Supporting Complex Decision-Making: Evidence from an Eye Tracking Study on In-Person and Remote Collaboration

Citation for the original published paper (version of record):
Wisiecka, K., Konishi, Y., Krejtz, K. et al (2023). Supporting Complex Decision-Making: Evidence from an Eye Tracking Study on In-Person and Remote
Collaboration. ACM Transactions on Computer-Human Interaction, 30(5).
http://dx.doi.org/10.1145/3581787

Supporting Complex Decision-Making: Evidence from an Eye Tracking Study on In-Person and Remote Collaboration

KATARZYNA WISIECKA, Institute of Psychology, SWPS University of Humanities and Social Sciences
YUUMI KONISHI, t2i Lab, Chalmers University of Technology and Koike Lab, Tokyo Institute of Technology
KRZYSZTOF KREJTZ, Institute of Psychology, SWPS University of Humanities and Social Sciences
MAHSHID ZOLFAGHARI and BIRGIT KOPAINSKY, System Dynamics Group, University of Bergen
IZABELA KREJTZ, Institute of Psychology, SWPS University of Humanities and Social Sciences
HIDEKI KOIKE, Koike Lab, Tokyo Institute of Technology
MORTEN FJELD, t2i Lab, Chalmers University of Technology and t2i Lab, University of Bergen

This article examines the attentional mechanism of in-person collaboration by means of System Dynamics-based simulations using an eye tracking experiment. Three experimental conditions were tested: in-person collaboration, remote collaboration, and single user. We hypothesized that collaboration focuses users' attention on key information facilitating decision-making. Collaborating participants dwelt longer on key elements of the simulation than single users. Moreover, in-person collaboration and single users yielded a strategy of decision-making similar to an optimal strategy. Finally, in-person collaboration was less cognitively demanding and of higher quality. The contribution of this article is a deeper understanding of how in-person collaboration on a large display can help users focus their visual attention on the most important areas. With this novel understanding, we believe collaborative systems designers will be better equipped to design more effective attention-guiding mechanisms in remote collaboration systems. The present work has the potential to advance the study of collaborative, interactive technologies.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **Collaborative interaction**; Empirical studies in interaction design;

78

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program–Humanities and Society (WASP-HS), funded by the Marianne and Marcus Wallenberg Foundation and the Marcus and Amalia Wallenberg Foundation. This work was partially supported by UiB strategic collaboration funds. This research is funded by MediaFutures partners and the Research Council of Norway (grant number 309339). Open access was financed by the European Union resources within the European Social Fund no. POWR.03.02.00-00-I054/16-00.
Authors' addresses: K. Wisiecka, K. Krejtz, and I. Krejtz, Eye Tracking Research Center, Institute of Psychology, SWPS University, ul. Chodakowska 19/31, 03-815 Warsaw, Poland; emails: {kwisiecka, kkrejtz, ikrejtz}@swps.edu.pl; Y. Konishi, t2i Lab, Chalmers University of Technology, Betongvagen 5, 412 58, Gothenburg, Sweden, and Koike Lab, Tokyo Institute of Technology, 2-12-1 W8-E401, Ookayama, Meguro-ku, Tokyo 152-8550, Japan; email: koni.forut@gmail.com; M. Zolfaghari and B. Kopainsky, System Dynamics Group, University of Bergen, P.O. Box 7800, 5020 Bergen, Norway; emails: mahshid.zolfaghari@gmail.com, birgit.kopainsky@uib.no; H. Koike, Koike Lab, Tokyo Institute of Technology, 2-12-1 W8-E401, Ookayama, Meguro-ku, Tokyo 152-8550, Japan; email: koike@c.titech.ac.jp; M. Fjeld, t2i Lab, Chalmers University of Technology, Betongvagen 5, 412 58, Gothenburg, Sweden and t2i Lab, University of Bergen, Museplassen 1, 5007 Bergen, Norway; email: morten.fjeld@uib.no.

1073-0516/2023/09-ART78 $15.00
https://doi.org/10.1145/3581787

Additional Key Words and Phrases: System dynamics simulation, visual attention, collaboration, eye tracking, natural resource management

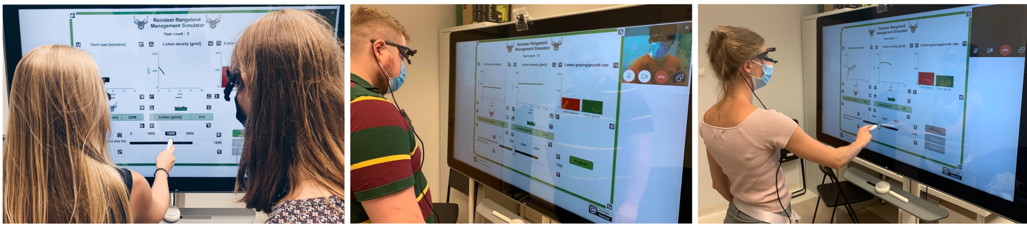
ACM Reference format:
Katarzyna Wisiecka, Yuumi Konishi, Krzysztof Krejtz, Mahshid Zolfaghari, Birgit Kopainsky, Izabela Krejtz, Hideki Koike, and Morten Fjeld. 2023. Supporting Complex Decision-Making: Evidence from an Eye Tracking Study on In-Person and Remote Collaboration. *ACM Trans. Comput.-Hum. Interact.* 30, 5, Article 78 (September 2023), 27 pages.
https://doi.org/10.1145/3581787

1 INTRODUCTION

As well-informed decision-making on natural resources, e.g., food, water, and construction material is getting more critical, employing natural-resource models and simulations to support knowledge workers in making decisions are becoming vital tools for good governance. Understanding and enhancing computer-supported collaboration is one of the most pressing needs of our society, where people are increasingly forced to solve complex tasks via remote collaboration. Olson and Olson [2000] predicted that remote collaboration in the new millennium would be empowered with high internet bandwidth and large displays. These advances in technology, which are currently in common use, were considered as technical requirements for remote collaborations "to come closer to some aspects of the face-to-face work" [Olson and Olson 2000, p.143]. Literature has demonstrated that large interactive surfaces improve performance and user satisfaction in various tasks involving collaborative decision-making, such as model design or data analysis [Butscher et al. 2018]. Recently, Mateescu et al. [2021] investigated collaboration on large interactive surfaces. Their findings suggest a "relatively clear advantage of the use of" large interactive surfaces over classic forms of collaboration, in particular over single-user environments (e.g., laptops). Mateescu et al. [2021] found positive effects of large interactive surfaces for knowledge gains and task-related outcomes. Collaborative decision-making requires mutual understanding of the task [Patel et al. 2012], which can be facilitated by focusing attention on key elements of the task. To our best knowledge, visual attention processes during in-person and remote collaborative decision-making have not previously been tested on large displays using the means of eye tracking.

In the present eye tracking study, participants were randomly assigned to three experimental conditions: in-person collaboration, remote collaboration, and single-user (Figure 1). Their main task, in all conditions, was to make several decisions when interacting with a **System Dynamics-based** (**SD-based**) simulation of reindeer rangeland management. The reindeer rangeland management task is representative of a large class of dynamic decision-making issues in natural resource management, such as deforestation and forest degradation, biodiversity loss, ecosystem degradation, reduction in soil quality, and fall in available water quantity [Tietenberg and Lewis 2019]. System Dynamics is a model-based approach to dynamic decision-making and policy analysis [Sterman 2000], often used to understand the parameters involved, their interactions, and how decisions on even a single parameter can affect complex system dynamics. The SD-based simulations are intended to help in this complex decision-making by visually presenting the parameters of the system and the effects of single decisions. Related research [Guy et al. 2013; Moxnes 1998b; Sterman and Sweeney 2007] shows that decision-makers tend to underestimate what it takes to restore depleting natural resources and that they rely on wait-and-see strategies, with, at times, disastrous long-term consequences.

Effective decision-makers need to focus their attention on key information of an SD-based simulation user interface [Bentzen et al. 2011; Milkman et al. 2009]. In the present study, we investigated visual attention distribution over key elements of the interactive simulation task by recording participants' eye movements during the task. In general, we hypothesized that collaboration would



(a) In-person collaboration (b) Remote collaboration (c) Single user

Fig. 1. Study participants collaborating on the system interface resolving a natural resource management task (Reindeer Task, [Moxnes 2014]) in (a) in-person collaboration, (b) remote collaboration, and (c) single user condition. Each participant is wearing a mobile eye tracking device.

facilitate focusing participants' visual attention on the most important information presented on the large display, and thus more likely to lead to better decisions. We present our detailed hypotheses after a brief review of relevant literature.

2 BACKGROUND

Collaboration has a strong potential to tackle problems posed in decision-making since it may facilitate emergent strategies [Cohen 1994] and novel problem perception and descriptions [Schwartz 1995; Shirouzu et al. 2002]. It is often described as a process of constructive problem exploration to find solutions that transcend each partner's individual point of view [Wood 1991]. However, just working in a group does not necessarily guarantee success. Often, there can be difficulties in group work leading to time wasted, discouragement, and, in the end, lack of progress [Barron 2003]. The success of problem-solving varies significantly among groups, even when group members have comparable knowledge levels [Hogan et al. 1999; Webb et al. 2002]. It is critically important to better understand visual attention mechanisms underlying collaboration on complex decision-making tasks in order to design effective tools and methods supporting remote collaboration.

2.1 Visual Attention during Collaboration

During collaboration, people share focus on an object, indicating to each other their course of attention. Individuals can either indicate things physically (deictic gestures), verbally (describing the object of interest), or by using their body position and orientation. Partners in social interaction also need to demonstrate awareness that they are working on something in common via nonverbal signals, such as gaze direction [Tomasello 1995]. Mutual gaze and gaze-following represent processes involving two individuals. Their actions lead to joint visual attention that is created by following and directing another person's gaze to a new target making a referential triangle [Pfeiffer et al. 2013]. This process is considered an important aspect of the understanding of other minds and building shared reality during collaboration [Echterhoff et al. 2009].

While there are many methods to study communication during interaction, such as head [Cognolato et al. 2018; Müller et al. 2018; Yu and Smith 2017] and gesture analysis [Gullberg and Holmqvist 1999; Kendon 1988], we chose to use mobile eye tracking to examine gaze patterns. Eye tracking enabled us to register participants' attention allocation during task performance as a measure of shared visual workspace [Fussell et al. 2000]. Eye tracking is considered to be an effective tool to study gaze behavior during interaction and cooperation [Pietinen et al. 2008a]. For example, gaze behavior measured with mobile eye tracking was used as an index of turn-taking during in-person dyadic [Jokinen et al. 2009, 2010] and triadic interaction [Holler and Kendrick 2015]. In another study, Gergle and Clark [2011] showed in a dyadic mobile eye tracking experiment that when pairs have shared visual reference they rely less on language in their communication. However, when their visual attention was not coordinated, specific referential forms such as deictic references can be used to direct their attention to the referent object during collaboration. Wang and Shi [2019] showed that awareness of a partner's gaze direction changes gaze behavior. Their results indicate that an interaction partner's gaze is more likely to lie on the referred object when the subject knows the helper can see his/her eyes. Awareness of the partner's gaze direction enhances joint attention and quality of performance, which has been demonstrated in eye tracking studies on in-person collaboration [Schneider et al. 2018].

In computer-supported collaboration, non-verbal communication such as gaze cueing to shared reference is often lost, especially in remote settings. In remote collaboration, visual attention has been taken into account both in early and recent works [Otsuki et al. 2017; Shiro et al. 2018; Vertegaal 1999]. Eye tracking studies examined methods of facilitating joint attention, e.g., by gaze display [Bednarik et al. 2011; D'Angelo and Gergle 2018; Kütt et al. 2019; Maurer et al. 2017; Velichkovsky 1995; Yao et al. 2018]. For instance, Velichkovsky [1995] used projection of eye fixations, which increased the efficiency of problem solving by experts and novices. However, it is still difficult to find such solutions in off-the-shelf systems. Eye tracking has also been used to uncover cognitive processing within visualization evaluation studies [Goldberg and Helfman 2011]. By combining eye movement data and interaction logs, Blascheck et al. [2018] studied how non-expert users "discover the functionality of an interactive visualization". Visual attention of two or more collaborators has also been taken into account in the design of large vertical and horizontal user interfaces during in-person collaboration on large displays [Kontogiorgos et al. 2018; Serim et al. 2018; van der Meulen et al. 2016]. Large displays enable simultaneous use by multiple users, providing at the same time high visibility to all the parties, and support joint work [Pietinen et al. 2008b; van der Meulen et al. 2016].

2.2 Computer-Supported Collaboration Tools

An extended body of literature points out that computer-supported collaboration tools have proven their capacity to facilitate collaboration by supporting workspace awareness, an "up-to-the-moment understanding of another person's interaction with a shared workspace" [Gutwin et al. 1996]. Lopez and Guerrero [2017] underlined the importance of the system providing information about partners' momentary actions during collaboration [Dourish and Bellotti 1992]. Similarly Heer and Agrawala [2007] underlined the role of workspace awareness and sense-making during computer-supported social interaction. They presented design considerations for collaborative visual analytics to improve workspace awareness. In line with this, Burch and Schmauder [2018] proposed several collaborative interaction techniques, e.g., the highlighting of user interface elements selected by other users for providing additional context for collaborative problem-solving on large high-resolution displays.

A large display could be used by single or multiple users as they have enough space to freely walk around and explore an individual sub-area of the display. Indeed, Kim and Snow [2013] showed that collaboration on a large shared display influences task efficiency, depending on how individuals effectively cooperate with others in the display context and with multiple inputs. Both asynchronous access and multiple input lead users to take on both separate and cooperative roles in task performance. Consequently, peer interaction increases attention on work practices and verbal communication in collaborating groups. Moreover, elements of non-traditional interaction techniques might support group collaboration, e.g., in collaborative learning situations [Schneider et al. 2016]. With such interaction techniques, e.g., sharing physical objects or space, students can intuitively explore complex systems, and collaborative groups find it easier to establish common ground for work [Falcao and Price 2011; Schneider et al. 2012; Valdes et al. 2012].

Along the same lines of thought, Arias et al. [2000] presented a large display prototype aiming to create "shared understanding" through discussion and negotiation, that ironically shifted attention away from the computer toward the interpersonal relationship and understanding between working partners. In an empirical study, Liu et al. [2016] demonstrated that wall-sized displays encouraged collaborative manipulation, reduced physical navigation and fatigue, and improved collaboration efficiency. Gorkovenko et al. [2018] showed that collaborative data exploration on a large, high-resolution display in comparison to a tablet-size display evoked less cognitive demands on users. However, large displays are not only beneficial for interaction with the system. Jakobsen and Hornbæk [2013] showed in two experiments that for making screen size beneficial, it is important to take into consideration the interaction of display size, information space, and scale ratio.

We postulate that decision-supporting systems like SD-based simulation presented on large displays, foster awareness of the problem and focus collaborating partners' attention on the most important information necessary to solve the collaborative task. Focusing visual attention on key elements of a simulation task during the decision-making process may be treated as a cognitive mechanism standing behind workspace awareness during computer-supported collaboration. In our study, we extend previous works through the inclusion of remote collaboration on large displays. Using a shared large display during in-person collaboration allowed for direct comparisons of attention mechanism during in-person and remote collaboration.

3 THE PRESENT STUDY

Our study examined visual mechanisms of collaboration during complex decision-making. This research was conceived in the context of an important societal challenge: sustainability in natural resource management. More specifically, we researched the management of natural resources that are threatened by over-exploitation and collapse. For our natural resource management simulation, we chose a reindeer rangeland management task. This task's structure is representative of a wide range of natural resource management tasks, including, for example, mitigation of climate change, and builds on benchmarking data from previous experiments [Derwisch et al. 2011; Moxnes 2004]. While most SD-based simulation tasks were designed for and tested on single users, a few studies have discussed the potential of multiplayer SD-based simulations [Happach and Schoenberg 2017]. In our work, we try to fill the gap by examining collaborative decision-making on large displays with the use of an eye tracking method. We investigate the visual attention process during collaborative decision-making on large interactive displays. By monitoring eye movements during collaboration, we uncover a cognitive mechanism underpinning workspace awareness during computer-supported collaboration. The new contribution of this study is a deeper understanding of how in-person collaboration on large displays can help users focus their visual attention on the most important areas [Bentzen et al. 2011]. With this novel understanding, we believe system designers will be better informed when designing attentive user interfaces [Vertegaal 2002, 2003] with potential for remote collaboration.

3.1 Hypotheses and Variables

The present study verifies four main hypotheses in relation to the three experimental conditions: in-person collaboration, remote collaboration, and single user condition (the between-subjects independent variable). First, we predicted that making a collaborative decision is highly cognitively demanding and can be manifested in the characteristics of eye movement fixations [Duchowski et al. 2020; Krejtz et al. 2020]. The *eye-mind assumption* suggests a close relationship between the average length of fixation duration and depth of cognitive processing [Just and Carpenter 1976].

Based on this assumption, we expected (hypothesis 1) to observe a longer average fixation duration reflecting deeper information processing while looking at the decision-related areas, compared to simulation control areas [Krejtz et al. 2016]. The within-subjects independent variable related to this hypothesis was areas-of-interest (AOI), which had two levels: simulation control vs. decision-making regions (see Figure 4(b)).

Second, we expected that collaboration would trigger the visual attention of partners to the most important visual information presented in simulation control areas, which, in turn, could influence system understanding and the decision-making strategy. We expected in this case to observe an attention bias toward the graph presenting the dynamics of lichen density (hypothesis 2). The attentional bias would manifest in a longer viewing time (dwell time as a dependent variable) while looking at the lichen density graph in comparison to other system dynamics visualizations. The within-subjects independent variable related to this hypothesis differentiated between simulation control graphs: lichen density vs. lichen growth vs. herd size graph (see Figure 4(b)).

Third, we aimed to examine how the process of collaborative decision-making and distribution of visual attention was affected by the distance between partners (in-person vs. remote collaboration condition). We anticipated that collaboration would lead to more effective task-solving than the single user condition. Related research [Guy et al. 2013; Moxnes 1998a, b; Sterman 2008; Sterman and Sweeney 2007] shows that decision-makers tend to underestimate what it takes to restore depleting natural resources and that they rely on wait-and-see strategies, with, at times, disastrous long-term consequences. That was reconfirmed in recent work [Gary and Wood 2016; Moxnes 2014; Nyam et al. 2020; Perissi et al. 2017], for which a detailed review would fall out of the scope of the present article. Our task was a simplified structure of a natural resource system where users could easily draw wrong conclusions about resource management. Specifically, we hypothesized that in-person collaboration would promote a strategy of decision-making similar to the optimal strategy derived from a dynamic model simulation which can be observed at the behavioral level in participants' decisions recorded with system logs (hypothesis 3). The first independent within-subjects variable related to this hypothesis was the decisions made during the simulation task (16 decisions). The second independent variable differentiated between optimal and experimental decisions.

Finally, we assessed the subjective quality of collaboration, task workload, and usability as a self-reported dependent variables, expecting to observe differences favoring in-person decision-making. This prediction was a natural consequence of task solving strategies presented in the preceding hypotheses, which presumed that deeper information processing during problem solving, focusing attention on key information for problem understanding and more optimal decisions, would lead to a higher subjective satisfaction from collaboration (hypothesis 4). Previous research found a link between complex problem solving and motivation, satisfaction, and attitudes toward collaboration [Albay 2019; Geister et al. 2006].

4 METHOD

To meet the study aims and verify the hypotheses, we conducted a mixed-design eye tracking experiment in which participants were randomly assigned to one of the three experimental conditions: single user, in-person collaboration, or remote collaboration. This study design allows for the triangulation of the data from three different sources: self-report questionnaires, behavioral decisions captured with the simulation system logs, and attentional captured with the eye tracking devices. A similar mixed-method approach has been used previously to study collaboration, for example by Mayer et al. [2018]. Next, we present details of the study method, including participants' descriptions and sampling, experimental procedure and task, study materials, and equipment.

Table 1. Demographic Characteristics of Participants in Each Experimental Condition

Variable	in-person collaboration	remote collaboration	single user	difference test
Number (n)	26 (13 pairs)	28 (14 pairs)	17	
Gender	18 females	16 females	7 females	$\chi^2(2) = 3.32, p = 0.19$
Handedness	3 left	2 left	0 left	$\chi^2(2) = 2.09, p = 0.35$
Age (years)	26.88(5.61)	27.14(4.55)	24.94(4.55)	$F(2, 54) = 1.14, p = 0.33$
Prior experience with:	$M(SD)$	$M(SD)$	$M(SD)$	difference test
Decisions systems	2.58(1.79)	1.82(1.54)	1.76(1.48)	$F(2, 54) = 1.44, p = 0.25$
ICT devices	2.65(0.65)	2.95(0.64)	2.75(0.53)	$F(2, 54) = 1.53, p = 0.23$
In-person collab.	3.81(1.39)	4.11(1.06)	4.18(1.19)	$F(2, 54) = 0.47, p = 0.63$
Remote collab.	4.04(1.25)	4.00(1.25)	3.76(1.25)	$F(2, 53) = 0.24, p = 0.79$
In-person games	2.73(1.25)	2.46(1.20)	2.12(1.11)	$F(2, 54) = 1.06, p = 0.35$
Remote games	2.37(1.19)	2.57(1.26)	2.00(1.00)	$F(2, 54) = 1.23, p = 0.30$

None of the differences for variables listed in the Table were statistically significant ($p > 0.05$). In the Table, the three first rows represent numbers (n) of participants in each experimental condition. The numbers presented in the following rows represent means and standard deviation for each sample characteristic in each experimental condition (M stands for mean value and *(SD)* for standard deviation value). The questions about prior experience with systems supporting complex decisions, in-person collaboration, remote collaboration, in-person games, remote games, and ICT devices usage were measured on the Likert-type scales (from 1-never to 5-everyday). The last column provides the results of statistical tests of difference between experimental conditions on respective dependent variables (chi-square tests or one-way between-subjects ANOVA tests were used accordingly to the dependent variable).

4.1 Participants

A total of 71 students volunteered to participate in the experiment in exchange for student activity credit points (30 Females, $M_{age} = 26.52$, $SD_{age} = 5.12$). Participants were recruited via an advertisement on a university recruitment system and social-media closed groups. Most participants ($n = 58$) were psychology students.

Participants were randomly assigned to one of three experimental conditions: single user (17 individuals), in-person collaboration (13 pairs), and remote collaboration (14 pairs). Since some reviews suggest that sex may influence team work [Bear and Woolley 2011], the collaborating pairs were of the same sex. Participants declared that their vision was normal or corrected to normal. The characteristics of the participants in each experimental condition are presented in Table 1. Power analysis was conducted with the use of *G*Power* software [El Maniani et al. 2016; Erdfelder et al. 1996]. The results showed that a sample of 75 participants would yield 0.95 power to detect effect sizes of 0.1 with the mixed-design **analysis of variance** (**ANOVA**) procedure.

4.2 Procedure

The flow of the experimental procedure is presented in Figure 2. Prior to testing, all participants signed an informed consent form. Participants were informed that they could resign from taking part in the experiment at any stage of the study. Then they filled out the pre-study online questionnaire. Next, the experimenter placed a mobile eye tracker on the participant's head and set three cameras (world camera and two eye cameras) adapting the device to the participants' eyes. Each person was informed how the eye tracker worked and was asked if cameras were occluding their vision. After a brief talk, during which the participants got accustomed to the eye tracker, they individually passed through a single marker calibration. The experimenter, at a distance of two meters and using a single circle marker, showed five calibration points around her head. Participants were asked to follow with their eyes the central point of the circle. After calibration, participants received instructions on how to maintain a stable position approximately one meter away from the large display and were instructed on how to proceed with the simulation task (Figure 2). Participants completed the task in an office environment either individually (single-user

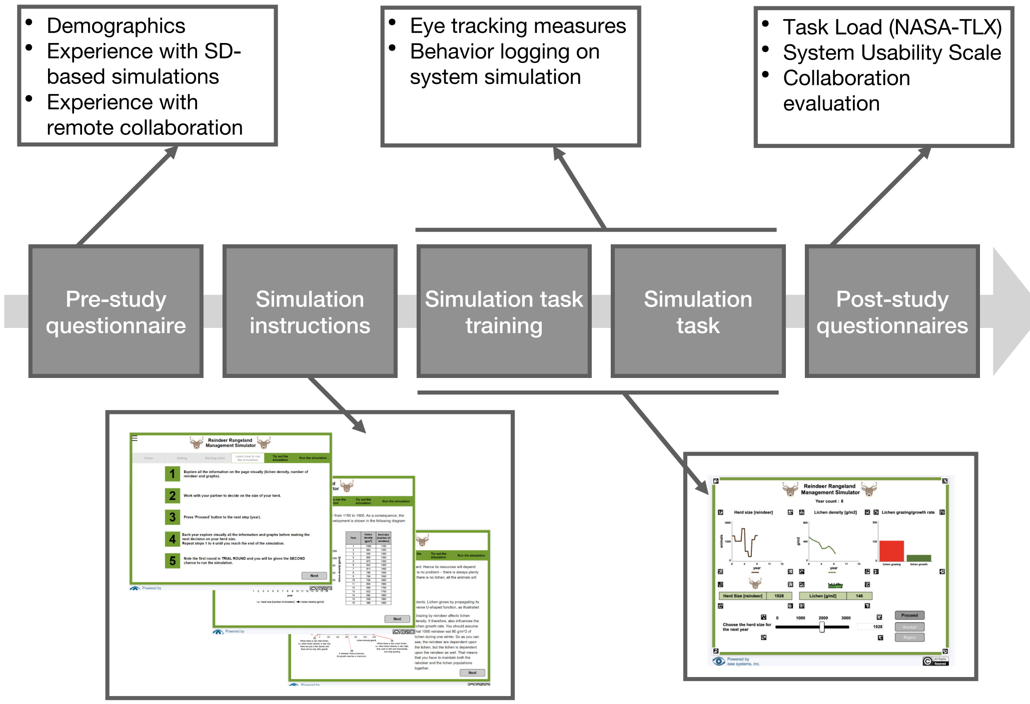


Fig. 2. Experimental procedure scheme. On the timeline (from left to right), we see steps followed by participants during the experiment. Above the timeline, the Figure presents the key measures used in subsequent steps of the experiment. Below the timeline, we present the exemplary user interface visible to the participants in the subsequent steps.

condition) or in pairs (in-person and remote collaboration conditions). Working in pairs, participants cooperated in one room on one interactive large display (in-person collaboration condition) or in two rooms on separate large displays (remote collaboration condition).

Participants could talk to each other and discuss their decisions during the entire task. In the remote condition, communication took place via Skype, where participants could talk and see each other in the window located in the upper right corner of the simulation user interface (Figure 1). All conversations were recorded using Pupil Labs Capture software.

After completion of the task, participants filled out post-study online questionnaires and were thanked for their participation. The entire procedure lasted about 30 minutes. The experimental procedure was approved by Ethical Committee No. 15/2020 of the first author's institution prior to the data collection.

4.3 Pre-Study Questionnaire

The pre-study questionnaire was completed by all participants prior to the simulation task. It consisted of 10 questions. The questionnaire was prepared with the use of the Qualtrix online questionnaire system. The first four questions were about basic demographic information (age, gender, handedness, sight problems). There were also six questions about participants' prior experience with systems supporting context decision: how often they use ICT devices (laptop, mobile phone, tablet, smartwatch, computer, smartphone), how often they collaborate in-person and remotely, and how often they play games in-person and remotely. Participants answered the six questions on a five-point Likert-type scale from 1 “never” to 5 “everyday”. Descriptive statistics of the pre-study questionnaire are presented in Table 1. Using one-way ANOVA, we tested differences among experimental conditions for ICT usage, prior experience, and average age. There were no significant differences between the experimental conditions. This suggests that participants' samples were similar in terms of the measured demographics across all three experimental conditions. The average pre-study questionnaire completion time was 4 minutes. All participants completed the questionnaires.

4.4 Experimental Task

For the experimental task, we used an SD-based simulation task developed by Moxnes [2004]. Participants played the role of sole owners of a reindeer herd. This task's structure is representative of a wide range of natural resource management tasks and provides ample benchmarking data from previous experiments [Derwisch et al. 2011; Moxnes 2004].

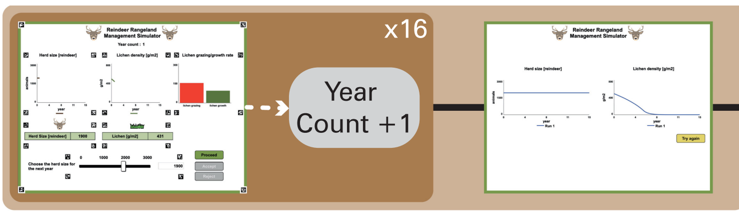
Time is a key dimension in natural resource management simulations, including understanding of historic data and planning for future decision-making. In the simulation task, participants were required to take over reindeer rangeland in an over-utilized state and decide on the herd size for each of 16 consecutive years (16 decisions). Their goal was to reach the maximum sustainable herd size within as few years as possible. Unlike climate change, in which most people have some knowledge and perhaps considerable interest, people usually do not have deep knowledge about reindeer rangeland management and thus focus more on the information provided in the experimental setting than on their own prior knowledge. Thus, prior to the simulation, participants were provided with detailed information about the rules of the rangeland simulation, i.e., the lichen grazing rate of the reindeer, the description of lichen growth rate, the historic record on reindeer herd size, and lichen density levels (Figure 2). After the detailed instructions, they started with two training runs of the simulation task prior to the main simulation task from which the data was analyzed (Figure 2).

Determination of the size of the reindeer herd is performed using a touch interface as shown in Figure 4(a). Participants make decisions for each year by checking the analytical information displayed on the user interface (herd size, lichen density, and lichen growth) and select the herd size for the following year using the slider located at the bottom of the user interface.

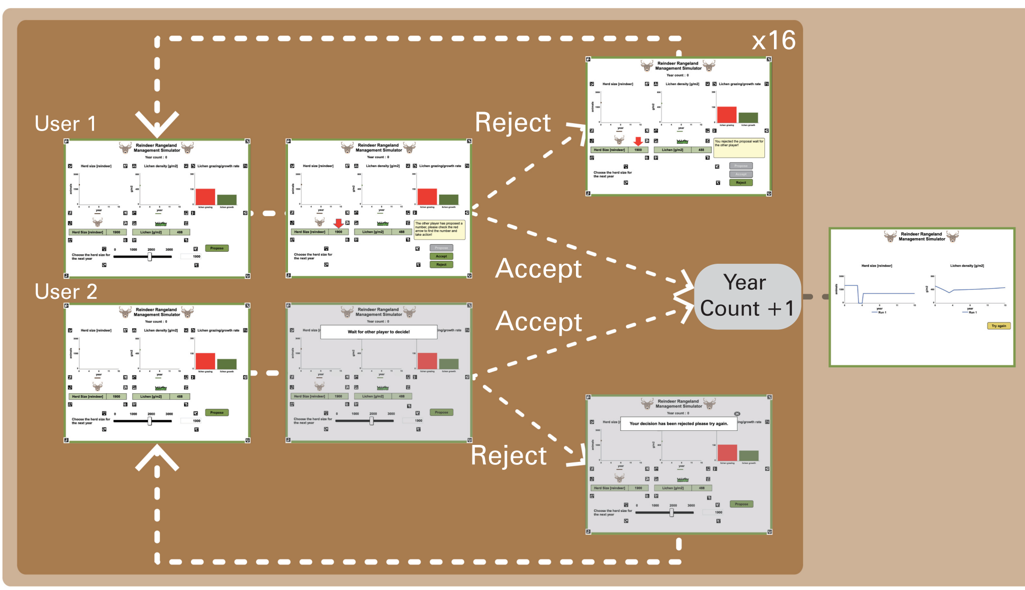
Performing an optimal simulation requires an understanding of the convex growth function of lichen density [Kopainsky and Sawicka 2011]. Thus, in a simulation achieving a maximum sustainable herd size, participants needed to first reduce the herd size dramatically so that the natural resource (lichen) could recover and reach maximum annual regeneration rates.

The simulation task workflow (Figure 3) was similar in all the experimental conditions, with small adjustments required by the collaboration conditions. In the single-user condition (Figure 3(a)), each participant decided on the reindeer herd size for each year. In the in-person and remote collaboration conditions, one of the two users sets the reindeer herd size for each year using the same interface as in the single-user condition (Figure 3(b)). However, both users needed to agree on the herd size decision. Therefore, when one participant in the collaboration pair proposed the size of the reindeer herd, the other was asked to accept or reject this proposal. If accepted, the task moved to the next year. If rejected, the decision on herd size had to be retaken until an agreement was achieved.

We built the lichen-reindeer simulation model and its interactive user interface using the Stella Architect software [isee systems inc. 2020]. This software supports web-based, remote collaboration, online data storage and analysis, and touch input. The software offers a drag-and-drop user interface builder. The simulation task was performed on a multitouch 55-inch 4K UHD Google



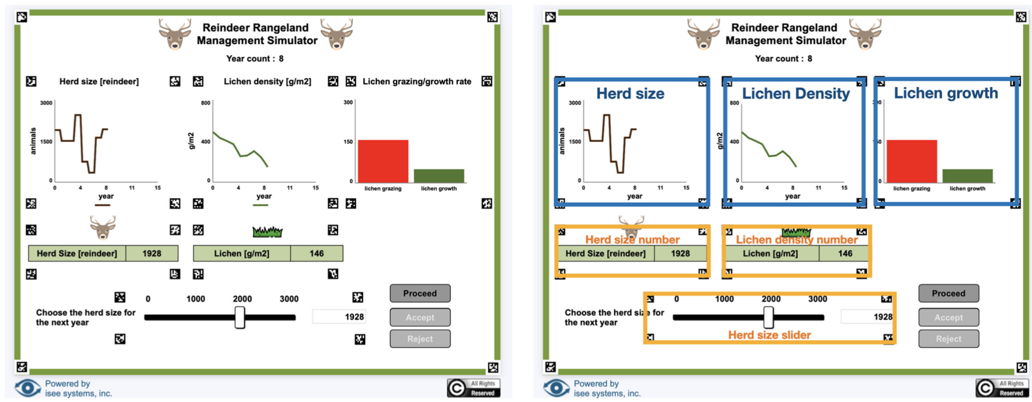
(a) Simulation task workflow used in a single-user experimental condition.



(b) Simulation task workflow used in remote and in-person collaboration experimental conditions.

Fig. 3. Simulation task workflow in all experimental conditions (single user, in-person collaboration, and remote collaboration). The user interface displays data such as herd size, lichen density, and lichen grazing rate. This data is presented in two ways: as a graph over time and as a numerical display. There are two input areas at the bottom of the user interface: one for users to enter their decisions on reindeer herd size and the other for users to interact with collaboration partner in the collaboration condition. The participants could enter the number of reindeer using either a slider or numeric input. Once they proposed their decision using the propose button, they were directed to a waiting page where they had to wait for the other collaboration partner to decide on the proposed number. The collaboration partner was notified that a number for the reindeer herd had been proposed. They could either accept or reject the proposed number using consensus buttons. If they were in agreement, they clicked the accept button, which advanced the simulation by one year. The reject button was used if the collaboration partner did not agree, whereby the partner who had proposed the number was notified and they had to agree upon a new number for the reindeer herd. Participants made 16 decisions in the simulation task as described above.

Jamboard large display with 1920 × 1080 resolution on the floor stand controlled by a standard laptop computer. The Google Jambord large display was used as a secondary display thus its native resolution was scaled to the laptop's maximum resolution. Participants interacted with the Google Jamboard using passive styluses (Figure 1). In the in-person condition, participants worked on the same large display, similar to a single user condition. In the remote condition, participants completed the task on two Google Jamboard large displays located in separate rooms.



(a) Simulation user interface

(b) Simulation user interface with marked AOIs

Fig. 4. Simulation task user interface. Figure 4(a) presents an exemplary screenshot with important information on the history and current status of the simulation (lichen density, reindeer size, and lichen growth graphs) and controls (buttons and slider) for current decision-making. The fiducial markers (squares with black and white patterns) were embedded in the interface to define AOIs for eye tracking data analysis. Figure 4(b) presents the same user interface with color-marked AOIs. The AOIs marked with a blue color relate to simulation control. AOIs marked with orange relate to decision-making (see Section 5). The AOIs marked in color on Figure 4(b) were not visible to the participants.

4.5 Eye Tracking Equipment

During the simulation task, participants' eye movements were recorded using mobile PupilLabs Core eye trackers (120*Hz* sampling rate). This is a non-invasive device that allows free gaze communication, does not obstruct the vision and does not cause discomfort to the participants during the cooperation due to its design [Schindler and Lilienthal 2017]. PupilLabs mobile eye trackers have been used before in dual eye tracking studies [Shvarts et al. 2018]. Recorded gaze data quality was above 0.80 confidence level (1.00 is the maximum). Gaze data were pre-processed and exported using the PupilLabs Player software. The fixation detection threshold was set to 80*ms*. The Apriltag visual fiducial markers [Olson 2011] were used to define AOIs on the key elements of the simulation user interface (Figure 4(b)). These are necessary to link eye-movement measures to selected parts of the interface [Hessels et al. 2016]. The user interface contained two types of visual information (see Figure 4(b)) defining two types of AOIs. The first type was related to simulation control. These were the three graphs for herd size, lichen density, and lichen growth, presenting momentary outcomes of the simulation (Figure 4(b), AOIs marked with blue color). The second type was related to decision-making processes. These were the three AOIs presenting actual herd size, lichen density, and the herd size choice slider (Figure 4(b), AOIs marked with orange color). We averaged eye movement-based dependent variables across these types of AOIs. Average fixation duration and dwell time (total viewing time) for simulation control AOIs and decision-making types of AOIs were considered as the main dependent measures of visual attention distribution during the simulation task.

4.6 Post-Study Questionnaires

After the simulation task, participants completed post-study questionnaires prepared with the use of the Qualtrics online questionnaire system. We used the following scales to assess their subjective perception of three aspects of collaboration during the experimental task (task load, usability, and collaboration quality). These questionnaires have frequently been used in previous HCI research [Duchowski et al. 2020].

(1) The **NASA Task Load Index** (**NASA-TLX**) is a six-item scale that measures a participant's reported level of task workload [Hart and Staveland 1988]. This is a subjective, multidimensional assessment tool rating perceived workload in assessing a task or a system. The range of responses was from 0 (very low) to 21 (very high). The sample item: *How hard did you have to work to accomplish your level of performance?* The reliability of the scale was satisfactory (Cronbach's $\alpha = 0.75$) [Tavakol and Dennick 2011].
(2) The **System Usability Scale** (**SUS**) is a 10-item attitude Likert scale with responses ranging from 1 (strongly disagree), to 5 (strongly agree), giving an overview of subjective assessments of usability [Brooke 1996, 2013]. The sample item: *I would imagine that most people would learn to use this system very quickly.* Following Hart and Staveland [1988] we subtracted 1 from each value of the item and multiplied it by 5. The overall score was the mean of all items. The reliability of the scale equaled 0.79 Cronbach's α.
(3) The **Collaboration Assessment Scale** (**CAS**) measures self-reported levels of collaboration quality from 1 (very bad) to 5 (very good) and the degree of being in control or being overwhelmed from 1 (not at all) to 5 (a lot). The CAS was completed only by participants in the collaboration conditions. The sample item: *Did you feel overloaded by the collaborator?* The reliability of the scale equaled 0.72 Cronbach's α. The average post-study questionnaires completion time was 6 minutes. All participants completed the questionnaires.

5 RESULTS

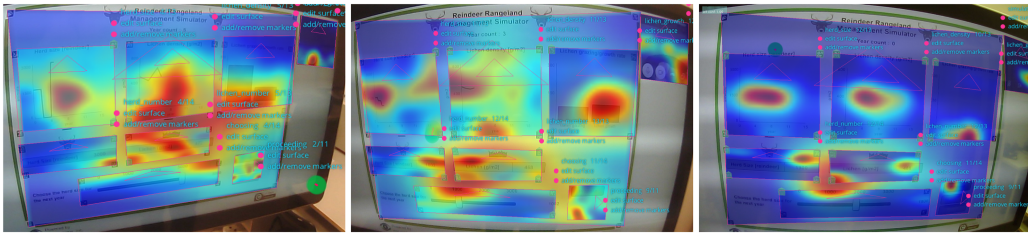
The Results section is divided into three parts: (a) depth of information processing (hypothesis 1), visual attention distribution (hypothesis 2), (b) simulation outcome analyses (hypothesis 3), and (c) subjective assessments of task workload, system usability, and collaboration quality (hypothesis 4). All the statistical analyses were performed in the R language for statistical computing [R Core Team 2020]. R language for statistical analysis has become an "industry standard" in the realm of data science [Weston and Yee 2017], and it covers all statistical procedures needed for hypothesis testing in the present study. To test the hypotheses, we decided to use ANOVA as a statistical test because it allows for obtaining statistical significance for means comparison between three experimental conditions in the present study, and also because it is a well-established standard statistical procedure for such comparisons [Field et al. 2012]. Any statistically significant effects obtained in the analyses were followed by *post-hoc* comparisons with HSD Tukey correction for multiple comparisons. The raw data used in the described analyses are available in the OSF repository (https://osf.io/bymak/).

Prior to the hypotheses testing, we examined differences between experimental conditions for demographic variables listed in Table 1. All comparisons were statistically insignificant ($p > 0.05$), therefore we did not include these demographic variables as covariates while testing the hypotheses (see Table 1 for the results of the statistical tests results of comparisons between experimental conditions on each demographic variable).

5.1 Distribution of Attention and Visual Information Processing

According to hypothesis 1, we expected that collaboration would increase the focus of attention to key information presented on the user interface. Understanding the mechanics of the simulation model and finding an optimal outcome require focusing attention on the simulation control AOIs.

In order to test our hypotheses related to differences in attention distribution, we conducted a series of mixed-design ANOVAs with experimental condition as the between-subjects



(a) In-person collaboration (b) Remote collaboration (c) Single user

Fig. 5. Exemplary heatmaps presenting attention distribution in three experimental conditions: in-person collaboration, remote collaboration, and single user. The heatmaps were prepared and normalized in Pupil Player software [Kassner et al. 2014]. Heatmaps provide an overview on the visual attention distribution over the screen and separate different levels of observation intensity. The longer the gaze on a certain region, the warmer the color is used to represent it [Špakov and Miniotas 2007]. Color codes represent the amount of time spent gazing on each AOI. *Note*: The heatmaps are projected onto images from the eye tracker's front camera when looking at the SD-based simulation in different experimental conditions.

independent variable and AOIs as the within-subjects independent variable. The analyses were conducted for average fixation duration (cognitive processing measure) and dwell time (attention distribution measure) as dependent variables. Details of the analyses are presented in the following subsections.

5.1.1 Depth of Information Processing: Simulation Control vs. Decision-Making AOIs. We tested the first hypothesis that decision-making evokes deeper cognitive processing than simulation control. We ran a two-way (2×3) mixed design ANOVA, with AOI type as the within-subjects dependent variable and experimental condition as the between-subjects independent variable, and average fixation duration as the dependent variable. In line with the hypothesis, the analysis revealed a significant main effect of an AOI type, ($F(1, 34) = 5.96,\ p = 0.020,\ \eta^2 = 0.037$). Average fixation duration was significantly longer while looking at decision-making AOI type ($M = 149ms,\ SE = 51.9$) than simulation control AOIs type ($M = 143ms,\ SE = 51.9$).

5.1.2 Distribution of Attention over Simulation Control AOIs. To verify if in-person and remote collaboration, in comparison to single user condition, facilitated a more thorough exploration of AOIs containing important information for controlling the simulation, we examined attention distribution over the three AOIs related to simulation control: the graphs presenting herd size, lichen density, and lichen growth (Figure 4(b)). Figure 5 presents exemplary differences in attention distribution over AOIs in the experimental conditions.

Two-way mixed-design (3×3) ANOVA was carried out with AOIs related to simulation control graphs (lichen density vs. lichen growth vs. herd size) as within-subjects independent variable and experimental condition as between-subjects the independent variable, and dwell time (attention distribution measure) as the dependent variable.

The ANOVA for attention distribution (dwell time) on AOIs related to the simulation control graphs revealed a main effect of experimental condition ($F(2, 32) = 5.80,\ p = 0.007,\ \eta^2 = 0.136$). As expected, participants collaborating in-person spent significantly more time examining the simulation control graphs ($M = 36668ms,\ SE = 4576$) than participants in the single user condition ($M = 16425ms,\ SE = 4246$). Also participants collaborating remotely tended to spend significantly ($t(32) = 2.43,\ p = 0.05$) more time examining the simulation control graphs ($M = 3592ms,\ SE = 4348$) than participants in the single user condition. The difference between remote and in-person collaboration was not significant.

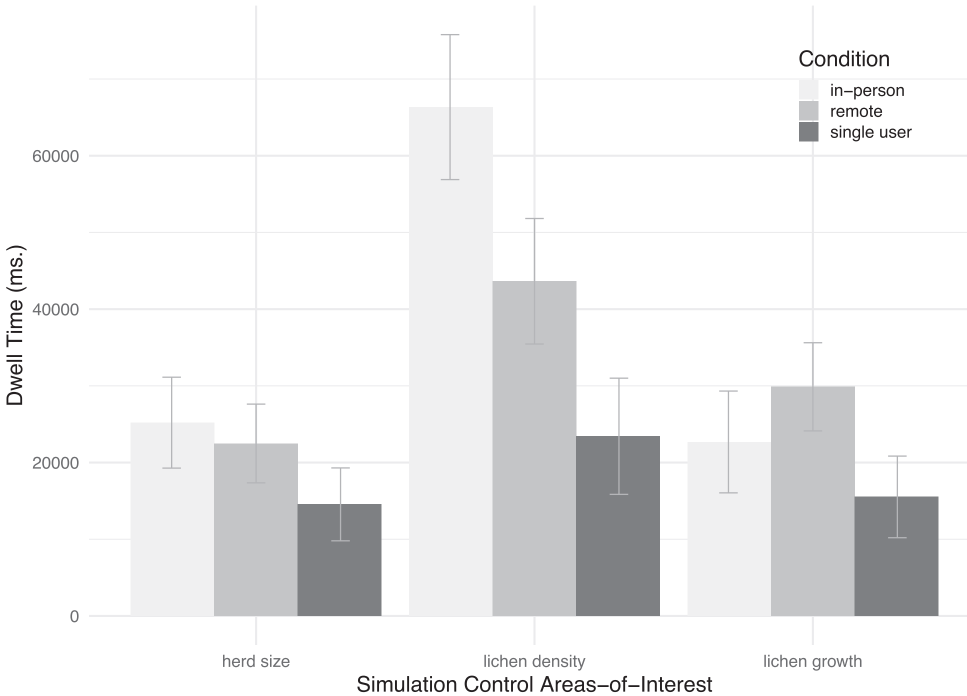


Fig. 6. Differences in distribution of attention (dwell time) over AOIs (herd size, lichen density, and lichen growth graphs) related to control of the simulation task between experimental conditions (in-person collaboration, remote collaboration, and single user condition). *Note*: the height of bars represents estimated means of dwell time and error bars represent ± 1 *SE*.

There was also a significant main effect of AOI, ($F(2, 64) = 13.73,\ p < 0.001,\ \eta^2 = 0.195$). In line with expectations, the longest dwell time was recorded for the lichen density graph ($M = 43060ms,\ SE = 3836$) when compared to the attention paid to the lichen growth graph ($M = 21287ms,\ SE = 3836;\ t(64) = 4.331,\ p < 0.001$) and the herd size graph ($M = 19338ms,\ SE = 3836;\ t(64) = 4.719,\ p < 0.001$).

The main effects were quantified by a significant interaction of AOI and condition ($F(4, 64) = 2.68,\ p = 0.039,\ \eta^2 = 0.086$), see Figure 6 for a comparison of means in dwell time on different AOIs between experimental conditions as well as Figure 5 for heatmaps of eye movements presenting the attention distribution over different AOIs in different experimental conditions. *Post-hoc* comparisons showed that in-person collaboration triggered significantly ($t(91.7) = 4.47,\ p < 0.001$) longer dwell time on the lichen density graph ($M = 64933ms,\ SE = 7254$) in comparison to the single user condition ($M = 22019ms,\ SE = 6196$). The difference between in-person and remote collaboration conditions was marginally significant ($t(91.7) = 2.29,\ p = 0.062$), and remote collaboration triggered marginally longer dwell times on the lichen density graph ($M = 42229ms,\ SE = 6532;\ t(91.7) = 2.27,\ p = 0.063$) than for the single user condition.

Next, we calculated *post-hoc* comparisons for the differences in attention distribution among simulation control graphs. In the single user condition, no significant differences in attention distribution on different elements of simulation control were found, suggesting more equal attention distribution in comparison to both collaboration conditions (Figure 6). In line with the second hypothesis, the collaboration conditions induced more attention to the lichen density graph, which is the most important part of the simulation control visualizations. In the in-person condition,

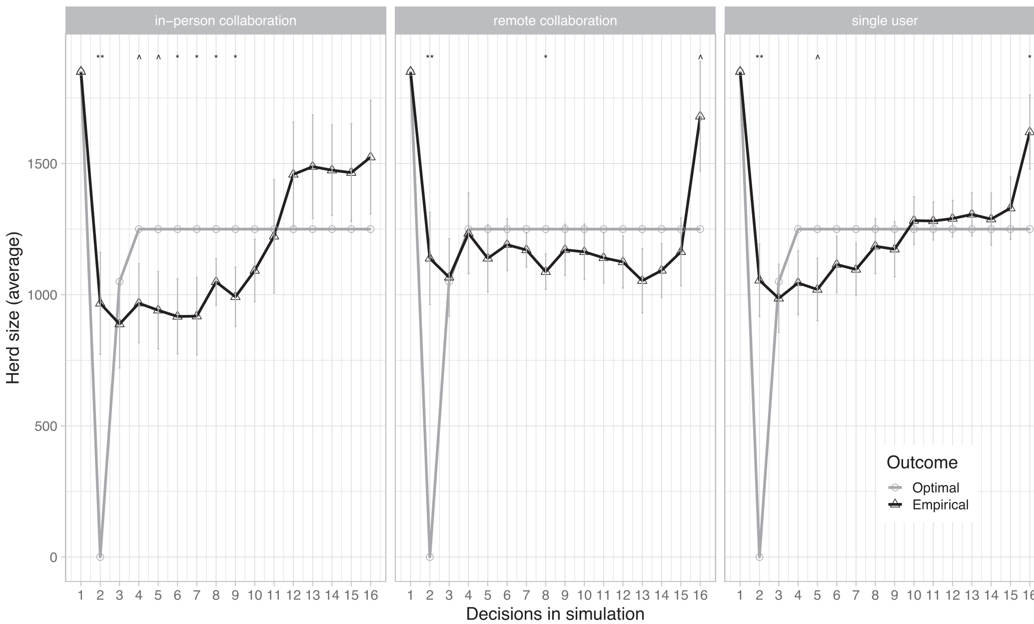


Fig. 7. Participants' decisions of herd size during the simulation task (black line) compared to an optimal outcome (gray line) for each experimental condition. The gray lines represent the optimal herd size derived from the model. Each graph represents the results of a separate ANOVA comparing herd size obtained in certain experimental condition compared to the optimal herd size outcome. *Note*: Error bars represent ±1 *SE*. Significant differences between empirical and optimal outcomes are annotated above the lines ($\hat{}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$).

participants focused their attention significantly more on the lichen density graph than on the herd size graph ($M = 23793ms$, $SE = 7254$; $t(64) = 4.22$, $p < 0.001$) and on the lichen growth graph ($M = 21278ms$, $SE = 7254$; $t(64) = 4.48$, $p < 0.001$). Similarly, in remote collaboration, participants spent significantly more time inspecting the lichen density graph than the herd size graph ($t(64) = 2.51$, $p = 0.039$), but the difference in dwell time between lichen growth graph and herd size graph was not significant.

5.2 Simulation Outcome

The third hypothesis predicted that in-person collaboration would promote a strategy of decision-making similar to the optimal strategy derived from the model. The optimal solution is not intuitive and requires understanding by users that, at the outset of the task, the rangeland is overutilized and that, for lichen density to recover, the number of reindeer (herd size) needs to be reduced drastically for a short period of time (see the gray line in Figure 7). The simulation outcome is indicated by lichen density in each simulation step, after each decision taken by participants regarding herd size. The optimal lichen density outcome is visible in the gray line on Figure 8. In order to test the third hypothesis regarding decisions taken by participants in each experimental condition, we conducted three mixed-design (2×16) ANOVAs with herd size as the dependent variable, separately for each experimental condition. A total of 16 decisions (steps of simulations) in the simulation task were entered as the within-subjects factor and the type of outcome (empirical vs. optimal) as the between-subjects factor.

In general, the empirical herd size did not differ from the optimal outcome ($M_{herd} = 1197$), neither in in-person collaboration ($F(1, 12) < 1$, $M_{herd} = 1201$, $SE_{herd} = 68.20$), nor in the remote

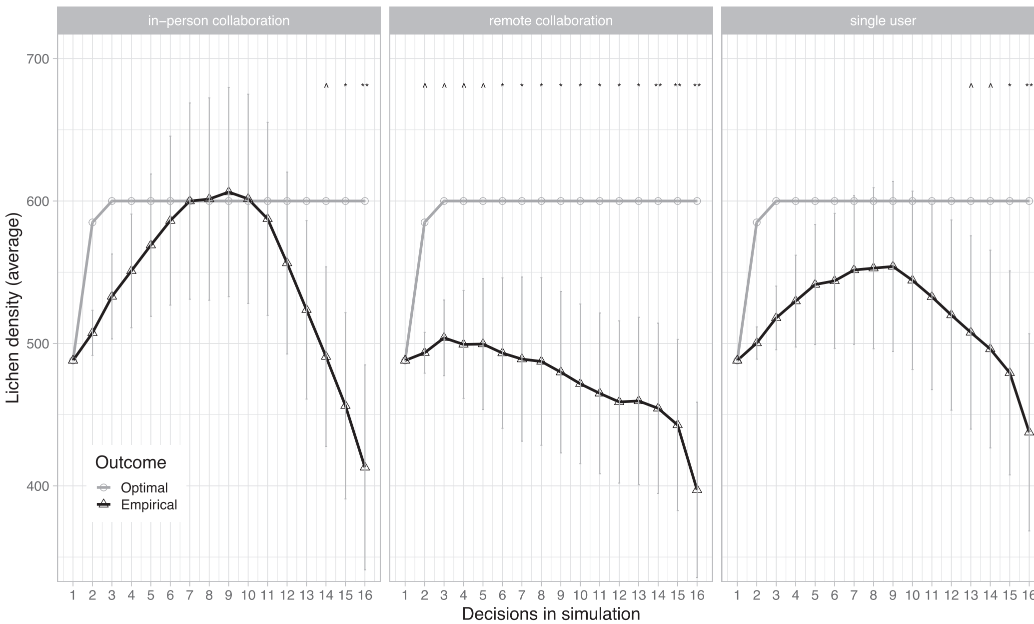


Fig. 8. Outcome of the simulation task represented in changes of lichen density (black lines) compared to an optimal outcome (gray line) for each experimental condition. The gray lines represent the optimal lichen density derived from the model. Each graph represents the results of a separate ANOVA comparing lichen density in certain experimental conditions compared to the optimal lichen density. *Note*: Error bars represent ±1 *SE*. Significant differences between the empirical and optimal outcome are annotated above the lines with ($\hat{}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$).

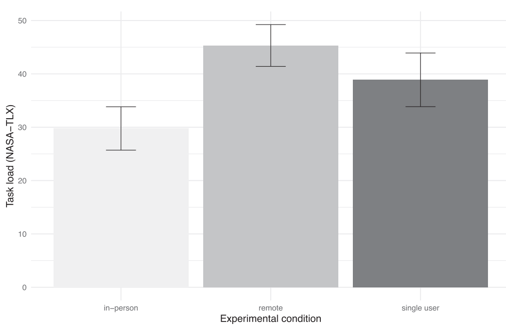
collaboration ($F(1, 13) < 1$), ($M = 1216$ $SE = 34.40$) and single-user condition ($F(1, 16) < 1$, $M_{herd} = 1245$, $SE_{herd} = 64.50$).

The interaction effects between outcome type and consecutive decisions were statistically significant in all three conditions: for in-person collaboration ($F(15, 180) = 5.27$, $p < 0.001$, $\eta^2 = 0.152$), remote collaboration condition ($F(15, 195) = 7.80$, $p < 0.001$, $\eta^2 = 0.337$), and single user condition ($F(3.59, 57.5) = 13.08$, $p < 0.001$, $\eta^2 = 0.265$). Each interaction effect was followed by pairwise comparisons of empirical and optimal herd sizes in each year (step of simulation) with HSD Tukey correction. For detailed comparisons see the interaction effects on Figure 7.

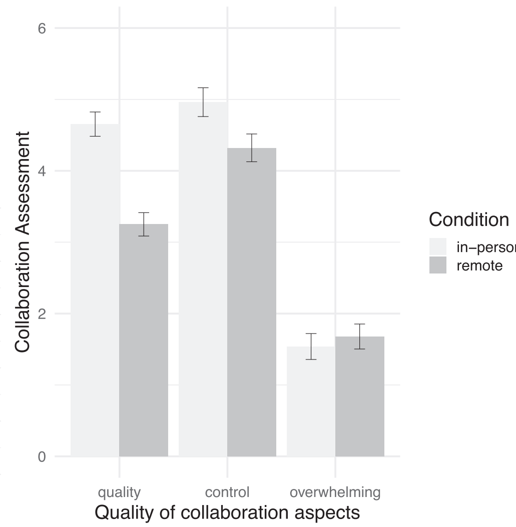
Next, we compared the average empirical lichen density (obtained as a result of participants' decisions) to the optimal outcome derived from the model. For each experimental condition, we conducted a mixed-design (2 × 16) ANOVA for lichen density as the dependent variable. A total of 16 decisions were entered as a within-subjects factor and the type of outcome (empirical vs. optimal) as a between-subjects factor.

The analyses revealed that the empirical outcome did not differ significantly from the optimal one, neither in the single-user condition ($F(1, 16) < 2.39$, $p = 0.141$, $M_{lichen} = 518.57$, $SE_{lichen} = 18.55$) nor in the in-person collaboration condition ($F(1, 12) < 1$, $M_{lichen} = 541.94$, $SE_{lichen} = 20.47$). However, the difference was statistically significant in remote collaboration condition ($F(1, 13) = 6.92$, $p = 0.021$, $\eta^2 = 0.265$), where participants' decisions led to significantly lower lichen density ($M = 473.93$ $SE = 17.25$) than the optimum ($M = 592.06$).

The interaction effect between outcome type and consecutive decisions was significant for in-person collaboration ($F(15, 180) = 3.88$, $p < 0.001$, $\eta^2 = 0.071$), remote collaboration ($F(15, 195) = 3.50$, $p < 0.001$, $\eta^2 = 0.042$), but not for the single user condition ($F < 1$). Both of the significant



(a) Task workload



(b) Collaboration assessment

Fig. 9. Subjective assessment of (a) task workload and (b) collaboration quality measured with three scales (quality, being in control, and feeling overwhelmed). *Note*: The error bars represent ± 1 *SE*.

interaction effects were followed by pairwise comparisons of empirical to the optimal outcome in each step of the simulation with HSD Tukey correction for multiple comparisons. Figure 8 presents estimated means with the significance of each comparison.

5.3 Subjective Assessments: Task Workload, Usability, and Collaboration Quality

In this section, we present subjective assessments of participants' reactions to the simulation task related to the fourth hypothesis. In a series of between-subjects one-way ANOVAs, we compared differences between experimental conditions in self-reported task workload, collaboration quality, and system usability. Collaboration quality was examined only in the collaboration conditions (in-person and remote).

System Usability Evaluation. The analysis for system usability did not indicate significant differences among the three experimental conditions, ($F(2, 68) = 1.48,\ p = 0.20$). System usability, on average, was similar in all experimental conditions (in-person: $M = 80.6,\ SE = 3.56$, remote: $M = 76.2,\ SE = 3.43$, and single user: $M = 70.9,\ SE = 4.40$).

Task Workload. We start by examining differences in participants' perceived task workload measured with NASA-TLX questionnaire. A one-way between-subjects ANOVA revealed a significant effect of experimental condition ($F(2, 68) = 3.77,\ p = 0.03,\ \eta^2 = 0.1$), see Figure 9(a). The *post-hoc* comparisons suggest a higher reported task load among participants working in the remote condition ($M = 45.1,\ SE = 3.92$) compared to the in-person condition ($M = 29.7,\ SE = 4.07$) and the single user condition ($M = 38.8,\ SE = 5.03$). The difference between single user and in-person condition in perceived task workload was not significant ($t(154) = 1.41,\ p = 0.34$).

Collaboration Quality. The collaboration assessment was conducted for remote and in-person collaboration experimental conditions. To check the differences between these conditions, three one-way ANOVAs were conducted separately for three dependent variables: quality of collaboration, feeling of being in control, and feeling of being overwhelmed.

The ANOVA for the quality of collaboration revealed that participants in the in-person condition rated the quality of pair-work significantly higher ($M = 4.64,\ SE = 0.18$) than participants in the remote condition ($M = 3.24,\ SE = 0.18$), $F(1, 52) = 21.4,\ p < 0.001,\ \eta^2 = 0.04$, (Figure 9(b)).

The ANOVA for the feeling of being in control suggested that participants felt significantly more in control while working in-person ($M = 4.95,\ SE = 0.18$) compared to working in the remote condition ($M = 4.31,\ SE = 0.18$), ($F(1, 52) = 5.22,\ p = 0.03,\ \eta^2 = 0.09$). Finally, there were no significant differences between the in-person and remote conditions for a feeling of being overwhelmed, ($t(154) = 0.55,\ p = 0.586$). In general, participants reported low levels of feeling overwhelmed during in-person ($M = 1.53,\ SE = 0.18$) and remote collaboration ($M = 1.67,\ SE = 0.18$).

6 DISCUSSION

In this article, we presented the results of an eye tracking experiment revealing the attentional mechanism underlying effectiveness and quality of in-person and remote collaboration during SD-based complex decision-making task solving. The experiment participants, all new to reindeer rangeland management, were given instructions on how to recover and grow lichen density and run the simulation task. During task performance, their eye movements were recorded. We hypothesized that collaboration would help to focus their visual attention on crucial information related to simulation control (hypothesis 1 and hypothesis 2). Consequently, we expected that collaboration would facilitate a more optimal outcome of the simulation task (hypothesis 3). The effectiveness of collaboration was traced via participants' decision outcomes and the self-reported quality of collaboration (hypothesis 4). In general, the triangulated results of eye movements, behavior, and self-report analyses supported our hypotheses.

6.1 In-Person Collaboration Fosters Attention to Key Task Information

To test the hypotheses about attentional bias to key information and provide insight into decision-making, we analyzed visual attention distribution and depth of processing over simulation control and decision-making AOIs. Participants in all three experimental conditions had significantly longer average fixation duration on decision-making AOIs than on simulation control AOIs. Referring to Just's and Carpenter's "*eye-mind assumption*" [1976] this may suggest that decision-making was more cognitively demanding than reading simulation control graphs. In line with the second hypothesis, when analyzing simulation control graphs, participants in both collaboration conditions focused longer on the lichen density graph than on the other two graphs (Figure 6). That was especially salient in the in-person collaboration condition, where participants focused more on the lichen density graph than in the other two experimental conditions. The lichen density graph was a critical piece of information for understanding the principles underlying the simulation task. In the single-user condition, participants' attention was evenly distributed across the simulation control graphs.

6.2 Closer to Optimal Decision-Making Strategy in In-Person Collaboration and Single-User Conditions

Previous System Dynamics research [Derwisch et al. 2011; Moxnes 2004] has shown that people have persistent difficulties with performing the reindeer rangeland management task. The task is formulated such that the sole owner of a reindeer herd (decision-maker) takes over the associated rangeland (pasture covered by lichen) in an overgrazed situation. The optimal solution to the task requires understanding that for lichen density to recover and grow to its maximum sustainable yield (which, in turn, allows for the maximum sustainable reindeer herd size), one needs to substantially reduce the number of reindeer in the first years and then gradually adjust the herd size up to the maximum sustainable number.

Overall, participants' decisions on herd size led to the outcome (lichen density) that was similar to the optimal task solution. Notably, participants in the in-person and single user conditions seemed to follow a slightly different strategy than participants in the remote condition. In the first years, in-person and single user participants reduced the reindeer herd size to below 1000 animals. This allowed lichen density to recover and grow closer to maximum sustainable levels. With this strategy, they followed the principles underlying an optimal solution more closely than participants in the remote condition who, at the beginning of the simulation, did not reduce herd size enough for lichen density to recover. In the last years, participants seemed to show some end-of-game behavior where they increased herd size again well above the maximum sustainable levels. This, in turn, led to the depletion of lichen density toward the end of the simulation. Similar performance improvements have been observed in studies with targeted instructional support for single users [Kopainsky and Alessi 2015; Kopainsky and Sawicka 2011], and we thus generalize that similar performance patterns can also be expected for other SD-based complex decision-making tasks.

6.3 Higher Self-Reported Collaboration Quality and Lower Task Load in the In-Person Condition

Even though all participants similarly rated system usability, participants in the in-person collaboration and single user condition reported less workload than people working remotely. Participants collaborating in person also reported a higher quality of collaboration and felt more in control of the simulation than participants during remote collaboration. Based on the results, we postulate that a better understanding of the task, reflected in attention bias toward the most important information and optimal outcome of own decisions, resulted in higher satisfaction from collaboration. However, future studies may want to add explicit measures of the model understanding level achieved during interactive simulation tasks [Happach and Schoenberg 2017].

6.4 The Role of In-Person Contact

Taken together, our findings lead us to conclude that collaboration in remote and in-person conditions facilitated the focus of users' attention on the information most important for understanding the simulation more than in the single-user condition. The findings align with the definition of collaboration as a construction of shared understanding through interaction with others, and commitment to problem-solving [Roschelle and Teasley 1995]. Importantly, an in-person collaboration facilitated even stronger attention focus on key information than remote collaboration.

Also, users' decisions and their outcomes were most similar to the optimal when participants collaborated in person. The in-person collaboration tested in our research thus seems to have improved understanding when compared to alternatives (without any instructional support) for single users [Moxnes 1998b, 2004]. The in-person collaboration yielded a lower task workload than remote collaboration. Presumably, this was why our participants valued in-person collaboration more than remote.

Remote collaboration is sometimes evaluated as more confusing, less satisfactory [Thompson and Coovert 2003], and less productive [Straus and McGrath 1994], which may result in lower participation satisfaction [Lipponen et al. 2003] than in-person collaboration. Limited access to non-verbal signals of communication in remote collaboration, such as gaze communication [Pfeiffer et al. 2013], or mimicking others' non-verbal behavior and postures [Shockley et al. 2003], may be responsible for the level of shared understanding and commitment in task-solving. Remote collaboration settings seem to lack several coordination behaviors between partners exhibited by, e.g., drawing both partners' attention to the same visual elements by pointing, gesturing, or gazing [Clark and Krych 2004]. Therefore, examining solutions enabling non-verbal communication in an online workspace is needed. For instance, Ashdown and Robinson [2005] have created a system called the Escritoire, allowing participants to gesture to each other in ample visual space to enhance their remote communication. To enhance the effectiveness of remote collaboration, interface designers may also guide users' attention to the most important elements of the collaboration interface using subtle graphical gaze cues (color, blinking) [Frischen et al. 2007; Lu et al. 2012] thereby activating the bottom-up mechanisms of visual attention [Posner 2004], and as a result, curtail haphazard attention distribution. Future work may focus on testing practical guidelines for designers of collaborative interfaces.

6.5 Conclusion and Outlook

Our research offers quantitative results demonstrating that collaboration may lead to a more successful SD-based simulation strategy, especially during in-person collaboration. We investigated the process of achieving an improved understanding of the simulation by registering visual attention distribution. The results of our research demonstrated that during in-person and remote collaboration, the visual attention of collaboration partners focused on key elements of a simulation task more than in the single user condition and that in-person collaboration yielded less cognitive workload and was perceived as of higher quality than remote collaboration.

The main contribution of this study is understanding of the attentional mechanism underlying the effectiveness of in-person collaboration on a large interactive display. In-person collaboration helps users to focus their visual attention on the most important information necessary for optimal decisions. Understanding this attentional mechanism may serve designers of attentive user interfaces [Vertegaal 2002, 2003] for remote collaboration interactive systems.

That is, based on our findings and using eye tracking technology, future remote collaboration systems may be able to detect and guide remote user's visual attention to decision-related areas. By using visual or even eyes-free, vibrotactile notifications to present attention cues to the remote user, improved effectiveness and efficiency of distributed decision-making can be expected. Inspired by eyes-free, vibrotactile notification principles, such as OmniVib [Alvina et al. 2015], ultra-low-cost prototypes, such as DragTapVib [Fang et al. 2022], can reliably provide dragging, tapping, and vibrating sensations to their users.

Complex Dynamic Systems are characterized by three factors: feedback processes, nonlinearities, and accumulations [Sterman 2000]. Most people, including experts, have difficulty in understanding and managing these factors [Brehmer 1992; Funke 1991; Jensen 2005; Moxnes 2004; Sterman 1989; Sterman and Sweeney 2007]. One of the main purposes of System Dynamics is to improve the design of computer simulations, as tools, to improve decision-making, problem-solving, and management. Existing SD research has mainly focused on task performance and less on understanding and its determinants [Kopainsky and Alessi 2015].

In the present study, we examined collaboration on large interactive displays. That work may be extended to very popular small screens, e.g., tablets or smartphones, where workspace awareness over small screen may be harder to achieve. Recent research prototypes offering dynamic bindings among multiple tablets or smartphones, e.g., RAMPARTS project [Wozniak et al. 2016], have shown how workspace awareness benefits from such bindings. Moreover, studies of visual application sharing mechanisms among large interactive displays and small screens, e.g., CamCutter project [Hagiwara et al. 2019], have demonstrated effective ways to integrate large and personal displays into a shared workspace. Employing application sharing principles like CamCutter [Hagiwara et al. 2019], visually pulling down results displayed on an interactive surface to participants personal devices, such as phones and tablets, can be a way to make insights outlast the process itself. The present work may also be extended by adding additional technologies which may increase workspace awareness, like haptic or tangible input devices [Fjeld et al. 2007; Patten et al. 2001]. Another direction worth exploring is the role of awareness in collaborative workspaces combining physical and VR environments [Kudo et al. 2021].

To extend the understanding of visual attention processes as a mechanism underlying effective and satisfactory collaboration supported by computer systems, e.g., SD-based simulations, future studies may aim at evaluating the role of joint attention in complex task understanding and optimal outcome achievement. Joint attention can be detected by tracking the gaze of interaction partners as they focus their attention on the same objects or events within a specific time frame [Andrist et al. 2015; Schneider and Pea 2013]. Following previous studies [Stahl et al. 2014; Webb et al. 2002], one may claim that joint attention may substantially improve the effectiveness and satisfaction of collaboration. In the future, advancing the understanding of visual attention distribution may play an important role for System Dynamics and eye tracking research. Broadening our knowledge base about cognitive processes in using various interfaces is needed. Therefore, appropriate visual and analytical methods to explore and analyze time-oriented data [Aigner et al. 2011] and eye-movement analysis are needed.

Future studies may focus on corroborating the current results with different collaborative tasks, e.g., joint text writing, mathematical and algorithmic problems solving, or software coding. Future studies may also address other collaboration settings, e.g., desktop computers with relatively small screens, as well as collaboration using mobile devices.

Our findings provide insights for designing collaborative ICT-based systems, e.g., computer-supported collaboration or remote teamwork, which may lead to novel technological solutions. Our findings can scale to collaborative problem-solving over large displays in related domains, such as natural resource planning, smart grid design, and logistics simulation. More generally, our results may also be applicable to collaborative learning and social discussion, and can help to inform the future design of collaborative technologies for groups, organizations, communities, and networks.

7 AUTHORS' STATEMENT REGARDING PRIOR WORK

This article entirely presents original work that has not been, entirely or in parts, published elsewhere.

ACKNOWLEDGMENTS

We are thankful to those who supported this research: Anna Redeł (data collection), Anastasiia Timshina (data collection)), Barbara Stuckey (editing), Philippa Beckman (proofreading), Maria Pamela David-Dobay (proofreading), and Aditya Giridhar (video).

REFERENCES

Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. 2011. *Visualization of Time-Oriented Data* (1st ed.). Springer Publishing Company, Incorporated, London Dordrecht Heidelberg New York, NY. DOI: https://doi.org/10.1007/978-0-85729-079-3

Eduard M. Albay. 2019. Analyzing the effects of the problem solving approach to the performance and attitude of first year university students. *Social Sciences & Humanities Open* 1, 1 (2019), 100006. DOI: https://doi.org/10.1016/j.ssaho.2019.100006

Jessalyn Alvina, Shengdong Zhao, Simon T. Perrault, Maryam Azh, Thijs Roumen, and Morten Fjeld. 2015. OmniVib: Towards cross-body spatiotemporal vibrotactile notifications for mobile phones. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2487–2496. DOI: https://doi.org/10.1145/2702123.2702341

Sean Andrist, Wesley Collier, Michael Gleicher, Bilge Mutlu, and David Shaffer. 2015. Look together: Analyzing gaze coordination with epistemic network analysis. *Frontiers in Psychology* 6 (2015), 1016. DOI: https://doi.org/10.3389/fpsyg.2015.01016

Ernesto Arias, Hal Eden, Gerhard Fischer, Andrew Gorman, and Eric Scharff. 2000. Transcending the individual human mind–creating shared understanding through collaborative design. *ACM Transactions on Computer-Human Interaction* 7, 1 (March 2000), 84–113. DOI: https://doi.org/10.1145/344949.345015

Mark Ashdown and Peter Robinson. 2005. Remote collaboration on desk-sized displays. *Computer Animation and Virtual Worlds* 16, 1 (2005), 41–51. DOI: https://doi.org/10.1002/cav.55

Brigid Barron. 2003. When smart groups fail. *Journal of the Learning Sciences* 12, 3 (2003), 307–359. DOI: https://doi.org/10.1207/S15327809JLS1203_1

Julia B. Bear and Anita Williams Woolley. 2011. The role of gender in team collaboration and performance. *Interdisciplinary Science Reviews* 36, 2 (2011), 146–153. DOI: https://doi.org/10.1179/030801811X13013181961473

Roman Bednarik, Andrey Shipilov, and Sami Pietinen. 2011. Bidirectional gaze in remote computer mediated collaboration: Setup and initial results from pair-programming. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Vertegaal Work*. ACM, New York, NY, 597–600. DOI: https://doi.org/10.1145/1958824.1958923

Eric Bentzen, John Christiansen, and Claus Varnes. 2011. What attracts decision makers' attention?: Managerial allocation of time at product development portfolio meetings. *Management Decision* 49, 3 (2011), 330–349. DOI: https://doi.org/10.1108/00251741111120734

Tanja Blascheck, Lindsay MacDonald Vermeulen, Jo Vermeulen, Charles Perin, Wesley Willett, Thomas Ertl, and Sheelagh Carpendale. 2018. Exploration strategies for discovery of interactivity in visualizations. *IEEE Transactions on Visualization and Computer Graphics* 25, 2 (2018), 1407–1420. DOI: https://doi.org/10.1109/TVCG.2018.2802520

Berndt Brehmer. 1992. Dynamic decision making: Human control of complex systems. *Acta Psychologica* 81, 3 (1992), 211–241. DOI: https://doi.org/10.1016/0001-6918(92)90019-A

John Brooke. 1996. SUS: A "quick and dirty" usability scale. In *Usability Evaluation in Industry* (1st ed.). Patrick W. Jordan, B. Thomas, Ian Lyall McClelland, and Weerdmeester Bernard (Eds.), CRC Press, London, 6. ISBN: 9780748404605. DOI: https://doi.org/10.1201/9781498710411

John Brooke. 2013. SUS: A retrospective. *Journal of Usability Studies* 8, 2 (2013), 29–40. DOI: https://doi.org/10.5555/2817912.2817913

Michael Burch and Hansjörg Schmauder. 2018. Challenges and perspectives of interacting with hierarchy visualizations on large-scale displays. In *Proceedings of the 11th International Symposium on Visual Information Communication and Interaction*. Andreas Kerren, Karsten Klein, and Yi-Na Li (Eds.), ACM, New York, NY, 33–40. DOI: https://doi.org/10.1145/3231622.3231635

Simon Butscher, Sebastian Hubenschmid, Jens Müller, Johannes Fuchs, and Harald Reiterer. 2018. Clusters, trends, and outliers: How immersive technologies can facilitate the collaborative analysis of multidimensional data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–12. DOI: https://doi.org/10.1145/3173574.3173664

Herbert Clark and Meredyth Krych. 2004. Speaking while monitoring addressees for understanding. *Journal of Memory and Language* 50, 1 (2004), 62–81. DOI: https://doi.org/10.1016/j.jml.2003.08.004

Matteo Cognolato, Manfredo Atzori, and Henning Müller. 2018. Head-mounted eye gaze tracking devices: An overview of modern devices and recent advances. *Journal of Rehabilitation and Assistive Technologies Engineering* 5 (2018), 2055668318773991. DOI: https://doi.org/10.1177/2055668318773991

Elizabeth G. Cohen. 1994. Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research* 64, 1 (1994), 1–35. DOI: https://doi.org/10.3102/00346543064001001

Sarah D'Angelo and Darren Gergle. 2018. An eye for design: Gaze visualizations for remote collaborative work. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–12. DOI: https://doi.org/10.1145/3173574.3173923

Sebastian Derwisch, Birgit Kopainsky, and Katharine Troeger. 2011. The management of trust for stimulating the adoption of improved maize seed in Malawi. In *Proceedings of the International Conference of the System Dynamics Society*. System Dynamic Society, Washington, DC, 24–28. Retrieved from https://proceedings.systemdynamics.org/2011/proceed/papers/P1226.pdf.

Paul Dourish and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work*. ACM, New York, NY, 107–114. DOI: https://doi.org/10.1145/143457.143468

Andrew T. Duchowski, Krzysztof Krejtz, Nina A Gehrer, Tanya Bafna, and Per Bækgaard. 2020. The low/high index of pupillary activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–12. DOI: https://doi.org/10.1145/3313831.3376394

Gerald Echterhoff, E. Tory Higgins, and John M. Levine. 2009. Shared reality: Experiencing commonality with others' inner states about the world. *Perspectives on Psychological Science* 4, 5 (2009), 496–521. DOI: https://doi.org/10.1111/j.1745-6924.2009.01161.x

Mohammed El Maniani, M. Rechchach, A. El Mahfoudi, M. El Moudane, and A. Sabbar. 2016. A calorimetric investigation of the liquid Bi-Ni alloys. *Journal of Materials and Environmental Science* 7, 10 (2016), 3759–3766.

Edgar Erdfelder, Franz Faul, and Axel Buchner. 1996. GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers* 28, 1 (1996), 1–11.

Taciana Falcao and Sara Price. 2011. Interfering and resolving: How tabletop interaction facilitates co-construction of argumentative knowledge. *International Journal of Computer-Supported Collaborative Learning* 6 (2011), 539–559. DOI: https://doi.org/10.1007/s11412-010-9101-9

Likun Fang, Ting Zhu, Erik Pescara, Yiran Huang, Yexu Zhou, and Michael Beigl. 2022. DragTapVib: An on-skin electromagnetic drag, tap, and vibration actuator for wearable computing. In *Proceedings of the Augmented Humans International Conference 2022*. ACM, New York, NY, 203–211. DOI: https://doi.org/10.1145/3519391.3519395

Andy Field, Jeremy Miles, and Zoe Field. 2012. *Discovering Statistics Using R*. SAGE Publications, London.

Morten Fjeld, Jonas Fredriksson, Martin Ejdestig, Florin Duca, Kristina Bötschi, Benedikt M. Voegtli, and Patrick Juchli. 2007. Tangible user interface for chemistry education: Comparative evaluation and re-design. In *Proceedings of the 2007 Conference on Human Factors in Computing Systems*. Mary Beth Rosson and David J. Gilmore (Eds.), ACM, New York, NY, 805–808. DOI: https://doi.org/10.1145/1240624.1240745

Alexandra Frischen, Andrew P. Bayliss, and Steven P. Tipper. 2007. Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin* 133, 4 (2007), 694.

Joachim Funke. 1991. Solving complex problems: Exploration and control of complex systems. In *Complex Problem Solving: Principles and Mechanisms* (1st ed.). R. J. Sternberg and P. A. Frensch (Eds.), Vol. 4, Psychology Press, New York, NY, 185–222. DOI: https://doi.org/10.4324/9781315807546

Susan R. Fussell, Robert E. Kraut, and Jane Siegel. 2000. Coordination of communication: Effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*. ACM, New York, NY, 21–30. DOI: https://doi.org/10.1145/358916.358947

Michael Shayne Gary and Robert E. Wood. 2016. Unpacking mental models through laboratory experiments. *System Dynamics Review* 32, 2 (2016), 101–129. DOI: https://doi.org/10.1002/sdr.1560

Susanne Geister, Udo Konradt, and Guido Hertel. 2006. Effects of process feedback on motivation, satisfaction, and performance in virtual teams. *Small Group Research* 37, 5 (2006), 459–489. DOI: https://doi.org/10.1177/1046496406292337

Darren Gergle and Alan T. Clark. 2011. See what I'm saying? Using dyadic mobile eye tracking to study collaborative reference. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*. ACM, New York, NY, 435–444. DOI: https://doi.org/10.1145/1958824.1958892

Joseph Goldberg and Jonathan Helfman. 2011. Eye tracking for visualization evaluation: Reading values on linear versus radial graphs. *Information Visualization* 10, 3 (2011), 182–195. DOI: https://doi.org/10.1177/1473871611406623

Katerina Gorkovenko, Lars Lischke, and Paweł W. Woźniak. 2018. Large displays and tablets: Data exploration and its effects on data collection. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*. ACM, New York, NY, 664–675. DOI: https://doi.org/10.1145/3240167.3240192

Marianne Gullberg and Kenneth Holmqvist. 1999. Keeping an eye on gestures: Visual perception of gestures in face-to-face communication. *Pragmatics & Cognition* 7, 1 (1999), 35–63. DOI: https://doi.org/10.1075/pc.7.1.04gul

Carl Gutwin, Saul Greenberg, and Mark Roseman. 1996. Workspace awareness in real-time distributed groupware: Framework, widgets, and evaluation. In *People and Computers XI*. Martina Angela Sasse, R. Jim Cunningham, and Russel L. Winder (Eds.), Springer London, London, 281–298. DOI: https://doi.org/10.1007/978-1-4471-3588-3_18

Sophie Guy, Yoshihisa Kashima, Iain Walker, and Saffron O'Neill. 2013. Comparing the atmosphere to a bathtub: Effectiveness of analogy for reasoning about accumulation. *Climatic Change* 121, 4 (2013), 579–594. DOI: https://doi.org/10.1007/s10584-013-0949-3

Takuma Hagiwara, Kazuki Takashima, Morten Fjeld, and Yoshifumi Kitamura. 2019. CamCutter: Impromptu vision-based cross-device application sharing. *Interacting with Computers* 31, 6 (2019), 539–554. DOI: https://doi.org/10.1093/iwcomp/iwz035

Roland Maximilian Happach and Billy Schoenberg. 2017. Linking simulator functionality with learning: An extension of the taxonomy of computer simulations to support learning. In *Proceedings of the 35th International Conference of the System Dynamics Society*, Vol. 2. System Dynamics Society, Cambridge, Massachusetts, 741–764. Retrieved from https://proceedings.systemdynamics.org/2017/proceed/supp/S1242.pdf.

Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in Psychology*. P.A. Hancock and N. Meshkati (Eds.), Vol. 52, Elsevier, North-Holland, 139–183. DOI: https://doi.org/10.1016/S0166-4115(08)62386-9

Jeffrey Heer and Maneesh Agrawala. 2007. Design considerations for collaborative visual analytics. *Information Visualization* 7, 1 (2007), 49–62. DOI: https://doi.org/10.1109/VAST.2007.4389011

Roy S. Hessels, Chantal Kemner, Carlijn van den Boomen, and Ignace T. C. Hooge. 2016. The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli. *Behavior Research Methods* 48, 4 (2016), 1694–1712. DOI: https://doi.org/10.3758/s13428-015-0676-y

Kathleen Hogan, Bonnie K. Nastasi, and Michael Pressley. 1999. Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction* 17, 4 (1999), 379–432. DOI: https://doi.org/10.1207/S1532690XCI1704_2

Judith Holler and Kobin H. Kendrick. 2015. Unaddressed participants' gaze in multi-person interaction: Optimizing recipiency. *Frontiers in Psychology* 6, 98 (2015), 1–14. DOI: https://doi.org/10.3389/fpsyg.2015.00098 isee systems inc. 2020. Stella®Architect Premium Modeling and Interactive Simulation. Retrieved from https://www.iseesystems.com/store/products/stella-architect.aspx. Accessed 12-01-2020.

Mikkel R. Jakobsen and Kasper Hornbæk. 2013. Interactive visualizations on large and small displays: The interrelation of display size, information space, and scale. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2336–2345. DOI: https://doi.org/10.1109/TVCG.2013.170

Eva Jensen. 2005. Learning and transfer from a simple dynamic system. *Scandinavian Journal of Psychology* 46, 2 (2005), 119–131. DOI: https://doi.org/10.1111/j.1467-9450.2005.00442.x

Kristiina Jokinen, Masafumi Nishida, and Seiichi Yamamoto. 2009. Eye-gaze experiments for conversation monitoring. In *Proceedings of the 3rd International Universal Communication Symposium.* ACM, New York, NY, 303–308. DOI: https://doi.org/10.1145/1667780.1667843

Kristiina Jokinen, Masafumi Nishida, and Seiichi Yamamoto. 2010. On eye-gaze and turn-taking. In *Proceedings of the 2010 Workshop on Eye Gaze in Intelligent Human Machine Interaction.* ACM, New York, NY, 118–123. DOI: https://doi.org/10.1145/2002333.2002352

Marcel Adam Just and Patricia A. Carpenter. 1976. The role of eye-fixation research in cognitive psychology. *Behavior Research Methods & Instrumentation* 8, 2 (1976), 139–143. DOI: https://doi.org/10.3758/BF03201761

Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* ACM, New York, NY, 1151–1160. DOI: https://doi.org/10.1145/2638728.2641695

Adam Kendon. 1988. How gestures can become like words. In *Cross-Cultural Perspectives in Nonverbal Communication.* F. Poyatos (Ed.), Hogrefe & Huber Publishers, Chicago, 131–141.

Henna Kim and Sara Snow. 2013. Collaboration on a large-scale, multi-touch display: Asynchronous interaction and multiple-input use. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion* . Amy S. Bruckman, Scott Counts, Cliff Lampe, and Loren G. Terveen (Eds.), ACM, New York, NY, 165–168. DOI: https://doi.org/10.1145/2441955.2441997

Dimosthenis Kontogiorgos, Vanya Avramova, Simon Alexanderson, Patrik Jonell, Catharine Oertel, Jonas Beskow, Gabriel Skantze, and Joakim Gustafson. 2018. A multimodal corpus for mutual gaze and joint attention in multiparty situated interaction. In *Proceedings of the 11th International Conference on Language Resources and Evaluation.* European Language Resources Association (ELRA), Miyazaki, 119–127. Retrieved from https://aclanthology.org/L18-1019.

Birgit Kopainsky and Stephen M. Alessi. 2015. Effects of structural transparency in system dynamics simulators on performance and understanding. *Systems* 3, 4 (2015), 152–176. DOI: https://doi.org/10.3390/systems3040152

Birgit Kopainsky and Agata Sawicka. 2011. Simulator-supported descriptions of complex dynamic problems: Experimental results on task performance and system understanding. *System Dynamics Review* 27, 2 (2011), 142–172. DOI: https://doi.org/10.1002/sdr.445

Krzysztof Krejtz, Andrew T. Duchowski, Izabela Krejtz, Agata Kopacz, and Piotr Chrząstowski-Wachtel. 2016. Gaze transitions when learning with multimedia. *Journal of Eye Movement Research* 9, 1 (2016), 1–17. DOI: https://doi.org/10.16910/jemr.9.1.5

Krzysztof Krejtz, Justyna Żurawska, Andrew Duchowski, and Szymon Wichary. 2020. Pupillary and microsaccadic responses to cognitive effort and emotional arousal during complex decision making. *Journal of Eye Movement Research* 13, 5 (2020), 1–15. DOI: https://doi.org/10.16910/jemr.13.5.2

Yoshiki Kudo, Anthony Tang, Kazuyuki Fujita, Isamu Endo, Kazuki Takashima, and Yoshifumi Kitamura. 2021. Towards balancing VR immersion and bystander awareness. *Proceedings of the ACM on Human-Computer Interaction* 5, ISS, Article 484 (2021), 22 pages. DOI: https://doi.org/10.1145/3486950

Grete Helena Kütt, Kevin Lee, Ethan Hardacre, and Alexandra Papoutsaki. 2019. Eye-write: Gaze sharing for collaborative writing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, 1–12. DOI: https://doi.org/10.1145/3290605.3300727

Lasse Lipponen, Marjaana Rahikainen, Jiri Lallimo, and Kai Hakkarainen. 2003. Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction* 13, 5 (2003), 487–509. DOI: https://doi.org/10.1016/S0959-4752(02)00042-7

Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, and Eric Lecolinet. 2016. Shared interaction on a wall-sized display in a data manipulation task. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, 2075–2086. DOI: https://doi.org/10.1145/2858036.2858039

Gustavo Lopez and Luis A. Guerrero. 2017. Awareness supporting technologies used in collaborative systems: A Systematic Literature Review. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing.* ACM, New York, NY, 808–820. DOI: https://doi.org/10.1145/2998181.2998281

Weiquan Lu, Been-Lirn Henry Duh, and Steven Feiner. 2012. Subtle cueing for visual search in augmented reality. In *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality.* 161–166. DOI: https://doi.org/10.1109/ISMAR.2012.6402553

Magdalena Mateescu, Christoph Pimmer, Carmen Zahn, Daniel Klinkhammer, and Harald Reiterer. 2021. Collaboration on large interactive displays: A systematic review. *Human–Computer Interaction* 36, 3 (2021), 243–277. DOI: https://doi.org/10.1080/07370024.2019.1697697

Bernhard Maurer, Alina Krischkowsky, and Manfred Tscheligi. 2017. Exploring gaze and hand gestures for non-verbal in-game communication. In *Proceedings of the Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play*. ACM, New York, NY, 315–322. DOI: https://doi.org/10.1145/3130859.3131296

Sven Mayer, Lars Lischke, Paweł W. Woźniak, and Niels Henze. 2018. Evaluating the disruptiveness of mobile interactions: A mixed-method approach. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–14. DOI: https://doi.org/10.1145/3173574.3173980

Katherine L. Milkman, Dolly Chugh, and Max H. Bazerman. 2009. How can decision making be improved? *Perspectives on Psychological Science* 4, 4 (2009), 379–383. DOI: https://doi.org/https:/doi.org/10.1111/j.1745-6924.2009.01142.x

Erling Moxnes. 1998a. Not only the tragedy of the commons: Misperceptions of bioeconomics. *Management Science* 44, 9 (1998), 1234–1248. Retrieved from https://www.jstor.org/stable/2634712.

Erling Moxnes. 1998b. Overexploitation of renewable resources: The role of misperceptions. *Journal of Economic Behavior & Organization* 37, 1 (1998), 107–127. DOI: https://doi.org/10.1016/S0167-2681(98)00079-1

Erling Moxnes. 2004. Misperceptions of basic dynamics: The case of renewable resource management. *System Dynamics Review* 20, 2 (2004), 139–162. DOI: https://doi.org/10.1002/sdr.289

Erling Moxnes. 2014. Discounting, climate and sustainability. *Ecological Economics* 102 (2014), 158–166. DOI: https://doi.org/10.1016/j.ecolecon.2014.04.003

Philipp Müller, Michael Xuelin Huang, Xucong Zhang, and Andreas Bulling. 2018. Robust eye contact detection in natural multi-person interactions using gaze and speaking behaviour. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ACM, New York, NY, 1–10. DOI: https://doi.org/10.1145/3204493.3204549

Yong S. Nyam, Julius H. Kotir, Andries J. Jordaan, Abiodun A. Ogundeji, Adetoso A. Adetoro, and Israel R. Orimoloye. 2020. Towards understanding and sustaining natural resource systems through the systems perspective: A systematic evaluation. *Sustainability* 12, 23 (2020), 1–20. DOI: https://doi.org/10.3390/su12239871

Edwin Olson. 2011. AprilTag: A robust and flexible visual fiducial system. In *Proceedings of the 2011 IEEE International Conference on Robotics and Automation*. IEEE, Shanghij, 3400–3407. DOI: https://doi.org/10.1109/ICRA.2011.5979561

Gary M. Olson and Judith S. Olson. 2000. Distance matters. *Human–Computer Interaction* 15, 2–3 (2000), 139–178. DOI: https://doi.org/10.1207/S15327051HCI1523_4

Mai Otsuki, Taiki Kawano, Keita Maruyama, Hideaki Kuzuoka, and Yusuke Suzuki. 2017. ThirdEye: Simple add-on display to represent remote participant's gaze direction in video communication. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 5307–5312. DOI: https://doi.org/10.1145/3025453.3025681

Harshada Patel, Michael Pettitt, and John R. Wilson. 2012. Factors of collaborative working: A framework for a collaboration model. *Applied Ergonomics* 43, 1 (2012), 1–26. DOI: https://doi.org/10.1016/j.apergo.2011.04.009

James Patten, Hiroshi Ishii, Jim Hines, and Gian Pangaro. 2001. Sensetable: A wireless object tracking platform for tangible user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 253–260. DOI: https://doi.org/10.1145/365024.365112

Ilaria Perissi, Ugo Bardi, Toufic El Asmar, and Alessandro Lavacchi. 2017. Dynamic patterns of overexploitation in fisheries. *Ecological Modelling* 359 (2017), 285–292. DOI: https://doi.org/10.1016/j.ecolmodel.2017.06.009

Ulrich J. Pfeiffer, Kai Vogeley, and Leonhard Schilbach. 2013. From gaze cueing to dual eye-tracking: Novel approaches to investigate the neural correlates of gaze in social interaction. *Neuroscience & Biobehavioral Reviews* 37, 10 (2013), 2516–2528. DOI: https://doi.org/10.1016/j.neubiorev.2013.07.017

Sami Pietinen, Roman Bednarik, Tatiana Glotova, Vesa Tenhunen, and Markku Tukiainen. 2008a. A method to study visual attention aspects of collaboration: Eye-tracking pair programmers simultaneously. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*. ACM, New York, NY, 39–42. DOI: https://doi.org/10.1145/1344471.1344480

Sami Pietinen, Roman Bednarik, Tatiana Glotova, Vesa Tenhunen, and Markku Tukiainen. 2008b. A method to study visual attention aspects of collaboration: Eye-tracking pair programmers simultaneously. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*. ACM, New York, NY, 39–42. DOI: https://doi.org/10.1145/1344471.1344480

Michael I. Posner. 2004. *Cognitive Neuroscience of Attention*. Guilford Press, New York, NY .

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

Jeremy Roschelle and Stephanie D. Teasley. 1995. The construction of shared knowledge in collaborative problem solving. In *Computer Supported Collaborative Learning*. C. O'Malley (Ed.), Springer, London, 69–97. DOI: https://doi.org/10.1007/978-3-642-85098-1_5

Maike Schindler and Achim Lilienthal. 2017. Eye-tracking and its domain-specific interpretation: A stimulated recall study on eye movements in geometrical tasks. In *Proceedings of the Annual Meeting of the International Group for the Psychology of Mathematics Education*. Vol. 4, PME, Singapore, 153–160.

Bertrand Schneider and Roy Pea. 2013. Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning* 8, 4 (2013), 375–397. DOI: https://doi.org/10.1007/s11412-013-9181-4

Bertrand Schneider, Kshitij Sharma, Sébastien Cuendet, Guillaume Zufferey, Pierre Dillenbourg, and Roy Pea. 2016. Using mobile eye-trackers to unpack the perceptual benefits of a tangible user interface for collaborative learning. *ACM Transactions on Computer-Human Interaction* 23, 6, Article 39 (2016), 23 pages. DOI: https://doi.org/10.1145/3012009

Bertrand Schneider, Kshitij Sharma, Sebastien Cuendet, Guillaume Zufferey, Pierre Dillenbourg, and Roy Pea. 2018. Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning* 13, 3 (2018), 241–261. DOI: https://doi.org/10.1007/s11412-018-9281-2

Bertrand Schneider, Megan Strait, Laurence Muller, Sarah Elfenbein, Orit Shaer, and Chia Shen. 2012. Phylo-Genie: Engaging students in collaborative 'tree-thinking' through tabletop techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3071–3080. DOI: https://doi.org/10.1145/2207676.2208720

Daniel L. Schwartz. 1995. The emergence of abstract representations in dyad problem solving. *The Journal of the Learning Sciences* 4, 3 (1995), 321–354. DOI: https://doi.org/10.1207/s15327809jls0403_3

Bariş Serim, Ken Pfeuffer, Hans Gellersen, and Giulio Jacucci. 2018. Visual attention-based access: Granting access based on users' joint attention on shared workspaces. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3, Article 133 (2018), 23 pages. DOI: https://doi.org/10.1145/3264943

Keisuke Shiro, Atsushi Okada, Takashi Miyaki, and Jun Rekimoto. 2018. OmniGaze: A display-covered omnidirectional camerafor conveying remote user's presence. In *Proceedings of the 6th International Conference on Human-Agent Interaction*. ACM, New York, NY, 176–183. DOI: https://doi.org/10.1145/3284432.3284439

Hajime Shirouzu, Naomi Miyake, and Hiroyuki Masukawa. 2002. Cognitively active externalization for situated reflection. *Cognitive Science* 26, 4 (2002), 469–501. DOI: https://doi.org/10.1207/s15516709cog2604_3

Kevin Shockley, Marie-Vee Santana, and Carol Fowler. 2003. Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology. Human Perception and Performance* 29, 2 (2003), 326–332. DOI: https://doi.org/10.1037/0096-1523.29.2.326

Anna Shvarts, Andrey Stepanov, and Dmitry Chumachenko. 2018. Automatic detection of gaze convergence in multimodal collaboration: A dual eye-tracking technology. *The Russian Journal of Cognitive Science* 5, 3 (2018), 4–17.

Oleg Špakov and Darius Miniotas. 2007. Visualization of eye gaze data using heat maps. *Elektronika Ir Elektrotechnika Medicine Technology Technologija* 74, 2 (2007), 55–58. Retrieved from https://eejournal.ktu.lt/index.php/elt/article/view/10372.

Gerry Stahl, Nancy Law, Ulrike Cress, and Sten Ludvigsen. 2014. Analyzing roles of individuals in small-group collaboration processes. *International Journal of Computer-Supported Collaborative Learning* 9, 4 (2014), 365–370. DOI: https://doi.org/10.1007/s11412-014-9204-9

John Sterman. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin McGrawHill, Boston, MA. Retrieved from http://www.mhhe.com/business/opsci/sterman/.

John D. Sterman. 1989. Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes* 43, 3 (1989), 301–335. DOI: https://doi.org/10.1016/0749-5978(89)90041-1

John D. Sterman. 2008. Risk communication on climate: Mental models and mass balance. *Science* 322, 5901 (2008), 532–533. DOI: https://doi.org/10.1126/science.1162574

John D. Sterman and Linda Sweeney. 2007. Understanding public complacency about climate change: Adults' mental models of climate change violate conservation of matter. *Climatic Change* 80 (2007), 213–238. DOI: https://doi.org/10.1007/s10584-006-9107-5

Susan Straus and Joseph McGrath. 1994. Does the medium matter? The interaction of task type and technology on group performance and member reactions. *The Journal of Applied Psychology* 79, 1 (1994), 87–97. DOI: https://doi.org/10.1037//0021-9010.79.1.87

Mohsen Tavakol and Reg Dennick. 2011. Making sense of Cronbach's alpha. *International Journal of Medical Education* 2 (2011), 53–55. DOI: https://doi.org/10.5116/ijme.4dfb.8dfd

Lori Thompson and Michael Coovert. 2003. Teamwork online: The effects of computer conferencing on perceived confusion, satisfaction, and postdiscussin accuracy. *Group Dynamics: Theory, Research, and Practice* 7, 2 (2003), 135–151. DOI: https://doi.org/10.1037/1089-2699.7.2.135

Tom Tietenberg and Lynne Lewis. 2019. *Natural Resource Economics: The Essentials*. Routledge, New York, NY. DOI: https://doi.org/10.4324/9780429299285

Michael Tomasello. 1995. Joint attention as social cognition. In *Joint Attention: Its Origins and Role in Development*. C. Moore and P. J. Dunham (Eds.), Lawrence Erlbaum Associates, Inc., New Jersey, NJ, 103–130. DOI: https://doi.org/10.4324/9781315806617

Consuelo Valdes, Michelle Ferreirae, Taili Feng, Heidi Wang, Kelsey Tempel, Sirui Liu, and Orit Shaer. 2012. A collaborative environment for engaging novices in scientific inquiry. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*. ACM, New York, NY, 109–118. DOI: https://doi.org/10.1145/2396636.2396654

Hidde van der Meulen, Petra Varsanyi, Lauren Westendorf, Andrew L. Kun, and Orit Shaer. 2016. Towards understanding collaboration around interactive surfaces: Exploring joint visual attention. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, New York, NY, 219–220. DOI: https://doi.org/10.1145/2984751.2984778

Boris M. Velichkovsky. 1995. Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics & Cognition* 3, 2 (1995), 199–223. DOI: https://doi.org/10.1075/pc.3.2.02vel

Roel Vertegaal. 1999. Designing awareness with attention-based groupware. In *Proceedings of the IFIP TC International Conference on Human-Computer Interaction*. M. Angela Sasse and Chris W. Johnson (Eds.), IOS Press, Edinburgh, 245–255.

Roel Vertegaal. 2002. Designing attentive interfaces. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*. ACM, New York, NY, 23–30. DOI: https://doi.org/10.1145/507072.507077

Roel Vertegaal. 2003. Attentive user interfaces. *Communications of the ACM* 46, 3 (2003), 30–33. DOI: https://doi.org/10.1145/636772.636794

Haofei Wang and Bertram E. Shi. 2019. Gaze awareness improves collaboration efficiency in a collaborative assembly task. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ACM, New York, NY, Article 88, 5 pages. DOI: https://doi.org/10.1145/3317959.3321492

Noreen M. Webb, Kariane Mari Nemer, and Stephen Zuniga. 2002. Short circuits or superconductors? Effects of group composition on high-achieving students' science assessment performance. *American Educational Research Journal* 39, 4 (2002), 943–989. DOI: https://doi.org/10.3102/00028312039004943

Sara J. Weston and Debbie Yee. 2017. Why You Should Become a UseR: A Brief Introduction to R. (2017). Retrieved from https://www.psychologicalscience.org/observer/why-you-should-become-a-user-a-brief-introduction-to-r. Accessed 12-01-2020.

Donna Wood. 1991. Toward a comprehensive theory of collaboration. *The Journal of Applied Behavioral Science* 27, 2 (1991), 139–162. DOI: https://doi.org/10.1177/0021886391272001

Paweł Wozniak, Nitesh Goyal, Przemysław Kucharski, Lars Lischke, Sven Mayer, and Morten Fjeld. 2016. RAMPARTS: Supporting sensemaking with spatially-aware mobile interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2447–2460. DOI: https://doi.org/10.1145/2858036.2858491

Nancy Yao, Jeff Brewer, Sarah D'Angelo, Mike Horn, and Darren Gergle. 2018. Visualizing gaze information from multiple students to support remote instruction. In *Proceedings of the Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 1–6. DOI: https://doi.org/10.1145/3170427.3188453

Chen Yu and Linda B. Smith. 2017. Hand–eye coordination predicts joint attention. *Child Development* 88, 6 (2017), 2060–2078. DOI: https://doi.org/10.1111/cdev.12730

Received 20 May 2021; revised 3 August 2022; accepted 20 November 2022