Applied Ergonomics 80 (2019) 175–186

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

An empirical study to investigate the efficacy of collaborative immersive virtual reality systems for designing information architecture of software systems

Shraddhaa Narasimha[a], Emma Dixon[a], Jeffrey W. Bertrand[a], Kapil Chalil Madathil[b,*]

[a] Department of Industrial Engineering, 110 Freeman Hall, Clemson, SC, 29634, USA

[b] Departments of Industrial and Civil Engineering, Clemson University, 110 Freeman Hall, Clemson, SC, 29634, USA

ARTICLE INFO

Keywords:
Computer-supported collaborative work
Virtual reality
Immersive systems
Information architecture design
Card sorting

ABSTRACT

The ability of Immersive Virtual Reality (IVR) systems to mimic the real world has made it possible to use this technology to create environments for remote collaborative work. This study aimed to understand the feasibility of immersive virtual reality when conducting a collaborative Information Architecture (IA) design task-card sorting, with geographically dispersed participants. Using a between-subjects experimental design, thirty groups of two individuals each completed a card sorting activity using conventional in-person, video screen-sharing method or immersive virtual reality methods. The dependent measures included total time, percentage match with master card set, usability, presence and perceived workload. Overall usability was found to be significantly higher for the immersive virtual reality condition when compared to conventional in-person card sorting. In addition, the new immersive virtual reality technology performed as well as the other two conditions for other dependent variables. Qualitative data from the participants also indicated a positive reaction to the use of immersive virtual reality for this task. Overall, the participants felt they were productive and enjoyed the IVR condition, indicating the potential of IVR-based approaches as an alternative to conventional approaches for IA design.

1. Introduction

Information system design and development is an interdisciplinary activity requiring collaborative involvement from multiple stakeholders (Choi et al., 2015; Eppinger and Ulrich, 2011). Collaboration is often required on user-centered design and usability testing procedures, generally involving the users, designers and other stakeholders to develop a product that fits the needs of the end user (Wentzel et al., 2016; Jaspers, 2009; Pagliari, 2007). User-centered design takes into account the needs, desires, and limitations of end users at all stages of system development to incorporate those requirements into the product (Juárez-Ramírez, 2017; Rubin and Chisnell, 2011). Information Architecture (IA) design, a subset of user-centered design process, is the design, implementation and evaluation of an information space to facilitate task completion and intuitive access to content (Dillon, 2002; Rosenfeld and Morville, 2002). The foundational components for creating an IA include organizing schemes and structures, labeling systems, navigating systems, and search systems (Rosenfeld and Morville, 2002). Card sorting is one commonly used IA design tool to understand how users categorize the information that will appear in an information space (Sinha, 2003; Spencer, 2009). Generally, in a card sorting session, users meet in person to organize topics into categories that make sense to them (Maguire, 2001; Wentzel et al., 2016). The results of a card sorting exercise are used to develop the Information Architecture of a software system. Literature suggests two techniques to conduct card sorting: open card sort and closed card sort. An open card sort allows participants to group the cards in as many categories as they wish. In an open card sort, participants are either asked to identify one of the existing cards in each group as the title card or create a title card for each group (Wentzel et al., 2016). A closed card sort gives users a set of predefined title cards (categories) and asks the users to organize the cards into those predefined categories (Karreman, Arendsen and van der Geest, 2010). Traditionally, card sorting involves participants sorting cards into meaningful groups and requires the physical presence of all parties involved (Eppinger and Ulrich, 2011; Wentzel et al., 2016; Zimmerman and Akerelrea, 2002).

For global enterprises, collaboration in a software system design can be difficult due to cost, time and other logistical constraints.

* Corresponding author. Departments of Industrial and Civil Engineering, Clemson University, Clemson, SC, 29634, USA.
E-mail address: kmadath@clemson.edu (K. Chalil Madathil).

https://doi.org/10.1016/j.apergo.2019.05.009
Received 6 September 2017; Received in revised form 6 December 2018; Accepted 20 May 2019
Available online 18 June 2019

Considering the vast potential of immersive virtual reality for remote technology assisted collaboration, the possibility of carrying out card sorting remotely remains an open question. Immersive Virtual Reality (IVR) is sometimes referred to as Immersive Computing Technology (ICT) and consists of technology that enables people to immerse themselves in a virtual world (Berg and Vance, 2017). Although IVR in its present form started roughly 50 years ago, only over the past 25 years has a vast amount of research been conducted in this area (Sun et al., 2015; Slater and Sanchez-Vives, 2016). Ivan Sutherland envisioned the potential for virtual reality and created some of the initial technology for immersive VR experiences in the 1960s (Sutherland, 1965; Slater and Sanchez-Vives, 2016). In the 1980s, NASA developed a full VR system called VIEW consisting of all the same elements used today, including a light-weight Head Mounted Display (HMD), audio, body tracking, tracking gloves to allow interaction with objects in the virtual environment, tactile and force feedback, and connection capabilities to telerobotic systems (Slater and Sanchez-Vives, 2016). Today, display technologies are available in various forms and sizes, including a single large projection screen, multiple connected projection screens known as CAVE (Cruz-Neira et al., 1993), stereo-capable monitors with desktop tracking, input device and head mounted displays (Berg and Vance, 2017; Nichols, 1999; Nichols and Patel, 2002). Audio is generally provided through headphones, single-speaker, or a full-surround system as appropriate for the environment (Berg and Vance, 2017). Immersive VR experiences vary depending on the level of simulation used and the corresponding technology leading to the need to choose the right technology to meet task demands (Azuma, 1997; Wilson, 1999).

The ability of virtual reality technology to mimic the real world, providing a sense of place and reality leading to a superior sensory experience, is its distinguishing factor (Burdea and Coiffet, 2003; Slater and Sanchez-Vives, 2016). This unique characteristic of virtual reality has resulted in a wide array of applications including the medical field which has benefited from this technology as a training platform for new surgical procedures and to train novices (Liu et al., 2003; Alaraj et al., 2011; Seymour et al., 2002; Slater and Sanchez-Vives, 2016). It has been used extensively in education as a platform for enhanced visualization, motivation, and interaction with a positive outcome among teachers and students (Gopinath and Tucker, 2015; Mikropoulos and Natsis, 2011; Pantelidis, 1995). Telepresence for remote communication, architectural visualization, psychology (Slater and Sanchez-Vives, 2016), molecular biology (3D view of molecular structures) (Lee et al., 2009), aircraft inspection training (Vora et al., 2002), transportation research (Deb et al., 2017) and VR-based ergonomic design of workstations (Nguyen et al., 2017) are among the areas where it has been applied. With an increase in computing power and decrease in the costs of this technology, VR now encompasses leisure and entertainment applications (Howarth and Finch, 1999; Zyda, 2005). Another important area of application of virtual reality is remote collaboration, such as factory planning, engineering education, engineering design, and new product design (Christian, Madathil et al., 2017; Christian, Galambos, Csapó, Zentay Baranyi, 2014; Bertrand et al., 2015; Bhargava et al., 2018; Berg and Vance, 2017 Gopinath and Tucker, 2015; Slater and Sanchez-Vives, 2016).

Research related to technology-assisted collaborative work focuses on work activity and seeking ways to distribute and coordinate the activity across geographically dispersed users (Benford et al., 2000). When people meet to carry out a task, awareness of group members and processes, referred to as group awareness, is an important aspect. Technological advances in communication and collaboration has attempted to recreate group awareness which forms the basis for Computer Supported Cooperative Work (CSCW) (Gross et al., 2005). Of the many different perspectives of CSCW, the technology-centric view is defined as developing ways to design computer technology that will further support collaboration and people working collectively (Greif, 1988; Mills, 2003). CSCW is synchronous when people work in the same room at the same time, distributed synchronous when people work at the same time but in different places. It is asynchronous when people work on a task at different times but at the same place, and distributed and asynchronous when the task is being performed in different places and at different times (Palmer and Fields, 1994).

Knowledge exchange, data visualization, and task complexity analysis are some areas in which computer supported remote collaboration can be used (Fleury et al., 2015). Although video conferencing and screen sharing are some methods of CSCW, recently virtual reality has taken precedence. It has been applied to, and studied in, the field of education with students participating through virtual avatars (De Lucia, Francese, Passero and Tortora, 2009). The potential advantage of virtual reality in substituting traditional meetings and conventional online meeting applications has also been studied for engineering and product design (Firat Ozkan and Greenstein, 2011).

Collaborative virtual reality systems are generally three-dimensional (3D) simulated environments that help users interact with each other in real time and can participate through their ‘avatars’ (Chalil Madathil & Greenstein, 2011, 2017; Narasimha et al., 2018a). With a simulated digital landscape, these collaborative virtual reality systems assist collaborative work between teams that are geographically separated (Churchill and Snowdon, 1998). Understanding collaboration in these 3D environments is necessary to evaluate and improve these systems. A review of the literature indicates the presence of multiple models designed to understand collaboration in these environments including the Awareness Evaluation Model (Neale et al., 2004). This framework (Fig. 1) identifies multiple variables to consider when evaluating collaborative activities in computer-supported environments (Neale et al., 2004). The variables include contextual factors, work coupling, communication, coordination, and common ground between group members (Neale et al., 2004). Each of these components combine to produce activity awareness within a group. This framework for assessing activity awareness can be used to evaluate collaboration in both face-to-face interactions and technology-assisted collaborative activities.

1.1. Motivation

The emergence of high speed internet technologies has resulted in the concept of the global village, and next generation products addressing its needs. In such a scenario when prospective users are wide-

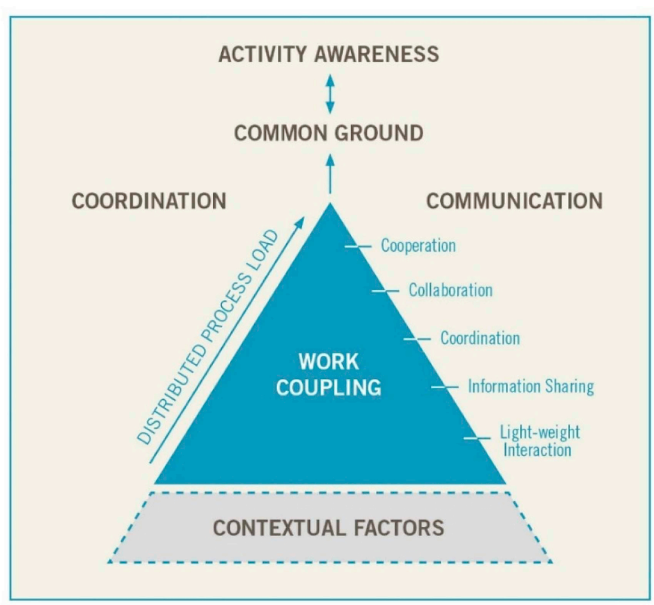

Fig. 1. Awareness evaluation model (adapted from Neale et al., 2004).

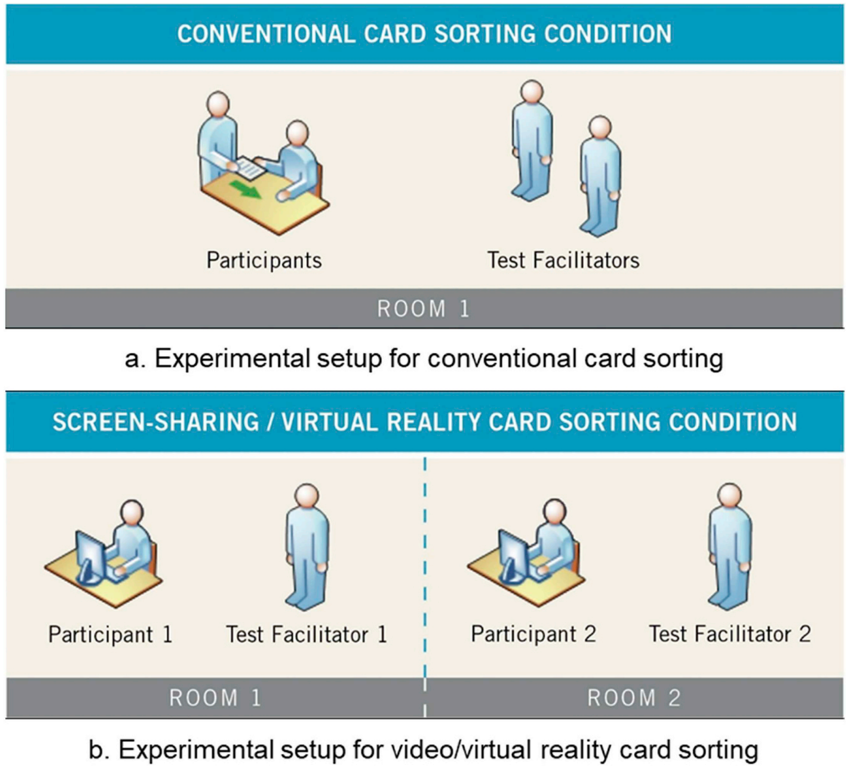

Fig. 2. Experimental setup for study conditions.

spread, across different countries and time zones, conducting traditional IA design studies creates challenges both from the cost and logistical perspective. Remote IA design with conventional computer platforms with users separated over space lacks the immediacy and sense of "presence" desired to support and create group awareness in collaborative IA design process. Other disadvantages include the distractions and interruptions experienced by the participants' in their native environment. The use of immersive three-dimensional (3D) virtual world applications may address some of these concerns. Such 3D applications mirror the collaboration among users when all members are physically present, potentially enabling IA design activities to be conducted more effectively when they are in different places. To address this need, this study compared the effectiveness of distributed synchronous IA design in an immersive 3D virtual meeting room with traditional in-person IA design and an online meeting tool.

2. Method

2.1. Participants

The study included two test facilitators who had previous experience using card sorting methods as well as immersive virtual reality systems. A total of 60 participants with a mean age of 24.03 (SD = 3.45) were recruited to complete the study who were then tested in pairs leaving 30 collaborative groups. These 30 groups were then randomly assigned to one of three different card sorting conditions, resulting in 10 pairs conducting the conventional card sorting method, 10 pairs conducting the video screen-share based card sorting method, and 10 pairs conducting the collaborative immersive virtual reality card sorting method. For the conventional card sorting method, both the participants as well as the two facilitators were in the same room observing the card sorting. Each of the two test facilitators monitored one of the pairs in two different rooms for the video and the immersive virtual reality conditions. The study was approved by Clemson University's Institutional Review Board.

2.2. Hypotheses

H1. In terms of percentage match with master card set, immersive virtual reality-based system is as good as conventional methods of performing collaborative information architecture activities.

H2. Immersive virtual reality-based system is as good as conventional methods of performing collaborative information architecture activities in terms of time taken to complete the task.

H3. Immersive virtual reality-based system is as good as conventional methods of performing collaborative information architecture activities for overall usability.

H4. Immersive virtual reality-based system is as good as conventional methods of performing collaborative information architecture activities for total presence.

H5. Immersive virtual reality-based system is as good as conventional methods of performing collaborative information architecture activities for total workload.

2.3. Experimental conditions and apparatus

The independent variable in this study was the information architecture design environment for card sorting at three levels - conventional card sorting, video-based card sorting, and immersive virtual reality-based card sorting. The cards for the study were selected from Amazon ("Amazon.com: Online Shopping for Electronics, Apparel, Computers," 1994), the online shopping website, which resulted in a closed card sort with pre-existing categories. The participants however, were not provided with this information and were directed to group cards based on their perception. The conventional card sorting method consisted of both participants as well as the test facilitators in one room to perform the card sorting activity (Fig. 2a). This method used a table, index cards with keywords, timer and a camera to take an image of the

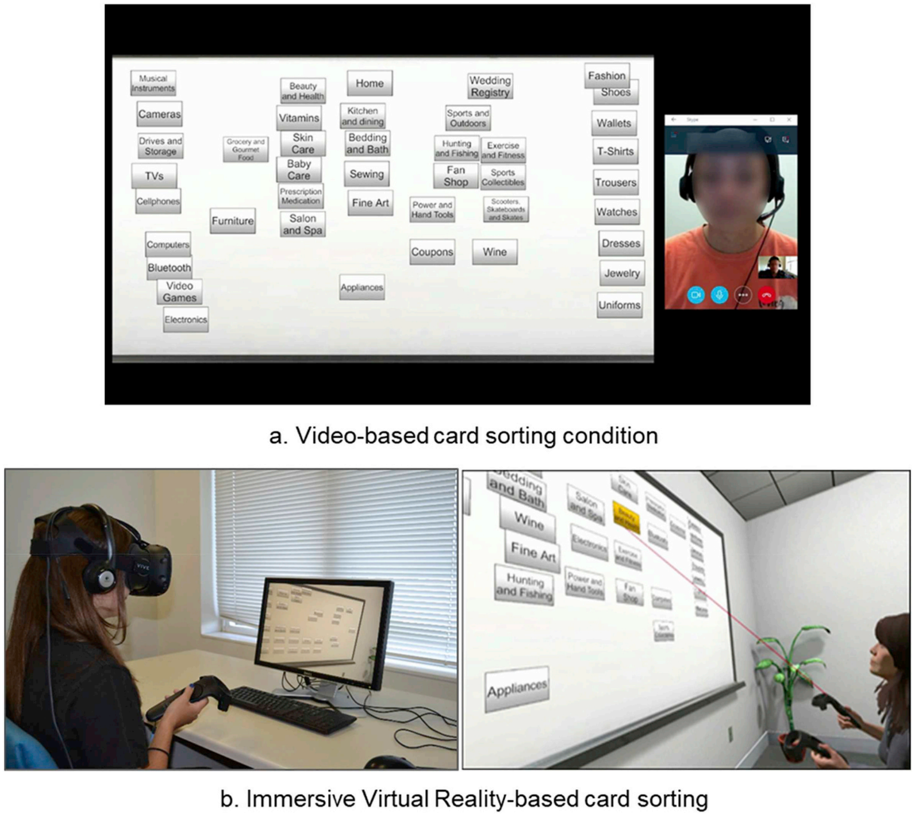

Fig. 3. Video and immersive Virtual Reality based card sorting conditions.

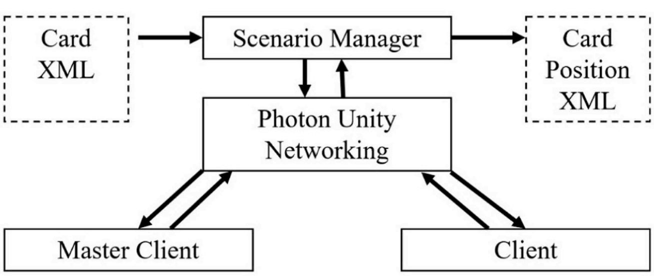

Fig. 4. System architecture view.

final grouping done by the participants.

The video screen-sharing based card sorting was the second condition. The two participants were in two different rooms with a test facilitator in each room monitoring the study (Fig. 2b). In each room, the participants communicated using Dell computers with 17" monitors and Logitech headphones to complete the card sorting task. Participants shared a common screen with the cards displayed on it and could see each other via a Skype window. They used a mouse to manipulate the position of the cards (Fig. 3a).

The immersive virtual reality card sorting method was the third condition. Each of the two participants were in separate rooms with a test facilitator monitoring the study progress (Fig. 2b). The participants in this condition were tasked to complete the card sorting using the HTC Vive Head Mounted Display, HTC Vive controllers ("VIVE," 2016) to move cards in the virtual environment, and Logitech headphones to communicate with each other. To move the cards, participants would lightly touch the trackpad on the Vive controller to activate a laser beam. When the beam intersected a card, they could push the trackpad down like a button to move the card around the whiteboard (Fig. 3b). Participants in the IVR condition could customize their avatars; meaning they could choose between a male and female avatar and could choose between four different skin tones. The avatars employed an Inverse Kinematics (IK) skeletal rig (Kallmann, 2008) for approximating the position of the player's body. The Vive controllers and HMD were used as targets for the IK system while the pelvis and feet targets remained stationary in a seated position. Therefore, only three tracked targets were used to infer the position and orientation of all the bones in the avatar. The position and orientation of each target was synced over the network so that the other participant could see the avatar moving without having to sync the entire skeletal rig.

The IVR and video conditions were both developed in the Unity3D game engine and all scripts were written in Microsoft C#. The avatars were created in Adobe Fuse ("Adobe Fuse CC (Beta)," 2014) and all other 3D models were created in Blender ("Home of the Blender project - Free and Open 3D Creation Software," 1998). See Fig. 4 for an overview of the system architecture. The Scenario Manager started by parsing an XML file containing the names of all of the cards. The Photon Unity Networking engine was used for networking and verbal communication between the two participants. The engine designated the first user as the master client or host of the application for controlling card instantiation. When participants indicated that they were finished, the application was closed and a file containing the final positions of the cards was exported into an XML file and later processed into images for review and analysis.

2.4. Experimental design

The study used a between-subjects experimental design. Time data was recorded using a timer in the conventional card sorting method and recorded by the simulations in the video and immersive virtual reality conditions. The final images of the sorted cards were photographed by the test facilitators in the conventional condition and the card placement was recorded by the simulation software for the video and immersive virtual reality conditions. The participant dyads were randomly assigned to one of the three test conditions. All participants, despite the

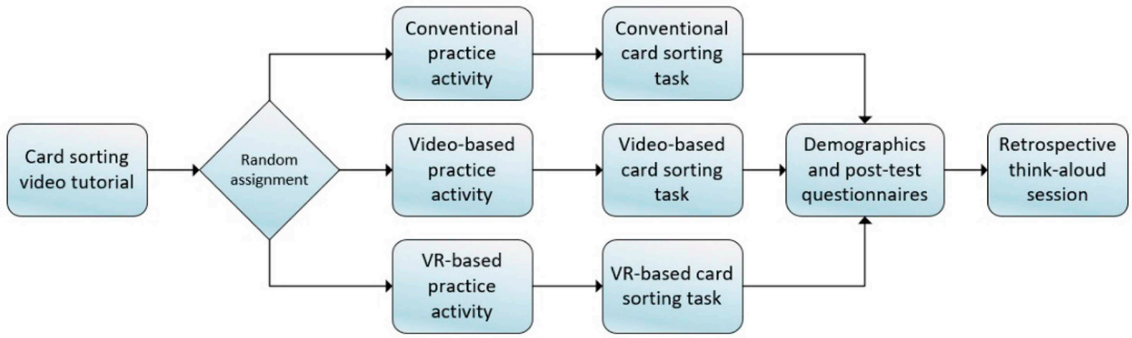

Fig. 5. Study procedure.

condition, were given the same set of cards to sort.

2.5. Procedure

The facilitators and participants followed similar procedures in all three environments as shown in Fig. 5. Each condition involved two participants performing the card sorting activity to simulate the collaborative environment. Two participants were randomly assigned to one of the test conditions and then introduced to one another. The participants were given a brief overview of the study and asked to read and sign an informed consent form. These introductions were followed by a 30-min video tutorial explaining card sorting to the participants.

The participants then performed a practice card sorting activity. The practice activity was carried out by the participants in the same test condition they would eventually use. During the training session, participants were also assisted by the test facilitators to ensure an accurate understanding of how to carry out the task. The participants first completed a practice activity with 30 cards followed by the actual study with 42 cards. During the conventional card sorting study, the facilitators made note of the time required to complete the test and at the end of the study photographed the final grouping to carry out analysis at a later stage. On completion of the test task, participants filled out a demographic questionnaire and a post-test questionnaire. This post-test questionnaire included the Witmer-Singer Presence Questionnaire (Witmer and Singer, 1998) to measure presence in the immersive VR environment, and the IBM-Computer System Usability Questionnaire (IBM-CSUQ) (Lewis, 1995) and NASA-Task Load Index (NASA-TLX) (Hart and Staveland, 1988) to understand system usability and workload as per past research (Narasimha, 2018a,b; Agnisarman et al., 2017; Chalil Madathil et al., 2013). At the end of the study, for all three card sorting conditions, participants were asked to provide their suggestions/concerns in a retrospective think-aloud session that was recorded by the facilitators using an audio recorder. All participants were given a $10 Amazon gift card at the end of the study irrespective of their performance.

2.6. Dependent measures

The three conditions for card sorting were compared using both objective and subjective measures. The objective dependent measures included average time for task completion and percentage agreement of card groupings with the master grouping obtained from Amazon online shopping website. The subjective measures were the ratings from the presence questionnaire, IBM-CSUQ and the NASA-TLX. Qualitative data was obtained through retrospective think-aloud session.

2.7. Analysis

The feasibility of IVR for the collaborative card sorting activity was based on the objective, subjective, and qualitative data obtained. IBM SPSS Statistics 23 was used to analyze the objective and subjective measures. To understand the difference between the three card sorting conditions, an Analysis of Variance (ANOVA) was carried out for percentage correct match of cards and time, and a multilevel linear model with individual participants at level 1 and card sorting condition at level 2 was carried out at 95% confidence interval with post-hoc Least Significant Difference (LSD) comparison for the other subjective measures. The method of Multilevel Linear Model was chosen since this study contained two levels contributing to variance - the individual and the group level.

3. Results

3.1. Demographics

Participants for this study ranged between 19 and 33 years of age (M = 24.03, SD = 3.45) with 33 male and 27 female participants. A summary of the demographics of the participant sample is presented in Table 1. Of the 60 participants tested in this study, 51.70% (n = 31) were Caucasian, 40.00% (n = 24) Asian, 5.00% (n = 3) African American and 3.30% (n = 2) reported as other races. More than five years of computer experience was reported by 96.70% (n = 58) of the participants and 3.30% (n = 2) reported 3–5 years of computer experience. In total, only 35% (n = 21) of the participants indicated having prior experience with immersive virtual reality. Of the 20 participants in the immersive virtual reality condition, only 35% (n = 7) reported prior experience in it. Also, 81.70% (n = 49) of the total participants had no previous experience with card sorting with only 18.30% (n = 11) of the participants having prior knowledge of card sorting.

Table 1
Demographic data.

Variable (n = 60)	Number count	Percentage
Gender		
Male	33	55.00%
Female	27	45.00%
Race		
Caucasian	31	51.70%
Asian	24	40.00%
African American	3	5.00%
Other	2	3.30%
Education level		
High school/GED	2	3.30%
Some College	21	35.00%
2-year college degree	2	3.30%
4-year college degree	15	25.00%
Masters	19	31.70%
Doctoral	1	1.70%

3.2. Objective and subjective measures

To carry out multilevel linear models, the Intra-Class Correlation 1 (ICC1) values were first determined (Table 2) which indicate the amount of nesting i.e., the amount of variability in a metric contributed by the testing condition, with individual members contributing the remaining variability for each metric. Very low ICC1 values (1% or less in this study) indicate that the variability is predominantly due to the individuals and not due to the condition they belonged to. In such cases (involvement score metric from Presence questionnaire, performance and frustration metrics from NASA-TLX), the analysis is a simple linear regression. Apart from this, the objective measures (time, percentage match with master card set) were analyzed using a one-way between-subjects ANOVA. Only those metrics that reached statistical significance are elaborated below with details of others mentioned in Table 3.

The outcome of the usability metrics from the IBM-Computer System Usability Questionnaire (IBM-CSUQ) (Fig. 8) showed that the interface quality metric was significantly different between the card sorting conditions, $F\ (2,27) = 5.78$, $p = 0.008$, multilevel model equivalent[1] of $\eta^2 = 75.42\%$. Interface quality decreased from IVR card sorting ($M = 19.30$, $SD = 2.00$) to conventional card sorting ($M = 16.35$, $SD = 3.44$) condition, and from video-based condition ($M = 18.05$, $SD = 2.37$) to conventional card sorting ($M = 16.35$, $SD = 3.44$). LSD post hoc analysis revealed that the mean decrease in interface quality from IVR to conventional card sorting condition (2.95, 95% CI [1.16, 4.74]) was significant ($p = 0.002$). Further, LSD post hoc comparison also indicated that the mean decrease in interface quality from video to conventional card sorting condition (1.70, 95% CI [-0.09, 3.49]) was marginally significant with $p = 0.06$. Overall usability metric was also statistically significant for the card sorting conditions with $F\ (2,27) = 4.34$, $p = 0.02$, multilevel model equivalent[1] of $\eta^2 = 59.71\%$. Overall usability decreased from immersive virtual reality card sorting ($M = 119.80$, $SD = 8.57$) to conventional card sorting ($M = 109.60$, $SD = 12.80$), and from video-based card sorting ($M = 116.00$, $SD = 9.98$) to conventional card sorting ($M = 109.60$, $SD = 12.80$). LSD post hoc analysis revealed that the mean decrease from IVR to conventional card sorting (10.20, 95% CI [3.02, 17.38]) was significant ($p = 0.007$). The mean decrease from video-based to conventional card sorting (6.4, 95% CI [-0.78, 13.58]) was also marginally significant ($p = 0.08$).

Among the presence questionnaire metrics (Fig. 9), statistically significant results were observed for sensory fidelity $F\ (2,27) = 7.63$, $p = 0.002$, multilevel model equivalent[1] of $\eta^2 = 82.83\%$. A decrease in sensory fidelity was observed from conventional condition ($M = 13.05$, $SD = 1.73$) to immersive virtual reality condition ($M = 9.60$, $SD = 2.74$), and from conventional condition ($M = 13.05$, $SD = 1.73$) to video condition ($M = 10.55$, $SD = 3.64$). LSD post hoc analysis revealed that the mean decrease in sensory fidelity from conventional to IVR condition (3.45, 95% CI [1.58, 5.32]) was significant ($p = 0.001$), as well as the mean decrease from conventional to video-based condition (2.50, 95% CI [0.63, 4.37]) was significant ($p = 0.01$). Total presence score was marginally significant with $F\ (2,27) = 2.95$, $p = 0.07$, multilevel model equivalent[1] of $\eta^2 = 52.42\%$. Total presence decreased from conventional card sorting condition ($M = 121.45$, $SD = 10.54$) to immersive virtual reality card sorting ($M = 113.30$, $SD = 8.11$). LSD post hoc comparison indicated that the mean decrease in total presence from conventional condition to IVR condition (8.15, 95% CI [1.13, 15.18]) was significant ($p = 0.03$).

[1] To estimate effect size in Multi-level Models for a level 2 predictor, a weighted effects coding is used (Cohen et al., 2003) in order to estimate the percent reduction in intercept variance from the null model to the target model.

Table 2
ICC (1) and Cronbach's alpha for IBM-CSUQ, Presence and NASA-TLX.

Subjective Measure	Intra Class Correlation (1)	Cronbach's alpha
IBM-Computer System Usability Questionnaire		
System Usability	35%	0.89
Information Quality	16%	0.73
Interface Quality	19%	0.85
Overall Usability	18%	0.88
Presence Questionnaire		
Involvement score	1%	0.83
Sensory fidelity score	23%	0.70
Immersion score	37%	0.47
Interface quality score	23%	0.31
Total presence score	12%	0.81
NASA-Task Load Index		
Mental demand	13%	–
Physical demand	19%	–
Temporal demand	31%	–
Performance	0%	–
Effort	35%	–
Frustration	0.2%	–
Total workload	22%	0.31

3.3. Qualitative data

A retrospective think-aloud session was carried out for each of the card sorting conditions. Participants were individually asked questions about the advantages, disadvantages and factors that affected performing the activity, and suggestions to improve the collaborative environment. Participants who were assigned to the conventional card sorting method often cited the advantage of being present with another person in the same place, indicating this was especially useful as they could read and understand each other's body language. Some participants said that the physical presence helped them understand the other person's thought process as they could see and move the cards physically and also acknowledged the ease of decision making in this method. When these participants were asked about their thoughts regarding completing the same task using technology as a medium, they said they may be slowed down by the use of controls, and communication may be affected due to the inability to see the other person's body language.

In the video-based card sorting condition, the participants' main observation was the ease of performing the activity on a shared screen noting that the video and audio capabilities were an additional advantage. Participants indicated that the video communication was advantageous, but lacked the ability to portray cues about their teammate such as stress and anxiety with one participant commenting that the interface "got the job done" but was not appealing. Many others indicated that collaboration was mainly facilitated by the audio communication and the video option was only an additional accessory and they did not pay attention to it. Some suggestions related to the system included the need to incorporate more color, the need to see the other person's cursor, and the need to incorporate the video screen with the shared card sorting screen.

The retrospective think-aloud protocol elicited positive reactions about the use of immersive virtual reality to perform this collaborative task remotely. Participants commented that the room really felt like a classroom and was devoid of distractions, which helped them focus on the task at hand. Participants felt it was very easy to learn and that the whiteboard on which they worked was an advantage and they were not limited by the size of the table. They also indicated the ease of coordinating in the virtual room as they could see the laser pointer from each other's controls and were not hindered by the use of a mouse in which case they would not know what their teammate was moving until they did so. The controls used were said to be satisfactory and did not show any lag, which helped them complete the task. Some limitations that were revealed included the need to improve the avatar design and

Table 3
Summary of statistical results of all dependent variables.

Dependent variable	Statistical analysis	Result
Percentage agreement with master card set	ANOVA	F (2,27) = 0.32 p = 0.72
Time	ANOVA	F (2,27) = 1.29 p = 0.29
IBM-Computer System Usability Questionnaire		
System usability	Multilevel linear model	F (2,27) = 1.51 p = 0.24
Information quality	Multilevel linear model	F (2,27) = 2.59 p = 0.09
Interface quality	Multilevel linear model	F (2,27) = 5.78 p = 0.008**
Overall usability	Multilevel linear model	F (2,27) = 4.34 p = 0.02*
Presence questionnaire		
Involvement score	Simple linear regression	F (2,57) = 1.54 p = 0.22
Sensory fidelity	Multilevel linear model	F (2,27) = 7.63 p = 0.002** multilevel model equivalent of η^2 = 82.83%
Immersion score	Multilevel linear model	F (2,27) = 0.19 p = 0.83
Interface quality score	Multilevel linear model	F (2,27) = 1.52 p = 0.24
Total presence score	Multilevel linear model	F (2,27) = 2.95 p = 0.07* Multilevel model equivalent of η^2 = 52.42%
NASA-Task Load Index		
Mental demand	Multilevel linear model	F (2,27) = 0.73 p = 0.49
Physical demand	Multilevel linear model	F (2,27) = 2.21 p = 0.13
Temporal demand	Multilevel linear model	F (2,27) = 1.26 p = 0.30
Effort	Multilevel linear model	F (2,27) = 0.84 p = 0.44
Performance	Simple linear regression	F (2, 57) = 0.10 p = 0.91
Frustration	Simple linear regression	F (2,27) = 0.88 p = 0.43
Total workload	Multilevel linear model	F (2,27) = 0.43 p = 0.66

Significance levels - *p < 0.05, **p < 0.005.

posture, inability to see the corners on the board in the virtual world, and the lack of physical movement were reported as issues. Participants pointed out that the graphics required more work but also that, for the task at hand, there was little emphasis on the graphics. Suggestions were mainly regarding the avatar present in the virtual room with participants indicating that the avatar felt unnatural and not like their partner. Although the avatar gave a sense of being in the room together, participants mentioned that it was devoid of expressions and nuances like body color variations, body hair and skin texture. The lack of similarity between their actual activity partner and the avatar made one particular participant feel like he could "brush him off easily". Participants indicated that a closer replication of their actual team member would make the avatar more effective.

In terms of presence in the virtual room, participants had mixed reactions. While some participants felt that the graphics were good and made them feel like they were in an actual room, others commented on how their placement in the virtual room gave them a sense of belonging. However, another participant felt the room was just a square box and did not provide a good sense of belonging. When asked about the ability to connect with the teammate, participants said that the avatar representation helped and also that it made them feel like they were in the same room, and the audio communication and ability to see the laser pointers was helpful. Although the avatar received multiple comments for improvement, many participants commented on how it helped them feel connected to the other person and the virtual environment (Table 4) provides a few excerpts from the retrospective think-aloud session for the three conditions.

4. Discussion

Although quantitative data showed little statistical significance, several of the outcomes can be rationalized. Percentage match with master card (Fig. 6) set which was not statistically significantly different still indicates that the virtual reality condition did not make the card sorting task more difficult than the conventional method thereby agreeing with our hypothesis. In addition, time data (Fig. 7) showed no significant difference, which, within the scope of this study, is again positive as it indicates that immersive virtual reality condition was not significantly different/worse than conventional and video methods.

For the dependent variable of IBM-Computer System Usability Questionnaire (Fig. 8), interface quality and overall usability was observed to be significantly better for immersive virtual reality card sorting, indicating that IVR fared much better than the other conditions. As seen in Table 4 participants indicated that the shared white screen in the IVR condition presented the cards better and, as indicated by one participant "when you are using the screen, people's hands don't get in the way." This input leads us to believe that the screen used to sort cards in both IVR and video conditions made the task easier, leading to better ratings for interface quality and overall usability in both conditions.

Sensory fidelity, which captures the multimodal sense-related input of information and self-movement within a simulated virtual environment, was higher for the conventional card sorting method. The use of controls, limited haptic feedback, and limited movement in the simulated world may be factors that impacted the observed output. An aggregate value of all metrics in the presence questionnaire– total presence score, was seen to be marginally significantly better for the conventional card sorting condition, which may have been affected in the IVR condition due to participants' expectation to have a more natural avatar or one that more accurately represented their actual team member (Table 4). Statistically, IVR card sorting was not worse than the conventional card sorting condition which is promising, but, the qualitative feedback from the participants suggests that some areas of the simulated environment require additional work to improve presence.

For the metrics under the NASA-Task Load Index (Fig. 10), the lack of statistical significance for physical and temporal demand was expected as the participants were under little to no physical or time-related pressure for the study. We also believe that mental demand, performance and frustration were similar for the three card sorting conditions due to task simplicity. These causes may have also affected the total workload measure which is an aggregation of the other metrics in the NASA-TLX. The lack of statistical significance is again a positive indication that IVR was not significantly different from conventional card sorting in terms of workload.

To summarize the quantitative data results, as per the hypotheses initially developed for this study, data indicated that immersive virtual reality was as good or in a few cases - percentage match with master card set, time taken, overall usability and NASA-TLX, better than conventional method in four variables. Immersive virtual reality,

Table 4
Qualitative data from retrospective think-aloud protocol.

Card sorting condition	Participant comments
Conventional card sorting	"It was simple, everything was right in front of you." "It was easy to go off of each other's body language. You could see what the other person was doing and talk to them in real time to figure out what that person was doing." "Just because you are in proximity, you can see what the other person is doing and quickly adapt." "In person we can like actually tell and explain better than in a virtual world." "Easy to make decision and natural to delegate tasks." "I might be a little bit slower using technology, just navigating my way around it would be difficult to get it going. This was just so easy and organic."
Video card sorting	"It was easy to learn. It didn't take very long to understand how the system worked." "The software was really clear, and pretty easy to do it." "I liked that we had the same exact screen. Whereas, if we were just over Skype, it would have been harder to communicate." "I liked that the cards would turn different colors that was nice. So I could know when he was moving something." "When it's like condensed like this in a smaller screen I can read it faster so I know everything on the screen and then start sorting." "You get less social cues, I couldn't see if he was showing physical signs of stress or anxiety which is harder to know what they're thinking." "When we were picking up the cards maybe it could be useful if we could see the other person's cursor." "It's workable. As in, if you need to get the job done, it's perfect. It could be more appealing." "It would be nice to have the whole chat and interface integrated into one system." "I liked seeing the other person but that wasn't near as much of a factor as hearing the other person." "It wasn't necessary for me to see, as long as the audio was there. I didn't think it was particularly helpful."
Immersive Virtual Reality card sorting	"It really looks like a classroom." "The accessibility is awesome for someone who's never used it. I feel like even my employees who never went to college, I could teach it to them in two minutes." "It was also nice that you were not limited by the table's size, you had like basically a wall. It was virtual cards that were on this huge wall that was moving around. It felt like I had more space." "..the pointers from the laser helped me just point at the cards I was talking about and this made it so much more easier than using like a mouse." "Area in which to work with slightly more freedom to move around would help." "When you look around the graphics are so poor but when you're in it you don't notice it. When you're in it because you're focused on your activity." "I felt like it was animated. Almost like it needs more dimension to it. If you look at my skin it's not just brown it has more colors to it. There's hair and stuff ..." "..because he was like a fake dude, he didn't look like him and so it was easy to brush him off." "A closer replication of her (the partner) would be nice. It would also be nice to see facial reactions to know when you're pissing someone off." "The avatar looked really weird. It didn't look human like at all. It was sort of boxy and just weird." "I'm a super visual person so being able to see a physical representation of him helped." "I naturally want to like look at people when we're trying to make a decision and I felt like he was sitting with me and we were sitting in a room." "It seemed pretty real because I could see the other person. We could interact with each other. I could see my own hands. I could see the room when I turned around."

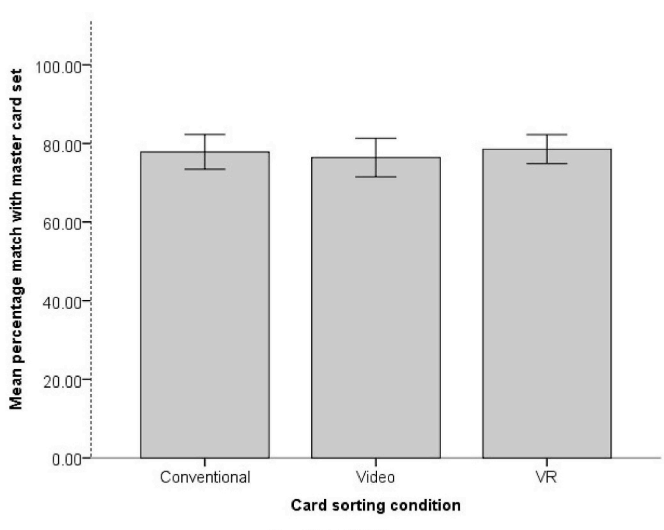

Fig. 6. Mean percentage agreement with master card set.

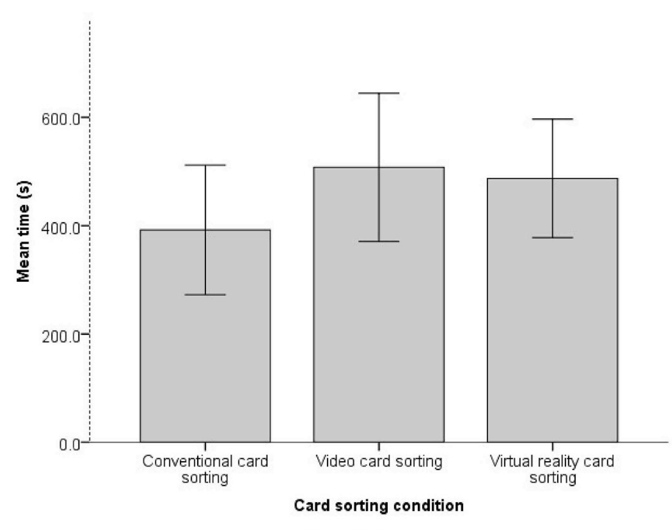

Fig. 7. Mean time taken to complete card sorting.

nevertheless, requires further evidence of performance in two constructs of the presence questionnaire: sensory fidelity and total presence.

Neale et al.'s (2004) Awareness Evaluation Model (see Fig. 1), primarily used to understand collaboration in simulation environments, was used to elaborate on the qualitative results obtained in this study to understand collaboration. According to this model, contextual factors are the basis for all collaborative work upon which the structure of the

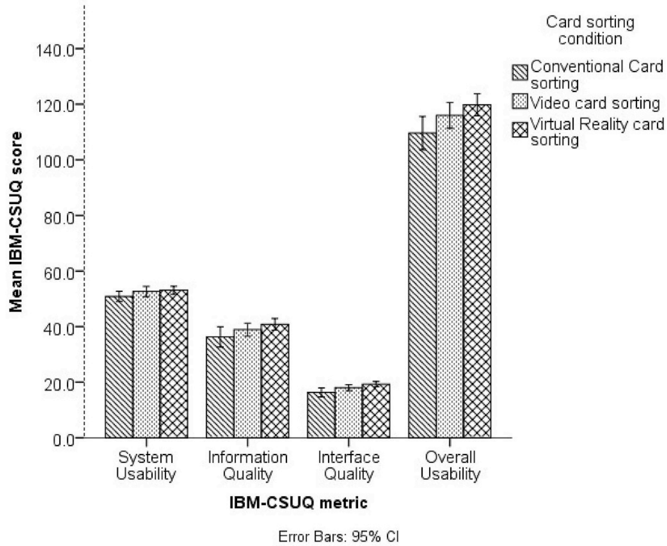

Fig. 8. Mean score for IBM-CSUQ metrics.

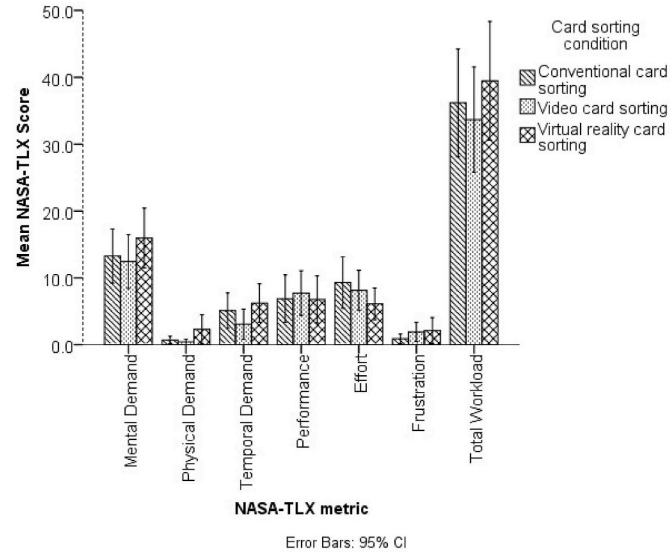

Fig. 10. Mean score for NASA-TLX metrics.

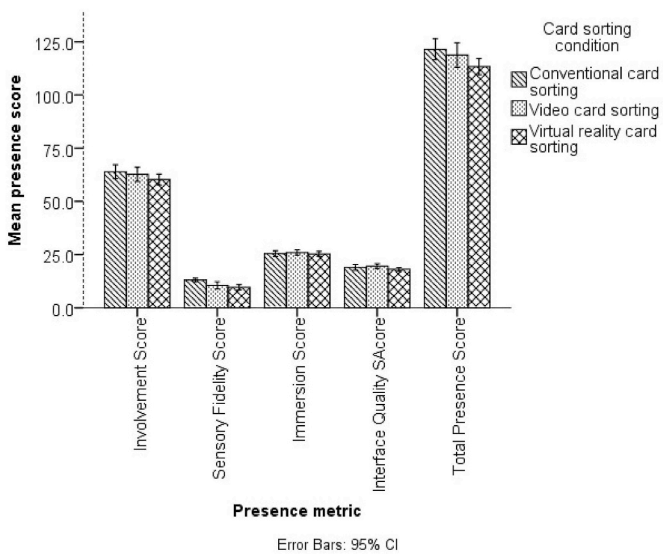

Fig. 9. Mean score for presence metrics.

activity is built. People build context in a specific situation based on their understanding of the task and with information about who is present with them, what they are doing, and how the parties involved are related to one another. This information helps users understand why people perform certain acts that are unexpected. Information leading to context is communicated mostly through lightweight interactions (Neale et al., 2004).

Work coupling, in the context of computer-supported collaborative work, refers to the amount of communication demanded by the task (Borghoff and Schlichter, 2000; Neale et al., 2004; Olson and Olson, 2000). Work coupling is further classified into – loosely coupled work, involving tasks that require less interaction and effortless communication between the participating entities, and tightly coupled tasks which require more communication with highly intertwined tasks dependent on the quality of communication between participants. Distributed process loss occurs when appropriate coordination does not take place. This model identifies five levels of work coupling: light-weight interactions, information sharing, coordination, collaboration, and cooperation (Neale et al., 2004).

To communicate, collaborate and coordinate, people share information. The awareness that teams share as a joint entity is said to be their common ground (Clark, 1996; Neale et al., 2004) and this is essential to carry out any collaborative activity. Team members must continuously update their information to have an updated common ground.

The various factors in this awareness evaluation model are interrelated and affect each other. Depending on the type of work coupling, the demands of communication may change (Neale et al., 2004). Based on this communication, information must be updated on a regular basis to have better common ground. Coordination, the mechanism used to manage collaborative activities, is directly related to information updates and common ground. Finally, the foundation for this common ground shared by the team members largely depends on contextual factors. These contextual factors act as an impetus for the amount of shared awareness that the team members have (Neale et al., 2004).

Having understood the three main concepts of the awareness evaluation model – contextual factors, work coupling and, common ground, the following section now provides a summary of the observations in this study as related to the model. The components of this study relatable to contextual factors in the activity awareness model include the card sorting task, the interaction of the two individuals performing the task, the artifacts in the virtual room, and the circumstances of people's lives outside of the immediate task (Table 5) includes selected quotes taken from the qualitative retrospective interviews concerning context within each of the test environments. Even within virtual environments, individuals can perceive different contextual factors leading to different levels of activity awareness. Context must be taken into consideration for the entirety of the study, as it is an integral part within every group's activity awareness.

Of the five levels of work coupling (qualitative comments in Table 5) defined by Neale et al. (2004), lightweight interaction includes any interaction having little to do with the task at hand. In this study, participants who were previously acquainted exhibited lightweight interaction by talking about their day or making light-hearted comments about the card sorting/organization skills of their partner during the study. Information sharing is the exchange of important background issues related to the task (Neale et al., 2004) as seen when participants shared their personal experiences about card sorting or the organization of the cards. The primary goal of sharing these experiences was to explain their sorting process to their partner.

The third category in work coupling, coordination, is defined by Malone as "the attempt by multiple entities to act in concert in order to achieve a common goal by carrying out a script/plan they all

Table 5
Qualitative data for context and work coupling.

Awareness model component	Card sorting condition	Participant comment
Context		
	Conventional card sorting	"I liked that I was actually there with the person and I could point to things and show them that, and that we could like not have to look at each other and talk."
	Video card sorting	"It didn't really feel like the person wasn't here because of the screen share. It felt like if you and I were sitting here discussing stuff."
		"You can read body language over seeing somebody or being with them. But video doesn't always capture body language like you could in person."
	Immersive Virtual Reality card sorting	"The way the environment was set up helped to establish the sense of this is where I am."
		"The fidelity of detail was not anywhere close to where it is in the real world."
Work coupling		
Coordination	Conventional card sorting	"We discussed which cards we thought might belong together. Then established headers to sort cards into."
		"We made the headers first and then added cards to each one as we went on."
	Video card sorting	"We analyzed the cards for a minute or so and tried to think of categories they could belong to before trying to find header cards to categorize the others."
		"We decided to put the category on top, then we moved all of the cards we thought fit under those categories."
	Immersive Virtual Reality card sorting	"We decided to look through the cards before deciding to move them and talked through the headers before deciding."
		"We picked the headers first and then worked together to sort the items."
Collaboration	Conventional card sorting	"We intuitively decided to divide and conquer. Choosing categories and finding most cards for that one category then communicating what we had done to the other."
	Video card sorting	"We communicated and delegated tasks."
		"We took a moment to soak in the material. Then we took one main word and tried to fill in each category, one at a time."
	Immersive Virtual Reality card sorting	"Each of us would begin by sorting separate groups first. Then we would work together to divide the remaining cards into their respective groups."
		"We scanned the list before sorting began, noticing general similarities between card items. As we began sorting the cards, our criterion for card groups became more specific."
Cooperation	Conventional card sorting	"Our strategy was established by recognizing the broadest categories first. Once we determined our header cards, we placed the sub-cards under their appropriate category based on what we perceived to be the most logical association. Once we had a sensible organization, we did some minor refining until we felt our arrangement was optimized."
		"We discussed what order we would like to take steps in, as well as what decisions should be made during the activity."
	Video card sorting	"We first looked for which cards we thought described categories. After that, we grouped cards based on the categories we selected, then we discussed, edited, and reviewed our selections."
	Immersive virtual reality card sorting	"We really just started picking the things that we felt should be grouped together and then started talking through the things that we weren't sure about. Afterwards we talked about some of the items we picked individually and moved them around."

understand" (Malone and Crowston, 1994). Many groups verbally communicated their strategies before diving into the activity. One group described their coordination as follows: "We discussed which cards we thought might belong together. Then established headers to sort cards into." Collaboration, the fourth category of work coupling, involves group members individually performing tasks to complete the overarching goal of the activity. It was observed that group members deliberately divided tasks and then came together at the end to finalize their results as was observed by a participant's comment about how they decided to divide and conquer by "choosing categories and finding most cards for that one category, then communicating what we had done to the other [team member]." The fifth and highest category in work coupling is cooperation, which demands the greatest amount and highest quality of communication (Neale et al., 2004). It involves shared goals, tasks, plans, and significant consultation with others before proceeding with the task. One participant stated in regards to cooperation: "We first looked for which cards we thought described categories. After that, we grouped cards based on the categories we selected, then we discussed, edited, and reviewed our selections." This group was able to develop a strategy, work individually to accomplish the goal, and used significant consultation to finalize the results of the work. This level of work coupling is often seen in face-to-face interaction as described by Neale et al. but was shown in our study throughout all three conditions (Neale et al., 2004). Table 5 includes quotes obtained from the collaborative questionnaire for three forms of work coupling.

Common ground, the common knowledge team members think they share with each other, is necessary in completing collaborative tasks. The participants were given an introduction to card sorting through a video tutorial, which formed the basis of contextual factors as the participants had enough knowledge about the task and knew their team member knew this. This information, however, had to be updated throughout the study as the participants sorted the cards. The change in the organization of the cards required the need to share and update each other's reasoning to place a card in a particular group. Although, exchanging information verbally contributes to the updating of common ground, non-verbal interactions with the environment can also contribute towards this (Neale et al., 2004). Towards this end, we believe the presence of the avatars within the virtual environment contributed towards the participants' awareness of what their team member was doing and about the progress of the task. While some groups found the presence of an avatar helpful as evident from their comment "When I turned, I could actually see a person, which could help when you're working to remember that that's an actual person."

Apart from the context of the Awareness Evaluation Model, from the qualitative data it was also evident that participants, in general, liked the new immersive virtual reality system. Participants' comments suggested that they liked the use of the virtual system for basic tasks/meetings as in the comments "If it was like in the preliminary stages of stuff just like gathering info or doing drafts or something it would replace all kinds of in person meetings". Another participant mentioned he/she felt more efficient using the virtual reality card sorting because "Personally, moving the cards was a lot quicker because physically moving cards would take more time than just pointing and clicking." Many participants also liked that they could use the avatar as a reference to their team member. One participant commented that the avatars made him/her realize that they were actually working in a team and the avatar served as a reference to his/her team member with another participant appreciating the avatar's ability to mimic his/her movements as “When I looked at something it moved with me and when I moved my hands it was real time.” Although there were comments about improving the avatars to be more real, the lack of expression on the avatar's face also seemed to be an advantage for one participant who mentioned - “Being able to see her [team member's] negative body language might have just added frustration to me when we had a task to do.” These were a few of the many comments we received from participants during our study indicating that overall the participants liked the use of immersive virtual reality.

When compared with the comments related to video-based card sorting and conventional card sorting, participants seemed to prefer the latter with participants mentioning the ease of delegating tasks, being able to see a person and his/her body language, and the ease of moving cards on a table. Many comments also pointed to the close similarities between in-person card sorting and immersive virtual reality such as the presence of an avatar, ability to see body movements, the ability to see the card the other person was referring to/moving and also the feeling of being in the same room (Table 4). The video-based card sorting condition, however, seemed to be less favored. Participants commented that it was frustrating that they could not see the card their partner was referring until they moved the card. One comment indicating this issue is “I would have really liked to be able to see what my partner was selecting, because this would affect what I was selecting”.

In summary, both the quantitative and qualitative data lead us to believe that the virtual reality system was at least as good as the conventional standard for card sorting—the in-person card sorting. Although there were very few quantitative variables showing statistical significance, the important take-away is that the new virtual reality system did not perform worse than the other testing methods. It was even shown to be better than the conventional and video-based conditions for a few variables. Analysis of the qualitative data for the different aspects of the awareness evaluation model (Neale et al., 2004) indicated that the participants' interaction on the virtual reality systems consisted of the important factors mentioned in the model leading us to believe that participants were aware of the task at hand, their team member, and the context in which they worked while coordinating and collaborating with each other similar to how they would in an in-person setting. Further, evidence of a positive reception and appreciation of the virtual reality system was also observed in the qualitative data. Overall, these outcomes suggest that there is evidence to suggest that a sense of presence and the ability to collaborate like in an in-person setting is possible within a virtual environment for collaborative IA tasks which was the question we set out to explore.

4.1. Limitations

This study was performed in a controlled setting with very little geographical separation and good internet connectivity. We had a small sample size of only 30 groups of dyads. This sample size may have affected the possibility of obtaining statistically significant results. In total, the participants sorted 42 cards during this study. The complexity of the task may vary depending on the website/task at hand in a real-world situation. The outcome of the study may also vary depending on the work setting and the hierarchical posts held by the participating entities in the virtual environment.

4.2. Future work and conclusion

Based on the outcomes of this study, the need to carry out further studies with larger groups of people is evident. Studies in the future should also involve true geographic separation to understand if there are other factors that may impact computer-supported collaborative work. For example, internet connectivity and technology availability may be issues requiring further attention. From the qualitative input obtained from the participants, another area requiring immediate attention is the avatars within the virtual world. More effort must be focused towards making the avatars and the virtual environment more relatable to the real world. The avatar design should also focus on incorporating more nuances that may make it more like the actual person with whom the activity is being performed. These factors, as seen from the retrospective think-aloud session, are of importance to the user when using a virtual environment to perform a collaborative task. These changes in the avatar may affect the presence felt by the participants in the collaborative simulated environment.

The study of such computer-supported collaborative work will further narrow the geographic distances felt in performing collaborative design tasks. A focus on the aspects within the virtual environment that affect the collaborative activity may, in time, lead to better systems that provide users a better sense of belonging. Such improvements may increase the scope of immersive virtual environments for remote collaborative tasks. To conclude, the evidence from this study suggests that IVR-based collaborative work could be a viable alternative to in-person collaborative work for geographically separated teams. These promising outcomes lay the foundation for further research investigating the possibility of IVR for collaborative work in areas other than Information Architecture design.

Acknowledgement

This material is based upon work supported by the National Science Foundation under awards #1501695 & #1700621. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apergo.2019.05.009.

References

Adobe Fuse CC (Beta), 2014. March 27. http://www.adobe.com/products/fuse.html, Accessed date: 17 August 2017.

Agnisarman, S., Chalil Madathil, K., Smith, K., Ashok, A., Welch, B., McElligott, J.T., 2017. Lessons learned from the usability assessment of home-based telemedicine systems. Appl. Ergon. 58, 424–434.

Alaraj, A., Lemole, M.G., Finkle, J.H., Yudkowsky, R., Wallace, A., Luciano, C., Charbel, F.T., 2011. Virtual reality training in neurosurgery: review of current status and future applications. Surg. Neurol. Int. 2, 52.

Amazon.com: online shopping for electronics, apparel, computers. July 5. https://www.amazon.com/, Accessed date: 17 August 2017.

Azuma, R.T., 1997. A survey of augmented reality. Presence Teleoperators Virtual Environ. 6 (4), 355–385.

Benford, S., Dourish, P., Rodden, T., 2000. Introduction to the special issue on human-computer interaction and collaborative virtual environments. ACM Trans. Comput. Hum. Interact. 7 (4), 439–441.

Berg, L.P., Vance, J.M., 2017. An industry case study: investigating early design decision making in virtual reality. J. Comput. Inf. Sci. Eng. 17 (1), 011001.

Bertrand, J., Brickler, D., Babu, S., Madathil, K., Zelaya, M., Wang, T., et al., 2015, March. The role of dimensional symmetry on bimanual psychomotor skills education in immersive virtual environments. In: Virtual Reality (VR), 2015 IEEE. IEEE, pp. 3–10.

Bhargava, A., Bertrand, J.W., Gramopadhye, A.K., Madathil, K.C., Babu, S.V., 2018. Evaluating multiple levels of an interaction fidelity continuum on performance and learning in near-field training simulations. IEEE Trans. Vis. Comput. Graph. 24 (4), 1418–1427.

Borghoff, U.M., Schlichter, J.H., 2000. Computer-supported cooperative work. In: Computer-Supported Cooperative Work. Springer, Berlin, Heidelberg, pp. 87–141.

Burdea, G.C., Coiffet, P., 2003. Virtual Reality Technology. John Wiley & Sons.

Chalil Madathil, K., Greenstein, J.S., 2011, May. Synchronous remote usability testing: a new approach facilitated by virtual worlds. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 2225–2234.

Chalil Madathil, K., Greenstein, J.S., 2017. An investigation of the efficacy of collaborative virtual reality systems for moderated remote usability testing. Appl. Ergon. 65, 501–514.

Chalil Madathil, K., Koikkara, R., Obeid, J., Greenstein, J.S., Sanderson, I.C., Fryar, K., Gramopadhye, A.K., 2013. An investigation of the efficacy of electronic consenting interfaces of research permissions management system in a hospital setting. Int. J. Med. Inform. 82 (9), 854–863.

Choi, S., Jung, K., Noh, S.D., 2015. Virtual reality applications in manufacturing industries: past research, present findings, and future directions. Concurr. Eng.: Research and Applications 23 (1), 40–63.

Christian, W., Galambos, P., Csapó, Á.B., Zentay, P., Baranyi, P.Z., 2014. Future internet-based collaboration in factory planning. Acta Polytechnica Hungarica 11 (7), 157–177.

Churchill, E.F., Snowdon, D., 1998. Collaborative virtual environments: an introductory review of issues and systems. Virtual Real. 3 (1), 3–15.

Clark, H.H., 1996. Using Language. Cambridge University Press.

Cohen, J., Cohen, P., West, S.G., Aiken, L.S., Others, 2003. Applied Multiple Regression/correlation Analysis for the Behavioral Sciences, third ed. erlbaum, Mahwah, NJ Retrieved from. https://epdf.tips/applied-multiple-regression-correlation-analysis-for-the-behavioral-sciences.html.

Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., 1993. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques. ACM, New York, NY, USA, pp. 135–142.

Deb, S., Carruth, D.W., Sween, R., Strawderman, L., Garrison, T.M., 2017. Efficacy of virtual reality in pedestrian safety research. Appl. Ergon. 65, 449–460.

De Lucia, A., Francese, R., Passero, I., Tortora, G., 2009/1. Development and evaluation of a virtual campus on Second Life: the case of SecondDMI. Comput. Educ. 52 (1), 220–233.

Dillon, A., 2002. Information architecture in JASIST: just where did we come from? J. Am. Soc. Inf. Sci. Technol. 53 (10), 821–823. Retrieved from. http://arizona.openrepository.com/arizona/handle/10150/105433.

Eppinger, S., Ulrich, K., 2011. Product Design and Development. McGraw-Hill Education.

Firat Ozkan, N., Greenstein, J.S., 2011. Effects of 2D online and 3D virtual world meeting spaces on the performance of a concept selection task by engineering design teams. In: Proceedings of the Human Factors and Ergonomics Society ... Annual Meeting Human Factors and Ergonomics Society. Meeting, vol. 55. pp. 1676–1680 1.

Fleury, C., Férey, N., Vézien, J.M., Bourdot, P., 2015. Remote collaboration across heterogeneous large interactive spaces. In: 2015 IEEE Second VR International Workshop on Collaborative Virtual Environments (3DCVE), pp. 9–10.

Gopinath, A.K., Tucker, C.S., 2015. Investigating the impact of interactive immersive virtual reality environments in enhancing task performance in online engineering design activities. In: ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers pp. V003T04A004–V003T04A004.

Greif, I., 1988. Overview. Computer-Supported Cooperative Work: A Book of Readings. Morgan Kaufmann Publishers, San Mateo, Calif, pp. 5–12.

Gross, T., Stary, C., Totter, A., 2005. User-centered awareness in computer-supported cooperative work-systems: structured embedding of findings from social sciences. Int. J. Hum. Comput. Interact. 18 (3), 323–360.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (task Load index): results of empirical and theoretical research. In: In: Hancock, Peter A., Meshkati, Najmedin (Eds.), Advances in Psychology, vol. 52. pp. 139–183 North-Holland.

Home of the Blender Project - Free and Open 3D Creation Software. . https://www.blender.org/, Accessed date: 17 August 2017.

Howarth, P.A., Finch, M., 1999. The nauseogenicity of two methods of navigating within a virtual environment. Appl. Ergon. 30 (1), 39–45.

Jaspers, M.W.M., 2009. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int. J. Med. Inform. 78 (5), 340–353.

Juárez-Ramírez, R., 2017. User-centered design and adaptive systems: toward improving usability and accessibility. Univers. Access Inf. Soc. 16 (2), 361–363.

Kallmann, M., 2008. Analytical inverse kinematics with body posture control. Comput. Animat. Virtual Worlds 19 (2), 79–91.

Karreman, J., Arendsen, R., van der Geest, T., 2010. Navigating on web sites: exploring the life events approach. In: 2010 IEEE International Professional Comunication Conference, . https://doi.org/10.1109/ipcc.2010.5530028.

Lee, J., Quy, P., Kim, J.I., Kang, L.W., Seo, A., Kim, H., 2009. A collaborative virtual reality environment for molecular biology. In: 2009 International Symposium on Ubiquitous Virtual Reality, pp. 68–71.

Lewis, J.R., 1995. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int. J. Hum. Comput. Interact. 7 (1), 57–78.

Liu, A., Tendick, F., Cleary, K., Kaufmann, C., 2003. A survey of surgical simulation: applications, technology, and education. Presence Teleoperators Virtual Environ. 12 (6), 599–614.

Maguire, M., 2001. Methods to support human-centred design. Int. J. Hum. Comput. Stud. 55 (4), 587–634.

Malone, T.W., Crowston, K., 1994. The interdisciplinary study of coordination. ACM Comput. Surv. 26 (1), 87–119.

Mikropoulos, T.A., Natsis, A., 2011. Educational virtual environments: a ten-year review of empirical research (1999–2009). Comput. Educ. 56 (3), 769–780.

Mills, K.L., 2003. Computer-supported cooperative work (CSCW). Encyclopedia of Library and Information Science 678–684.

Narasimha, S., Agnisarman, S., Chalil Madathil, K., Gramopadhye, A., McElligott, J.T., 2018a. Designing home-based telemedicine systems for the geriatric population: an empirical study. Telemed. J. e Health: The Official Journal of the American Telemedicine Association 24 (2), 94–110.

Narasimha, S., Scharett, E., Madathil, K.C., Bertrand, J., 2018b. WeRSort: preliminary results from a new method of remote collaboration facilitated by fully immersive virtual reality, No. 1. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 62. SAGE Publications, Los Angeles, CA, pp. 2084–2088 Sage CA.

Neale, D.C., Carroll, J.M., Rosson, M.B., 2004. Evaluating computer-supported cooperative work. In: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work - CSCW '04, . https://doi.org/10.1145/1031607.1031626.

Nguyen, H., Pontonnier, C., Hilt, S., Duval, T., Dumont, G., 2017. VR-based operating modes and metaphors for collaborative ergonomic design of industrial workstations. J. Multimodal User Interfac. 11 (1), 97–111.

Nichols, S., 1999. Physical ergonomics of virtual environment use. Appl. Ergon. 30 (1), 79–90.

Nichols, S., Patel, H., 2002. Health and safety implications of virtual reality: a review of empirical evidence. Appl. Ergon. 33 (3), 251–271.

Olson, G.M., Olson, J.S., 2000. Distance matters. Hum. Comput. Interact. 15 (2), 139–178.

Pagliari, C., 2007. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J. Med. Internet Res. 9 (2), e15.

Palmer, T.D., Fields, N.A., 1994. Computer supported cooperative work. Computer 27 (5), 15–17.

Pantelidis, V.S., 1995. Reasons to use virtual reality in education. VR in the Schools. 1. pp. 9 1.

Rosenfeld, L., Morville, P., 2002. Information Architecture for the World Wide Web. O'Reilly Media, Inc.

Rubin, J., Chisnell, D., 2011. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. John Wiley & Sons.

Seymour, N.E., Gallagher, A.G., Roman, S.A., O'Brien, M.K., Bansal, V.K., Andersen, D.K., Satava, R.M., 2002. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann. Surg. 236 (4), 458–463 discussion 463–464.

Sinha, R., 2003. Persona development for information-rich domains. In: CHI '03 Extended Abstracts on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 830–831.

Slater, M., Sanchez-Vives, M.V., 2016. Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI. 3. pp. 74.

Spencer, D., 2009. Card Sorting: Designing Useable Categories. Rosenfeld Media.

Sun, H.-M., Li, S.-P., Zhu, Y.-Q., Hsiao, B., 2015. The effect of user's perceived presence and promotion focus on usability for interacting in virtual environments. Appl. Ergon. 50, 126–132.

Sutherland, I.E., 1965. The ultimate display. Multimedia: from wagner to virtual reality. Retrieved from. https://books.google.com/books?hl=en&lr=&id=adtKCdwzwqIC&oi=fnd&pg=PA232&dq=Sutherland+1965&ots=frLnflJv79&sig=LWTE8_U121ya-BZpdjI2T9FIB1M.

VIVE, 2016. April 5. https://www.vive.com/us/, Accessed date: 17 August 2017.

Vora, J., Nair, S., Gramopadhye, A.K., Duchowski, A.T., Melloy, B.J., Kanki, B., 2002. Using virtual reality technology for aircraft visual inspection training: presence and comparison studies. Appl. Ergon. 33 (6), 559–570.

Wentzel, J., Müller, F., Beerlage-de Jong, N., van Gemert-Pijnen, J., 2016. Card sorting to evaluate the robustness of the information architecture of a protocol website. Int. J. Med. Inform. 86, 71–81.

Wilson, J.R., 1999. Virtual environments applications and applied ergonomics. Appl. Ergon. 30 (1), 3–9.

Witmer, B.G., Singer, M.J., 1998. Measuring presence in virtual environments: a presence questionnaire. Presence Teleoperators Virtual Environ. 7 (3), 225–240.

Zimmerman, D.E., Akerelrea, C., 2002. A group card sorting methodology for developing informational Web sites. In: Proceedings. IEEE International Professional Communication Conference, pp. 437–445.

Zyda, M., 2005. From visual simulation to virtual reality to games. Computer 38 (9), 25–32.