/publication/239552433_Afrekenen_op_rekenen_over_de_rekenvaardigheid_van_pabo-studenten_en_de_toetsing_daarvan/links/0a85e534398e6a21b2000000.pdf

Suthers D.D. (2012) Computer-Supported Collaborative Learning. In: Seel N.M. (eds) Encyclopedia of the Sciences of Learning. Springer, Boston, MA

Underwood, J., Underwood, G., & Wood, D. (2000). When does gender matter? *Learning and Instruction*, *10*(5), 447–462. https://doi.org/10.1016/s0959-4752(00)00008-6

Universiteit Utrecht. (z.d.). *Handout 5a. Toetsanalyse*. Retrieved February 12, 2020, from https://tauu.uu.nl/wp-content/uploads/2015/03/Handout-5a-Toetsanalyse.pdf

van den Bogert, N., van Bruggen, J., Kostons, D., & Jochems, W. (2014). First steps into understanding teachers' visual perception of classroom events. *Teaching and Teacher Education*, *37*, 208–216. https://doi.org/10.1016/j.tate.2013.09.001

van den Bosch, R. M., Espin, C. A., Chung, S., & Saab, N. (2017). Data-Based Decision-Making: Teachers' Comprehension of Curriculum-Based Measurement Progress-Monitoring Graphs. *Learning Disabilities Research & Practice*, *32*(1), 46–60. https://doi.org/10.1111/ldrp.12122

van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of technology and teacher education*, *10*(4), 571-596. Retrieved from https://pdfs.semanticscholar.org/c253/d8a8436583754a70c862ddea603366d71665.pdf

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, *24*(2), 244–276. https://doi.org/10.1016/j.tate.2006.11.005

van Leeuwen, A., & Rummel, N. (2020, March). Comparing teachers' use of mirroring and advising dashboards. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 26-34).

van Leeuwen, A., Rummel, N., & van Gog, T. (2019). What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations? *International Journal of*

Computer-Supported Collaborative Learning, *14*(3), 261–289. https://doi.org/10.1007/s11412-019-09299-x

Verbert, K., Govaerts, S., Duval, E., Santos, J. L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing*, *18*(6), 1499-1514. https://doi.org/10.1007/s00779-013-0751-2

Wasson, B., & Hansen, C. (2015). Data literacy and use for teaching. In *Measuring and Visualizing Learning in the Information-Rich Classroom* (pp. 72-89). Routledge.

Wasson, B., Hansen, C., & Netteland, G. (2016). Data literacy and use for learning when using learning analytics for learners. CEUR Workshop Proceedings, 1596, 38-41.

Zurita, G., & Nussbaum, M. (2004). Computer supported collaborative learning using wirelessly interconnected handheld computers. *Computers & Education*, *42*(3), 289–314. https://doi.org/10.1016/j.compedu.2003.08.005

Appendix A

Data literacy conceptualisation

Table 4

Data Literacy Conceptualisation

Knowledge area	Competency	Knowledge/tasks
Data evaluation	Data interpretation (understanding data)	· Reads and understands charts, tables, and graphs · Identifies discrepancies within the data
	Identifying problems using data	· Uses data to identify problems in practical situations (e.g., workplace efficiency) · Uses data to identify higher-level problems (e.g., policy, environment, scientific experimentation, marketing and economics)
	Data visualization	· Evaluates effectiveness of graphical representations · Critically assesses graphical representations for accuracy and misrepresentation of data
	Data-Driven Decisions Making (DDDM) (Making decisions based on data)	· Converts data into actionable information · Weighs the merit and impacts of possible solutions/decisions · Implements decisions/solutions

Data application	Critical thinking	· Aware of high-level issues and challenges associated with data · Critically thinks when working with data
	Data culture	· Recognizes the importance of data · Supports an environment that fosters the critical use of data for learning, research, and decision-making

Note. Adapted from "Strategies and best practices for data literacy education: Knowledge synthesis Report" by Ridsdale et al. (2015), p. 38

Appendix B

Questionnaire with background information

Geslacht: man / vrouw

Leeftijd: ……...

In welke groep geeft u les?: …….

Indien u momenteel geen les geeft; hoeveel maanden geleden heeft u voor het laatst les gegeven? ……...

Heeft u ervaring met lesgeven over breuken?

- 0-5 lessen
- 6-10 lessen
- 11-15 lessen
- meer dan 15 lessen

Heeft u ervaring met samenwerkend leren en dashboards? (samenwerkend leren gaat over het bereiken van een gemeenschappelijk begrip met anderen, waarbij de groepsleden bij dezelfde doelen en taken betrokken zijn).
……...……...……...
……...……...……

Maak een schatting van het aantal lessen dat je samenwerkend leren hebt ingezet dit schooljaar:

- 0-5 lessen
- 6-10 lessen
- 11-15 lessen
- meer dan 15 lessen

Ervaring met dashboards. Met zogenaamde "Dashboards" bedoelen we computerschermen die de activiteiten van leerlingen automatisch volgen en de leerkracht hierover informatie geven tijdens de les, bijvoorbeeld het leerkrachtscherm van Snappet.

Heb je ervaring met dashboards? Ja/Nee

Indien je ervaring hebt met Dashboards, kun je toelichten om welke software het gaat en hoeveel ervaring je ermee hebt? ……..

Appendix C

Questionnaire data literacy

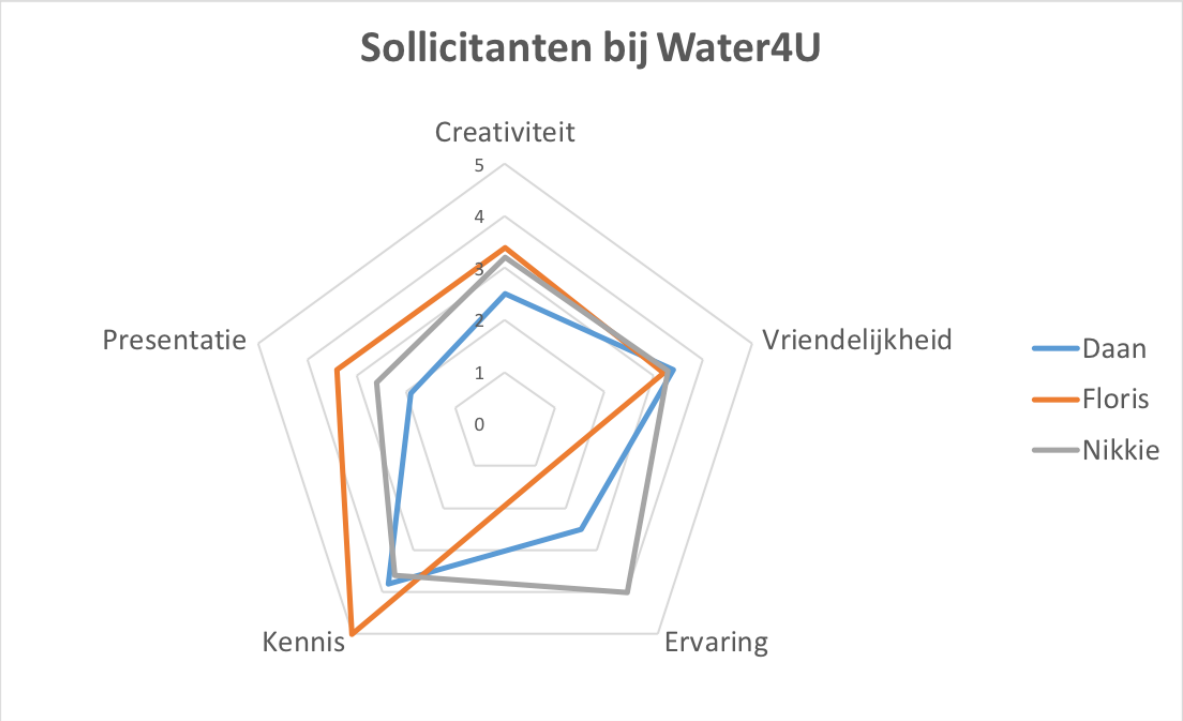
Deze vragenlijst gaat over data. Data is een benaming voor digitale gegevens, hierbij kan u denken aan de gegevens op een computer, maar ook aan de gegevens die u ophaalt bij het afnemen van een vragenlijst.

Eerst volgen er acht vragen die u zo goed mogelijk moet beantwoorden. Tot slot volgen er zes stellingen over uw houding tegenover data.

De gegevens in onderstaande figuren zijn fictief, deze zijn verzonnen.

Vraag 1

Sollicitanten bij Water4U



De baas van Water4U heeft 3 mensen laten solliciteren en zijn indruk in een radargrafiek gezet. Hij vindt alle vaardigheden even belangrijk. Wie zou hij op basis hiervan kiezen?

A. Daan
B. Floris
C. Nikkie

Vraag 2

Informatie over het klimaat in Nederland

Parameter	Maandag	Dinsdag	Woensdag	Donderdag	Vrijdag
Luchtdruk (hPa)	1013	1015	1017	1019	1018
Gemiddelde windsnelheid (m/s)	3,4	3,4	3,3	3,3	3,2
Zonneschijnduur (h)	2,3	4,5	3,9	4,3	3,4
Gemiddelde temperatuur (°C)	18,2	19,2	18,6	14,5	16,7
Gemiddelde relatieve luchtvochtigheid (%)	72	73	72	75	74
Neerslagtotaal (in mm)	2,3	1,2	1,1	2,4	4,2

Met spreekt van aangename omstandigheden als:

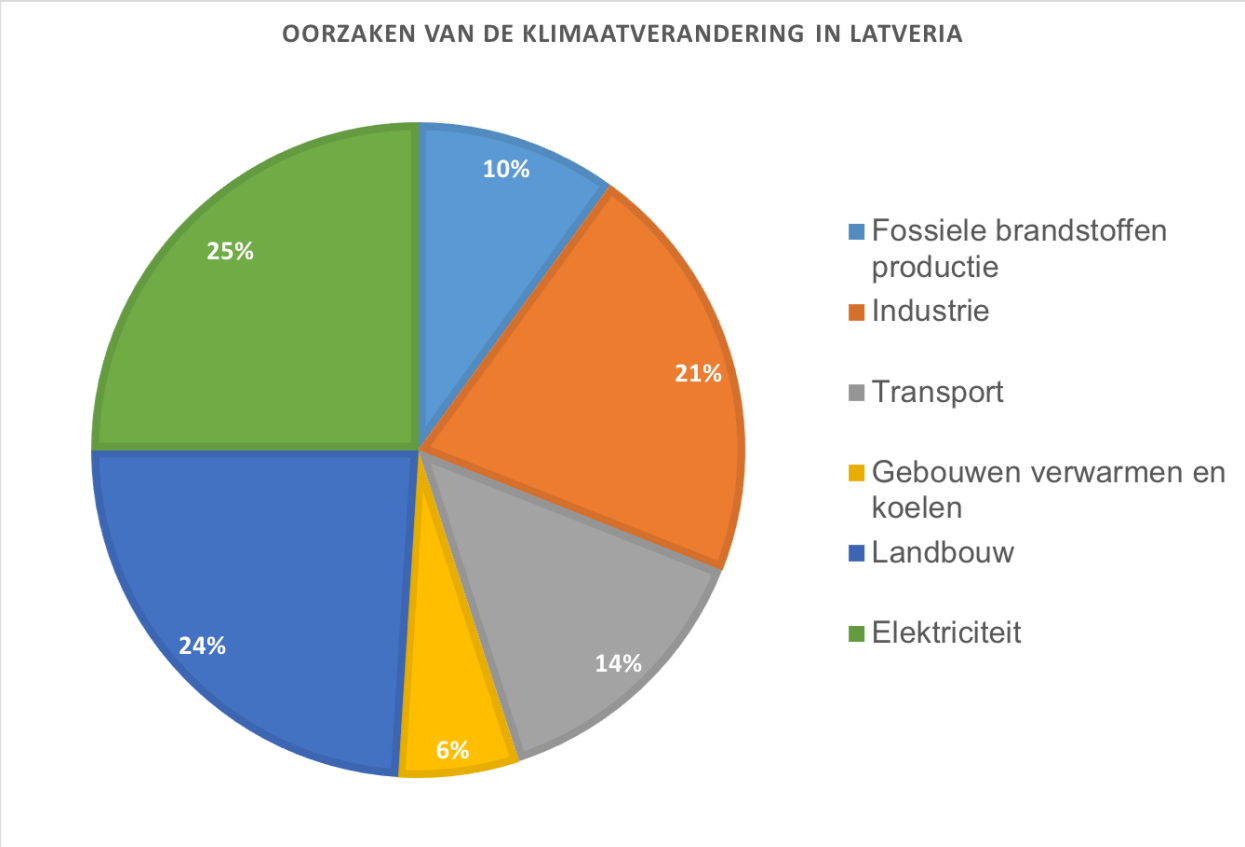
- De luchtdruk hoger is dan 1014 hPa
- Er meer dan 4 uur zonneschijn is per dag
- En de luchtvochtigheid lager of gelijk is aan 73%

Welke dag had de meest aangename omstandigheden?

A. Donderdag
B. Dinsdag
C. Woensdag

Vraag 3

Oorzaken van de klimaatverandering in Latveria



Welke van de volgende stellingen over het figuur is juist?

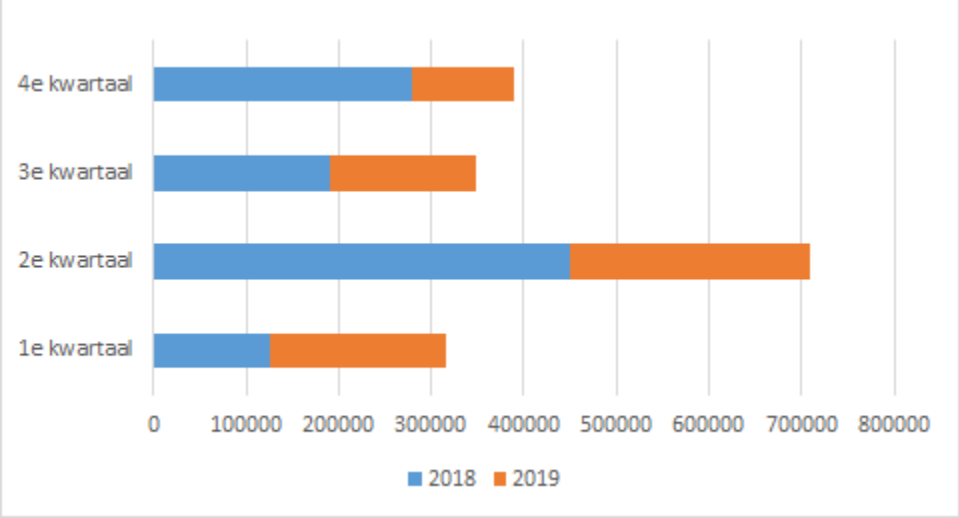
A. In Latveria wordt minder dan 50% van de klimaatverandering veroorzaakt door transport, industrie en elektriciteit.

B. *In Latveria wordt meer dan 40% van de klimaatverandering veroorzaakt door landbouw en industrie.*

C. Het verwarmen en koelen van gebouwen en industrie veroorzaken samen 26% van de klimaatverandering in Latveria.

Vraag 4

Omzet Water4U 2018 en 2019

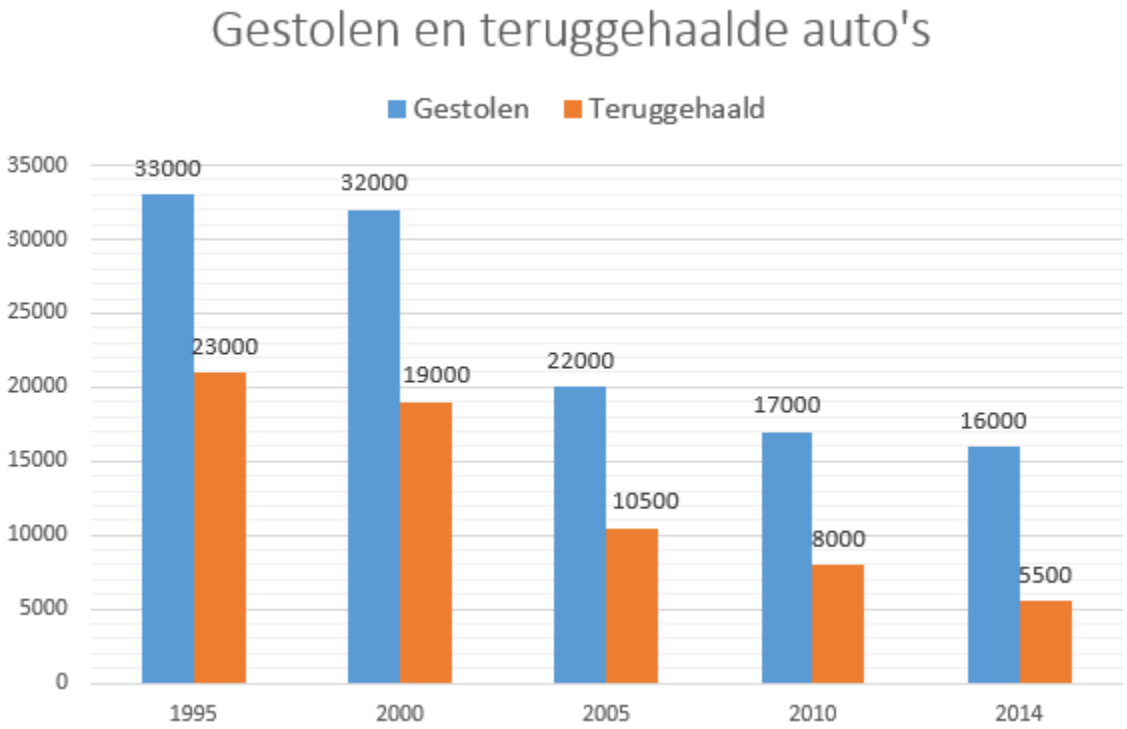


Wat kan geconcludeerd worden uit dit figuur?

A. Water4U doet het ongeveer even goed in 2018 als in 2019
B. In 2019 is er in het tweede kwartaal meer omzet gemaakt dan in 2018
C. Water4U heeft in 2018 meer omzet gemaakt dan in 2019

Bij deze vragen is het de bedoeling dat je kritisch kijkt naar de figuren.

Vraag 5

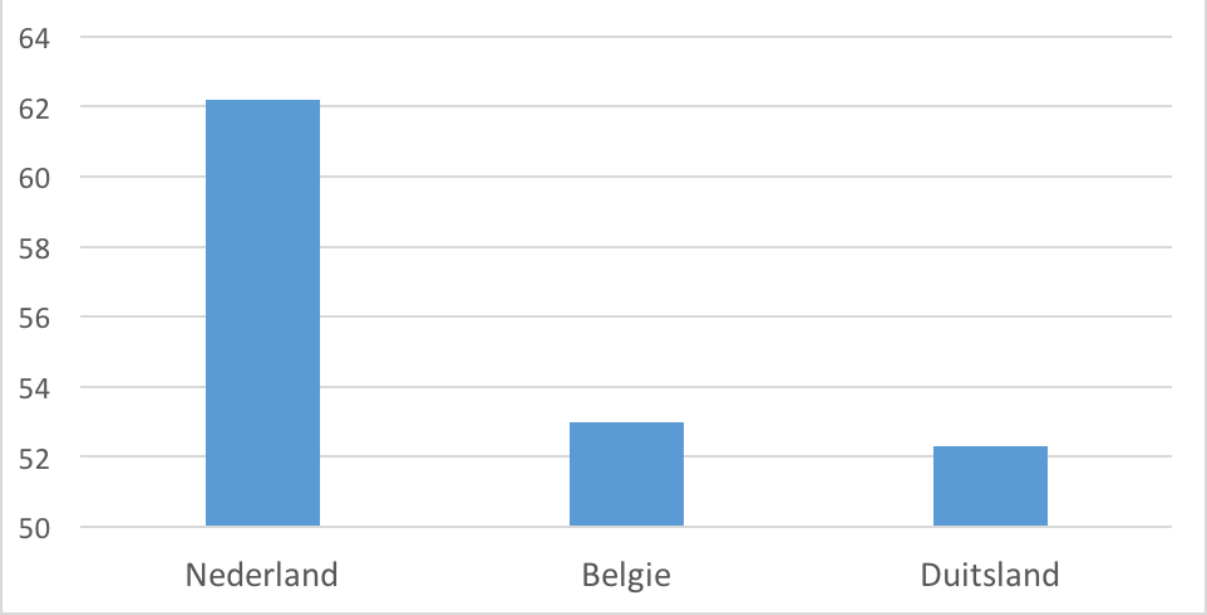


Welk van de drie onderstaande onderdelen van de staafgrafiek bevat geen fout?

A. De x-as
B. De y-as
C. De waardes van de staven

Vraag 6

Percentage gezond gewicht per land (BMI 18,5 - 24,9)



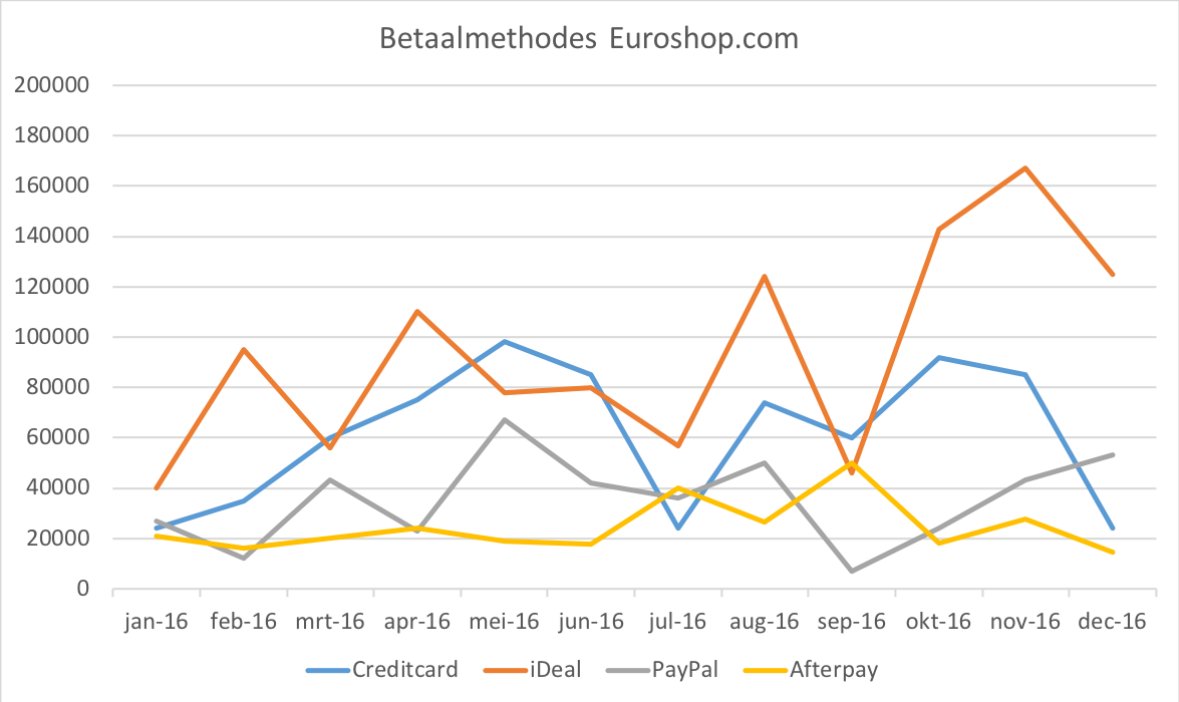
Bovenstaand figuur geeft het percentage gezond gewicht per land weer. Welke van de onderstaande stellingen is juist?

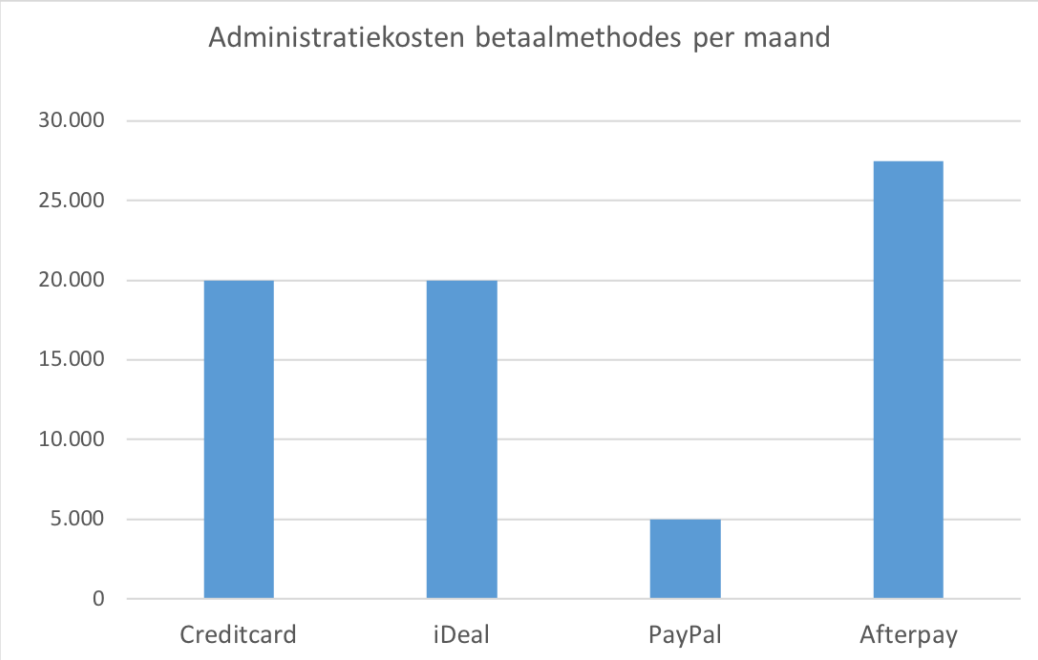
A. Dezelfde data kon beter in een cirkeldiagram weergeven worden.
B. *De y-as maakt de figuur misleidend.*
C. Nederland is minstens 3x zo gezond als België en Duitsland.

Vraag 7

In de volgende grafieken staan de betaalmethodes die gebruikt worden bij Euroshop.com en de administratiekosten van de betaalservices per maand. Bij Euroshop.com kunnen klanten maximaal 1 product bestellen van 1 euro. Welke van de onderstaande stellingen is juist?

A. In september 2016 heeft de betaalmethode iDeal minder opgeleverd dan dat het gekost heeft.
B. Euroshop.com heeft relatief weinig administratie kosten aan PayPal, maar maakt over het gehele jaar toch verlies met deze betaalmethode.
C. Om kosten te besparen, kan Euroshop.com het beste Afterpay afschaffen.


Betaalmethodes Euroshop.com
200000
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
jan-16
feb-16
mrt-16
apr-16
mei-16
jun-16
jul-16
aug-16
sep-16
okt-16
nov-16
dec-16
Creditcard
iDeal
PayPal
Afterpay


Administratiekosten betaalmethodes per maand
30.000
25.000
20.000
15.000
10.000
5.000
0
Creditcard
iDeal
PayPal
Afterpay

Vraag 8

Samenwerkingscores bij Water4U op schaal van 1 tot 5

	Henk	Babs	Kim	Jordy	Jan	Freek	Gerda	Marlies	Chris	Pim
Henk	-	2	3	3	4	5	4	3	4	5
Babs	2	-	3	1	5	2	3	1	2	3
Kim	3	3	-	5	4	3	2	3	4	3
Jordy	3	1	5	-	5	4	2	3	3	4
Jan	4	5	4	5	-	5	4	4	5	5
Freek	5	2	3	4	5	-	4	3	2	3
Gerda	4	3	2	2	4	4	-	3	4	2
Marlies	3	1	3	3	4	3	3	-	2	3
Chris	4	2	4	3	5	2	4	2	-	2
Pim	5	3	3	4	5	3	2	3	2	-

De baas van Water4U heeft de samenwerking tussen elke medewerker gescoord van laag naar hoog (1 t/m 5). Hij wil graag een medewerker aanstellen als teamleider. Het is van belang dat deze persoon met iedereen goed kan samenwerken.
Welke van de volgende stellingen is juist?

A. *De kans dat Jan wordt aangesteld als teamleider is het grootst*
B. Marlies heeft de kleinste kans om aangesteld te worden als teamleider
C. De kans dat Henk wordt aangesteld als teamleider is het grootst

Nu volgen er een aantal stellingen.

9. Als ik kijk naar figuren of grafieken, ben ik altijd kritisch of dit de juiste manier is waarop de gegevens weergeven kunnen worden.

Helemaal oneens - oneens - neutraal - eens - helemaal eens

10. Ik neem niet zomaar data over van anderen, zonder er nauwkeurig naar te kijken.

Helemaal oneens - oneens - neutraal - eens - helemaal eens

11. Ik ben mij bewust van de risico's die komen kijken bij grote hoeveelheden data, zoals een datalek of privacygevoelige datasets.

Helemaal oneens - oneens - neutraal - eens - helemaal eens

12. Ik werk graag bij een organisatie waar er veel met data wordt gewerkt.

Helemaal oneens - oneens - neutraal - eens - helemaal eens

13. Ik werk graag bij een organisatie waar data wordt gebruikt voor het maken van beslissingen.

Helemaal oneens - oneens - neutraal - eens - helemaal eens

14. Ik ben elke dag bezig met data, voor mijn werk of in mijn vrije tijd.

Helemaal oneens - oneens - neutraal - eens - helemaal eens

Appendix D

Questionnaire Teacher experience with collaboration

Tijdens samenwerkend leren vind ik het belangrijk dat…

1. **leerlingen problemen bespreken met klasgenoten die zich voordoen tijdens het leren**

Nooit - zelden - soms - vaak - altijd

2. **leerlingen hun notities en leermiddelen delen met elkaar**

Nooit - zelden - soms - vaak - altijd

3. **leerlingen feedback geven op de ideeën die hun klasgenoten hebben**

Nooit - zelden - soms - vaak - altijd

4. **leerlingen hun leerervaringen met elkaar uitwisselen**

Nooit - zelden - soms - vaak - altijd

5. **leerlingen met elkaar discussiëren**

Nooit - zelden - soms - vaak - altijd

6. **leerlingen gebruik maken van de informatie die ze elkaar geven**

Nooit - zelden - soms - vaak - altijd

7. **leerlingen samen leren/lezen voordat zij een toets maken**

Nooit - zelden - soms - vaak - altijd

Appendix E

Example of a dashboard trial

Situatieschets

In deze situatie geef je als invaller les aan een **groep 6**. De leerlingen hebben al een les in duo's aan breukenopgaven gewerkt, waarbij de leerlingen **op niveau gekoppeld** waren aan elkaar. Je hebt zojuist instructie gegeven aan de klas en de duo's in dezelfde samenstelling verder laten werken. Omdat geen enkel duo hun vinger opsteekt, pak je het dashboard erbij om te kijken hoe het met de groepjes gaat.

Klik op "Start" om het dashboard te zien en uit te zoeken of er een groep is waar een probleem speelt. **Doe dit zo snel maar nauwkeurig mogelijk, zodat je daarna de klas weer kunt observeren**. Als je klaar bent, klik je op "klaar met bekijken", en wordt je gevraagd enkele vragen te beantwoorden over wat je gezien hebt.

In welke groep was er in deze situatie volgens jou iets aan de hand? (multiple-choice)

- Groep 1
- Groep 2
- Groep 3
- Groep 4
- Groep 5
- Met geen enkele groep was iets aan de hand.

Geef een toelichting op wat volgens jou het probleem was. Wees alsjeblieft zo specifiek mogelijk over wat je in de situatie opgemerkt hebt.

……..……
…..……..……
……..……..……

Appendix F

Overview of dashboards situations

Table 5

Overview of Dashboard Situations.

Situation	Problem with one dyad
1	The dyad is stuck on an assignment.
2	The dyad shows off-task behaviour.
3	One of the students is dominating the collaboration.
4	The dyad is alternating the tasks.
5	The dyad is having ineffective discussions, where they do not consider each other's ideas.
6	The dyad displays trial-and-error behaviour.
7	None.
8	None.

Note. Adapted from "What information should CSCL teacher dashboards provide to help teachers interpret CSCL situations?" by van Leeuwen, A., et al., 2019, *International Journal of Computer-Supported Collaborative Learning*, 14(3), pp. 272-273

Appendix G

Coding scheme teacher interpretation of CSCL dashboards

Table 6

Coding scheme teacher interpretation of CSCL dashboards

Code	Levels
Specificity	Specific elements: specific fraction skill, a specific student name, a specific number from one of the indicator graphs, a comparison of multiple groups of students, or a comparison of one group over multiple time points (Van Leeuwen et al., 2019) 0. There is no specific element mentioned 1. 1 specific element is mentioned 2. 2 specific elements are mentioned 3. 3 specific elements are mentioned 4. 4 specific elements are mentioned 5. More than 5 specific elements are mentioned
Attitude	0. The problem is not described 1. The problem is described as isolated 2. The problem is described and associated with a broader learning principle.
Use of data widgets	Data widgets are: completed assignments, attempts per assignment, chance of trial-and-error behaviour, amount of talk, skill proficiency and activity timeline. It should be clear that they have used the data widget, it is not necessary to literally name the data widget. The wording should show that they base their answer on a data widget 0. The problem is not explained with reference to a data widget

1. It explains what the problem is with the help of 1 data widget

2. It explains what the problem is with the help of 2 or more data widgets.

3. It explains what the problem is with the help of 2 or more data widgets and those are linked with each other.

Weights
Score specificity = 1 * (determined level)
Score attitude = 2.5 * (determined level)
Score data widgets = 1.66 * (determined level)

Score #participant interpretation = (9 * (specificity + attitude + data widgets) / 15) + 1

Appendix H

Questionnaire about the study

De volgende vragen gaan over hoe je het onderzoek hebt ervaren. Omcirkel het antwoord dat bij jou van toepassing is.

1. **De procedure van het onderzoek was duidelijk**

   Helemaal oneens - oneens - neutraal - eens - helemaal eens

2. **Ik had tijd genoeg om de situaties adequaat in te schatten.**

   Helemaal oneens - oneens - neutraal - eens - helemaal eens

3. **Ik zou graag een soortgelijk dashboard in mijn klas willen hebben**

   Helemaal oneens - oneens - neutraal - eens - helemaal eens

4. **Ik had het idee dat ik over de juiste vaardigheden beschikte om het dashboard te gebruiken.**

   Helemaal oneens - oneens - neutraal - eens - helemaal eens

5. **Kun je aangeven welke strategie je in het algemeen gebruikte om de situaties in te schatten? (Waar keek je naar, wat vond je belangrijk om op te letten, etc)**

   ……...……

   ……...……

   ……...……...

   ……...……...

   ……...……...

6. **Heb je eventueel nog toelichting op bovenstaande antwoorden of andere opmerkingen over het onderzoek? Die kun je hieronder kwijt:**

……...……

……...……

……...……...

……...……...

……...……...

Appendix I

Improvement Plan Adjusted Questionnaire

Table 7

Test analysis data evaluation part

	Alternatives				Test analysis				
Question	A	B	C	D	N_participants	p-value	Corrected p-value	Rit-value	Rir-value
1	0	0	4	**47**	51	0,92	0,88	0,52	0,35
2	3	0	0	**48**	51	0,94	0,91	0,23	0,06
3	**40**	3	3	5	51	0,78	0,68	0,42	0,13
4	1	14	**36**	0	51	0,71	0,56	0,59	0,30

	Alternatives						Test analysis				
Question	A	B	C	D	E	F	N_participants	p-value	Corrected p-value	Rit-value	Rir-value
5*	13	1	**31**	**30**	**42**	1	51	0,22	-0,17	0,45	0,15
6**	13	3	3	**39**	3	5	51	0,57	0,36	0,35	-0,13

	Alternatives										Test analysis				
Question	Henk	Babs	Kim	Jordy	Jan	Freek	Gerda	Marlies	Chris	Pim	N_participants	p-value	Corrected p-value	Rit-value	Rir-value
7	0	**45**	0	1	0	1	2	2	0	0	51	0,88	0,82	0,38	0,15

8	2	0	0	1	**43**	3	0	0	0	2	51	0,84	0,76	0,61	0,40

* Combinations of answers are possible, good if only C, D and E are given

** Combinations of answers are possible, good if only D is given

Here, a plan will be discussed for improving the data literacy questionnaire. The questionnaire consisted of two parts: data evaluation with test questions and data application with statements. The reliability analysis shows that the reliability of the data evaluation part is .41 and data application is .60. The results will be discussed in detail.

Data evaluation

De Kuder Richardson (an alternative to Cronbach's alpha for dichotomous items) gives the reliability of the questionnaire. As mentioned, Kuder Richardson is .41 for the data evaluation part. An alpha higher than 0.8 is desirable, but 0.6 is often sufficient for tests. Items 1, 4, and 8 have good inter-item correlation > 0.3.

The questionnaire is not reliable enough yet to use it in the study. That is why some general adjustments will be discussed below.

There are different response scales for the data evaluation part. It has been decided to adjust all response scales to three options instead of four. Research by Paes and ten Cate (2009) examined several papers and came to the conclusion that three-choice questions are preferable to questions with four or five options. Items 5, 6, 7 and 8 will also be changed to multiple-choice items with three answer options. In addition, the description "Figuur x" will be removed for all questions because the number does not match the question, which causes confusion.

A test analysis was carried out to further inspect the items. The Rir values are especially taken into account. The Rit value has also been calculated, but is less applicable to the data evaluation part (< 20 items) because the score weighs quite heavily. The results of the test analysis can be found in Table 7.

Each item will now be analyzed individually:

Competency Data interpretation (Understanding data)

Item 1 'Sollicitanten bij Water4U': The Rir value of this question is high, 0.35, which means that the good students have the question right and the bad students choose the distractor. Because the inter-item correlation of this question is also good, it has been decided to keep this question and only adjust it to three multiple-choice options instead of four. This also means that one name will be taken from the radar chart, this will be Meike. However, because the p-value (0.88) of the item is rather high, the graph will be made a bit more difficult by making the outliers somewhat less clear.

Item 2 'Plantengroei': The Rir value of this question is low, 0.06, and the p-value is too high, 0.91. This can indicate a question that can be solved with common sense and is therefore too easy (Universiteit Utrecht, z.d.). This question will therefore also be made more difficult by expanding the table with several plant groups in the row and more columns about other plant specifications (e.g. hours of light they received).

Competency Identifying problems using data

Item 3 'Oorzaken van de klimaatverandering in Latveria': Because the Rir value of this question is too low, 0.13, the question will be adjusted. This also has to do with the fact that the number of multiple-choice options is also adjusted to three. The two statements are replaced by asking the question as follows: "Which of the following statements is correct?". The statements I and II will be used for this purpose, and another distractor will be added. To improve the readability of the figure, the font of the legend and the numbers will be adjusted. Because the p-value is quite high (0.78), the distractor will be difficult.

Item 4 'Omzet Water4U 2018 en 2019': This item has a high Rir value 0.30 and the p-value is also optimal (0.56). Therefore, like all other questions, this question will only be adjusted to three answer options by omitting answer options D.

Competency Data Visualisation

Item 5 'Inbraken in Nederland': Rir value scored 0.15 and the corrected p-value is even negative (-0.17). This probably has to do with the combinations that were involved in this question. Only 22% of the respondents made the right combinations. Several respondents indicated that they experienced the answer options as confusing, for example the answer option 'de figuur duidt op een foutieve toename'.
This question will be improved by converting it into a multiple-choice question with three answer options. The question will be changed into 'Welke van de volgende uitspraken is juist?' and the answer options that are currently being used may also be used in combination with each other as potential answers. There will be only one correct answer.

Item 6 'Percentage gezond gewicht per land (BMI 18.5 - 24.9)': The Rir values of this question are negative: -0.13, and the p-value is lower than average but acceptable (0.36). This means that other alternatives are also plausible, but this could also have to do with the fact that multiple answer options are also possible with this question. Answer option A is a good distractor (13 participants chose this option), but is incorrect. Furthermore, the data 'BMI 18.5 - 24.9' provides superfluous information. This will therefore also be removed from this question.
The question will be improved by transforming it into an MC question with three answer options and formulating the question as follows: "Welke van de volgende uitspraken is juist?".

Competency Data-Driven Decision Making

Item 7: 'Samenwerkingsscores bij Water4U': This item scored a fairly high p-value (0.88) and the Rir value was rather low (0.15). Because items 7 and 8 are too similar, it has been decided to remove the item with the worst Rir value. This means that item 7 will get a completely different shape with a different 'widget'. Just like item 8, this question will tie in with the Data-Driven Decision Making competency.

Item 8 'Samenwerkingscores op schaal van 1 tot 5': This item has an acceptable Rir (0.40) and p-value (0.76), so it has been decided to retain this question as good as possible. Just like the other items, this question will be adjusted to a question with three answer options. The introduction will be removed and the participant has to pick the correct statement. In addition, the title will be adjusted to 'Samenwerkingsscores bij Water4U op een schaal van 1 tot 5'.

Data application

The items in the questionnaire are not numbered consecutively. This is also an improvement for the next version. For now, these items will be numbered consecutively to prevent confusion.

Competency Critical thinking

9. Als ik kijk naar figuren of grafieken, ben ik altijd kritisch of dit de juiste manier is waarop de gegevens weergeven kunnen worden.

10. Ik neem niet zomaar data over van anderen, zonder er nauwkeurig naar te kijken.

11. Ik ben mij bewust van de risico's en de uitdagingen die komen kijken bij grote hoeveelheden data.

Competency Data culture

12. Ik werk graag bij een organisatie waar er veel met data wordt gewerkt en dit gebruikt wordt voor het maken van beslissingen.

13. Ik zie data als een belangrijk onderdeel van het dagelijks leven.

14. Ik denk dat ik zonder het gebruik van data fenomenen kan onderzoeken (reversed item).

A reliability analysis has been performed for the statements in SPSS. The Cronbach's alpha of the competency data application was 0.6, which can be considered as questionable.

However, it is more important to look at the inter-item correlation instead of Cronbach's alpha, because there are fewer than 10 items (Pallant, 2020). For each item the inter-item correlation is higher than > 0.3, except for items 13 and 14. The content of these statements will be adjusted, for example Item 13 could be: 'Ik ben elke dag bezig met data voor mijn werk of in mijn vrije tijd'. In addition, some participants provided feedback that item 14 was formulated confusingly. This will also be taken into account.
The factor analysis with oblimin rotation (because the factors were correlated) shows that the items do not yet correlate to the two factors they belong to. The validity of the statements can therefore still be improved. The statements will be changed so that they are more in line with the concept they belong to, and that better distinction between the factors can be made.

Pilot

Now that many suggestions have been made following the first pilot, the improved questionnaire will be made and tested in a second pilot. For the second pilot, the same target group will be chosen as the first pilot: anyone over the age of 18 can participate. In this way the difference in data literacy remains as large as possible. However, people who have already participated in the first pilot are now excluded from participation. Because a large number of participants took part in the first pilot, the number of participants will now be around 20. The pilot will again be conducted with Google Forms and the results will be analyzed in the same way. It is expected that the reliability and validity after the above adjustments will be higher than before. After approval of the research plan, the aim is to immediately carry out the pilot and to have the results within a week. In this way, follow-up steps can be taken quickly and the real investigation can begin.

Appendix J

Item analysis data evaluation

Table 8

Item analysis data evaluation

	Alternative			Item analysis			
Question	A	B	C	N_participant	P-value	Corrected p-value	RIR-value
1	0	3	**26**	29	0,90	0,84	
2	2	**27**	0	29	0,93	0,90	.51
3	0	**28**	1	29	0,97	0,95	.10
4	2	2	**25**	29	0,86	0,79	.27
5	2	**27**	0	29	0,93	0,90	.51
6	3	**25**	1	29	0,86	0,79	
7	5	4	**20**	29	0,69	0,53	.24
8	**29**	0	0	29	1,00	1,00	

Appendix K

Informatiebrief onderzoek over teacher dashboards

Proefpersoneninformatie voor deelname aan (sociaal)-wetenschappelijk onderzoek
Teacher dashboards

Utrecht, 15 maart 2020

Geachte heer, mevrouw,

Leuk dat u mee wilt doen aan het onderzoek!

Middels deze brief wil ik uw toestemming vragen om mee te doen aan het onderzoek over ***teacher dashboards***. Met zogenaamde "teacher dashboards" bedoelen we computerschermen die de activiteiten van leerlingen automatisch volgen en de leerkracht hierover informatie geven tijdens de les, bijvoorbeeld het leerkrachtscherm van Snappet. In dit onderzoek worden teacher dashboards weergeven waarop leerkrachten de voortgang van hun samenwerkende leerlingen kunnen zien. Er worden straks teacher dashboards weergeven van leerlingen die ***samenwerken in tweetallen***, waarbij zij ***breukopgaven*** maken. Voor de juiste weergave van de teacher dashboards, is het belangrijk dat u dit onderzoek invult ***op de computer***.

Achtergrond onderzoek

De laatste jaren wordt er steeds meer technologie in de klas gebruikt. Ook zijn er steeds meer technologieën die de samenwerkingsprocessen van leerlingen op de computer weergeven. Dit vraagt echter wel om nieuwe vaardigheden voor docenten. Wat deze vaardigheden precies zijn, is nog weinig onderzocht.

Opzet van het onderzoek

Dit onderzoek bestaat uit meerdere vragenlijsten en een onderdeel waarbij er naar teacher dashboards gekeken wordt. Het onderzoek begint met vragenlijsten over uw achtergrondkenmerken, uw interpretatie en houding met betrekking tot data en uw ervaring met samenwerken. Daarna zal u op de computer verschillende teacher dashboards bekijken en vervolgens een aantal vragen hierover beantwoorden. Het onderzoek eindigt met een vragenlijst over wat u van het onderzoek vond.

Wat wordt van u als participant verwacht

Dit onderzoek vraagt ***ongeveer 30 minuten*** van uw tijd. Het is de bedoeling dat u het onderzoek zo serieus mogelijk invult.

Voordelen van het onderzoek

Door mee te doen aan dit onderzoek draagt u bij aan de wetenschap over teacher dashboards. Daarnaast is het voor u een mogelijkheid om een teacher dashboard te ervaren, en mogelijk te gebruiken bij de beslissing of u dit zou willen implementeren in de klas. Ook worden er onder de deelnemers van dit onderzoek ***twee Bol.com cadeaukaarten*** (2x 20 euro) verloot.

Vertrouwelijkheid verwerking gegevens

Voor dit onderzoek is het nodig dat wij een aantal persoonsgegevens van u verzamelen. Deze gegevens hebben wij nodig om de onderzoeksvraag goed te kunnen beantwoorden, dan wel om u te kunnen benaderen voor vervolgonderzoek. De persoonsgegevens worden op een andere computer opgeslagen dan de onderzoeksgegevens zelf (de zgn. ruwe data). De computer waarop de persoonsgegevens worden opgeslagen is volgens de hoogste normen beveiligd en alleen betrokken onderzoekers hebben toegang tot deze gegevens. De gegevens zelf zijn ook beveiligd d.m.v. een beveiligingscode. Alle data zal geanonimiseerd worden. Als u wilt stoppen met het onderzoek mag u dat zonder opgaaf van reden op elk moment aangeven. Uw gegevens zullen voor minimaal 10 jaar bewaard worden. Dit is volgens de daartoe bestemde richtlijnen van de VSNU. Meer informatie over privacy kunt u lezen op de website van de Autoriteit Persoonsgegevens: https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/avg-europese-privacywetgeving

Vrijwilligheid deelname

Deelname aan dit onderzoek is vrijwillig. U kunt op elk gewenst moment, zonder opgave van reden en zonder voor u nadelige gevolgen, stoppen met het onderzoek. De tot dan toe verzamelde gegevens worden wel gebruikt voor het onderzoek, tenzij u expliciet aangeeft dit niet te willen.

Contactpersoon en klachtenfunctionaris

Als u vragen of opmerkingen over het onderzoek heeft, kunt u met mij contact opnemen via ***l.gerrits@students.uu.nl***. Als u een officiële klacht heeft over het onderzoek, dan kunt u een mail sturen naar de klachtenfunctionaris via klachtenfunctionaris-fetcsocwet@uu.nl

Als u na het lezen van deze informatiebrief besluit tot deelname aan het onderzoek zou ik u willen vragen om toestemming te geven en door te gaan naar het onderzoek.

Met vriendelijke groet, Laura Gerrits

- ❑ Hierbij verklaar ik de informatiebrief m.b.t. onderzoek teacher dashboards gelezen te hebben en akkoord te gaan met deelname aan het onderzoek.

Appendix L

FETC Study registration form

Section 1: Basic Study Information

1. Name student:

Laura Gerrits

2. Name(s) of the supervisor(s):

Anouschka van Leeuwen

3. Title of the thesis (plan):

A Study On Teacher Characteristics and The Use Of CSCL Dashboards

4. Does the study concern a multi-center project, e.g. a collaboration with other organizations, universities, a GGZ mental health care institution, or a university medical center?

No

5. Where will the study (data collection) be conducted? If this is abroad, please note that you have to be sure of the local ethical codes of conducts and permissions.

Due to the COVID-19 virus, the data collection took place online via the online environment Gorilla. All participants received information about the study prior to the research and subsequently they signed an online informed consent.

Section 2: Study Details I

6. Will you collect data?

Yes □ Continue to question 11

7. Where is the data stored?

8. Is the data publicly available?

Yes / No If yes: Where?

9. Can participants be identified by the student? (e.g., does the data contain (indirectly retrievable) personal information, video, or audio data?)

Yes / No If yes: Explain.

10. If the data is pseudonymized, who has the key to permit re-identification?

Section 3: Participants

11. What age group is included in your study?

The age of the participants will differ between 16 and 67 years.

12. Will be participants that are recruited be > 16 years? Yes
13. Will participants be mentally competent (wilsbekwam in Dutch)? Yes
14. Does the participant population contain vulnerable persons? (e.g., incapacitated, children, mentally challenged, traumatized, pregnant) No
15. If you answered 'Yes' to any of the three questions above: Please provide reasons to justify why this particular groups of participant is included in your study.

My participants are PABO or ALPO students and teachers with a degree who are currently teaching. So, that can be beginning students or teachers who have been in the profession for a long time.

16. What possible risk could participating hold for your participants?

There is a risk that the participants will find the investigation too long and want to stop.

17. What measures are implemented to minimize risks (or burden) for the participants?

As described in the informed consent form, every participant can stop when they want to. Besides that, the length of the questionnaires was taken into account. These have been kept as short as possible.

18. What time investment and effort will be requested from participants?

The participant must indicate whether he or she wants to participate in the study and reserve time for this in their agenda. The study will take approximately 35 minutes of their time.

19. Will be participants be reimbursed for their efforts? If yes, how? (financial reimbursement, travelling expenses, otherwise). What is the amount? Will this compensation depend on certain conditions, such as the completion of the study?

Two Bol.com vouchers of 2x 20 euros were raffled among the participants. The vouchers will only be raffled among the participants who have completed the research completely.

20. How does the burden on the participants compare to the study's potential scientific or practical contribution?

Very little research has been done on data literacy in general. There has also been little research into the relationships between experience with collaboration and data literacy on the use of CSCL dashboards. Participating in the research would therefore greatly contribute to the scientific gap we currently have.

21. What is the number of participants? Provide a power analysis and/or motivation for the number of participants. The current convention is a power of 0.80. If the study deviates from this convention, the FERB would like you to justify why this is necessary.

(Note, you want to include enough participants to be able to answer your research questions adequately, but you do not want to include too many participants and unnecessarily burden participants.)

The number of participants to reduce a type II error was determined through a power analysis with G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). With a medium effect size of 0.5 and a significance level of ($\alpha = .05$), two-tailed, the ideal sample size is ≥ 29 participants)

22. How will the participants be recruited? Explain and attach the information letter to this document.

This study will use the self-selected sampling method, which means participants with certain characteristics can participate voluntarily. The acquisition of participants will be mainly done by contacting schools and asking teachers.

23. How much time will prospective participants have to decide as to whether they will indeed participate in the study?

Two weeks

24. Please explain the consent procedures. Note, active consent of participants (or their parents) is in principle mandatory. Enclose the consent letters as attachments. You can use the consent forms on Blackboard.

All participants received an active informed consent form to take into account the ethical guidelines. The purpose and procedure of the investigation are discussed here. It also states that all data will be treated confidentially and anonymously and that every participant can indicate when they want to stop. The informed consent letter is also enclosed in this research paper.

25. Are the participants fully free to participate and terminate their participation whenever they want and without stating their grounds for doing so? Explain.

Yes, participating is voluntary and participants can stop whenever they want to without giving a reason for it.

26. Will the participants be in a dependent relationship with the researcher?

The participants will in part be acquaintances of the researcher. The researcher does not know the other participants.

27. Is there an independent contact person or a general email address of a complaint officer whom the participant can contact?

yes, klachtenfunctionaris-fetcsocwet@uu.nl

Section 4: Data management

28. Who has access to the data and who will be responsible for managing (access to) the data?

During the writing of the master thesis, only the researcher will have access to the data. If the master thesis is to be submitted, the data and analyzes will be shared with the supervisor and the second assessor in a secure environment called YoDa.

29. What type of data will you collect or create? Please provide a description of the instruments.

This research mainly collects quantitative data through surveys. Some open questions will be coded.

Questionnaire with background information. In this questionnaire, participants were asked about their age, gender, in which group they teach, their experience with fractions, experience with collaboration and experience with dashboards to report the descriptive statistics (see Appendix B). The questionnaire consisted of 10 items.

Questionnaire data literacy. The data literacy of the participants was tested with a questionnaire, consisting of 14 items (see Appendix C). Since there was no existing questionnaire about data literacy, it was decided to construct a questionnaire based on the conceptual framework of Ridsdale et al. (2015). The adapted knowledge areas were data evaluation and data application (see Appendix A).

- ***Data evaluation*.** In this knowledge area, the competencies were measured with a test. For each question, the participants were presented a figure and then received a question about it. The questionnaire from Van den Bosch, Espin, Chung and Saab (2017) about graph literacy has been used as inspiration. The questions were multiple-choice, with three answer options (See Figure 2). In total, the knowledge area data evaluation consisted of eight items. The data literacy score on data evaluation was calculated using the following formula: grade = (9 * number of correct answers / 8) + 1.
- ***Data application.*** In this knowledge area, the competencies were measured with statements on a five-point Likert scale (1 = completely disagree, 5 = completely agree). The six statements were based on the adjusted conceptual framework of Ridgdale et al., (2015), see Appendix A. Data application consisted of critical thinking (e.g., 'I am aware of the risks involved with large amounts of data, such as a data breach or privacy-sensitive datasets') and data culture (e.g., 'I like working at an organization where data is used to make informed decisions'). Both competencies consisted of three items each (statement 1 - 3 represented critical thinking and 4 - 6 represented data culture). A high score on a statement indicated a high application of data literacy. The average score of all statements was calculated for each participant.

Questionnaire experience with collaboration. The tool used to measure teachers' experiences with collaboration was based on the Lee and Tsai (2011) scale. This is an

instrument on student perceptions of experience with collaboration, self-regulated learning and information seeking in a traditional classroom setting and internet-based learning. The validity of the instrument was good, and the reliability was satisfactory (between .84 and .91). It was decided to use the questions on student perceptions of experience with collaboration in a traditional classroom setting, and convert these student perceptions into teacher perceptions (See Appendix D). A five-point Likert scale (1 = never, 5 = very frequent) measured the ideas that teachers have about collaborative learning and consisted of seven items (e.g. "I think it is important that students discuss problems that arise during learning with their classmates"). A high score on this scale indicated that teachers had specific ideas about collaboration, since they know what to look for when observing and assessing the class during collaborative learning, due to their developed knowledge structures about good collaboration.

Dashboard trials. Dashboard trials were used to test whether data literacy and teacher experience with collaboration were related to how teachers use the dashboard (Appendix E). These dashboard trials were adapted from the research by Van Leeuwen et al. (2019) in which they presented five dyads by means of three types of dashboards. The three types of dashboards were different in terms of support: mirroring, alerting and advising. In this study, only the mirroring dashboard was used, where the results were displayed without support or advice being given to the participant. This type of dashboard makes it possible to see whether data literacy could make a difference. The information on the dashboards was the result of dyads working together on assignments about fractions. Eight different dashboards were displayed to the participant with a description of the situation before each dashboard trail. The participant was given the task to look at the different dyads on the dashboard. The dashboard displayed six data widgets per dyad (see Figure 3). In the classroom, teachers do not have the time to study a dashboard extensively, due to other tasks that the teacher has to do while teaching (van Leeuwen & Rummel, 2020). The participants were therefore given 50 seconds per dashboard to imitate the classroom situation.

- ***Detection.*** The participants first received one question related to *detection*: in which group did they think something was going on in this situation. In all dashboard trials, there was only one correct answer. As described in the theoretical background, there are six examples of problems that can arise in a group during collaboration. In each dashboard, there was one dyad where such a problem arose (see Appendix F). The score of all detection questions was calculated using the following formula: grade = (9 * number of correct answers / 8) + 1.
- ***Interpretation.*** Next, participants were asked about the *interpretation* of CSCL dashboards: what did they think was the problem. This open question was quantified using a coding scheme (see Appendix G). The coding scheme consisted of three dimensions: specificity, attitude and use of data widgets. Based on research by Van Leeuwen et al. (2019), naming many specific elements results in higher scores on specificity. If the situation was described and linked to a learning principle, this resulted in a high score on attitude. Finally, if more than two data widgets were used and associated with each other, this led to a high score on the use of data widgets. Since

each dimension had a different number of points that could be allocated, a weight was assigned to the dimensions (specificity = 1, attitude = 2.5, data widgets = 1.66). As a result, the total score on interpretation can be calculated by means of a grade (score interpretation = (9 * (specificity + attitude + data widgets) / 15) + 1. Two assessors rated 25% of the comments on the interpretation question. There was a 75.56% agreement between the assessors, which according to Cicchetti and Sparrow (1981) can be considered as excellent. It was therefore decided to have the rest of the answers coded by a single assessor.

Questionnaire about the study. Participants received a questionnaire at the end of the investigation about how they experienced it, based on the study by Van Leeuwen et al. (2019) (see Appendix H). One item was added to the questionnaire, "I had the idea that I had the right skills to read the dashboard". The questionnaire therefore consisted of a total of six items. There were four statements with a five-point Likert scale (1 = completely disagree, 5 = completely agree), such as "The investigation procedure was clear" and two open questions, e.g., "Can you indicate which strategy you generally used to deal with situations on the dashboard?"

30. Will you be exchanging (personal) data with organizations/research partners outside the UU?

No

31. If so, will a data processing agreement be made up?

32. Where will the data be stored and for how long?

The data will be stored on YoDa and on my personal computer for 7 years.

33. Will the data potentially be used for other purposes than the master's thesis? (e.g., publication, reporting back to participants, etc.)

Yes, when participants are interested in the results, reporting back is an opportunity.

Appendix M

Timetable

Wat	**Wanneer**
Literatuur over teacher dashboards bekijken	10 - 13 november
Onderzoeksmogelijkheden uitwerken en schrijven van een outline	13 - 21 november
Verder met outline en feedback verwerken	21 november - 2 december
Begeleidingsgesprek	2 december
Probleemstelling schrijven en nadenken over instrumenten	3 december
Beginnen theoretisch kader	5 december - 11 december
Methode en instrumenten ontwikkelen	12 - 17 december
Poster presentatie en overleg	*18 december*
Feedback verwerken	21 en 22 december
Theoretisch kader	26 december
Theoretisch kader	27 december
Data literacy vragenlijst	28 december
Experience with collaboration vragenlijst	29 december
Werken aan conceptversie	1 - 5 jan
Feedback session	15 januari
Feedback verwerken	15 - 20 januari

Pilot uitzetten	21 januari
Codeerschema maken	22 - 23 januari
Methodesectie Pilot analyseren Theoretisch kader	24 - 25 januari
FETC form en informed consent Alles nalopen	26 januari
Feedback session	5 februari
Verwerken feedback	7 - 8 februari
Test analyse uitvoeren	10 - 12 februari
Improvement plan schrijven	13 februari
Onderzoeksplan verbeteren	15 - 16 februari
Nieuwe feedback verwerken	18 - 19 februari
Laatste stappen definitief onderzoeksplan	20 - 21 februari
Pilot 2 voorbereiden	28 februari
Data verzamelen pilot 2	28 februari - 3 maart
Begeleidingsgesprek	11 maart
Participanten werven	12 - 16 maart
Definitieve vragenlijst opstellen	16 maart
Instrumenten in Gorilla	17 - 22 maart

Start dataverzameling	23 maart
Participanten werven en herinneringen sturen	23 maart - 7 april
Theoretisch kader verbeteren	8 april
Round tables	8 april
Begeleidingsgesprek	9 april
Eind dataverzameling	11 april
Betrouwbaarheid vragenlijsten analyseren	14 - 17 april
Digitale doorschrijfweek - UU	20 - 24 april
Resultatensectie schrijven	20 - 25 april
Discussiesectie schrijven	27 april - 1 mei
Discussie en theoretisch kader	4 - 7 mei
Theoretisch kader perfectioneren	11 mei
Methode en resultaten perfectioneren	12 mei
Discussie perfectioneren	14 mei
Engelse zinnen verbeteren	15 mei
Deadline concept master thesis	16 mei
Discussie	19 mei
Discussie	20 mei
Methode en resultaten Slides bespreking Nikol voorbereiden	21 mei

Theoretisch kader opmerkingen verwerken Slides bespreking Nikol voorbereiden	22 mei
Theoretisch kader opmerkingen	25 mei
Theoretisch kader afmaken	26 mei
Syntax opstellen	27 mei
Engels verbeteren	28 mei
Engels verbeteren	29 mei
Syntax opstellen en discussie	31 mei
Discussie	1 juni
Theoretisch kader afmaken Appendix Scrolformulier	2 juni
Discussie	3 juni
Resultaten	4 juni
Laatste aanpassingen	5 juni
Deadline master thesis	8 juni