SEEING EYE TO I?

THE INFLUENCE OF SELF-VIDEO DISPLAY SIZE ON VISUAL ATTENTION AND COLLABORATIVE PERFORMANCE IN PEER-TO-PEER VIDEO CHAT

A Thesis Presented to the Faculty of the Graduate School of Cornell University

In Partial Fulfillment of the Requirements for the Degree of Master of Science

by

Kelton Ray Minor

January 2017

ABSTRACT

This thesis examines the influence of self-video size in video chat conversations on visual attention, collaborative performance, grounding, comfort and distraction during a brainstorming task. Twenty pairs of female university students were randomly assigned to either a large or small self-video condition. Two eye tracking systems were used to simultaneously record pairs of participants' gaze across 4 areas-of-interest spanning a 15-minute task. Participants with larger self-video gazed at themselves longer but did not spend a significantly different percentage of the conversation gazing at their partner. Participants sufficiently estimated how long they looked at each other, but significantly overestimated how long they, and their partners, gazed at their own self-video. A majority of participants found their self-video to be comforting, and participants with larger displays found it to be more distracting than those with smaller displays. Over a third of participants would prefer to chat without their self-video visible.

BIOGRAPHICAL SKETCH

Kelton Minor is a graduate student in Human Environment Relations in the Department of Design and Environmental Analysis within the College of Human Ecology at Cornell University. He was previously a US Denmark Fulbright Fellow at the Royal Danish Academy of Fine Arts: School of Design.

This thesis is dedicated to everyone who has ever mistaken pixels for people or a person for a pixel.

ACKNOWLEDGMENTS

While all of my research was physically conducted on the Cornell campus under the patient supervision of Professor Alan Hedge and James Cutting, the majority of my analysis and writing transpired on the other side of the Atlantic in the Psychology Department at København's Universitet and the Co-Design Research Center at the Royal Danish Academy of Fine Arts: School of Design. I would like to sincerely thank all of the participants who took part in this project and conjured up new ideas together, as well as the College of Human Ecology and Department of Design and Environmental Analysis for their generous support of this endeavor. I am especially appreciative of Tracey Sherwood for being my eyes and ears on the ground in Ithaca during my extended leave. Above all, I am thoroughly grateful for the enduring support of my Boulder, Ithaca and Copenhagen friends and families. It is a little too fitting that many of our conversations the last three years were via video chat.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH	III
DEDICATION	IV
ACKNOWLEDGEMENTS	V
TABLE OF CONTENTS	VI
LIST OF FIGURES	VIII
LIST OF TABLES	IX
1. LITERATURE REVIEW	p.1
1.1 Introduction	p.1
1.2 Video Chat 2016	p.2
1.3 Video Chat Before 2016	p.4
1.4 Visual Attention	p.10
1.5 Eye Movements	p.13
1.6 Eye Tracking	p.14
1.7 Attending to Faces	p.16
1.8 Conversational Attention	p.17
1.9 Computer-Supported Cooperative Work Via VC	p.23
2. OBJECTIVES & HYPOTHESES	p.30
2.1 Research Questions	p.30
2.2 Confirmatory Hypotheses	p.30
2.3 Exploratory Hypotheses	p.32
3. METHODS	p.35
3.1 Participants	p.35
3.2 Statistical Power and Sample Size	p.35
3.3 Randomization Procedure	p.36
3.4 Experimental Setup	p.36
3.5 Calibration and Partial Blinding	p.40
3.6 Video Chat Brainstorming Task	p.41
3.7 Visual Attention Measures	p.43
3.8 Data Treatment and Track Analysis	p.45
4. RESULTS	p.50
4.1 Demographics	p.50
4.2 Participant Flow	p.50
4.3 Analysis Procedure	p.50
4.4 Personal Visual Attention	p.51

4.5 Interpersonal Visual Attention p.52
4.6 Collaborative Outcomes p.53
4.7 Environmental Visual Attention p.55
4.8 Retrospective Gaze Awareness p.56
4.9 Self-video Display Evaluation p.58
5. DISCUSSION p.61
5.1 Visual Attention p.61
5.2 Collaborative Outcomes p.63
5.3 Gaze Awareness p.64
5.4 Self-video Display Evaluation p.66
6. LIMITATIONS p.67
6.1 Limitations and Future Research p.67
7. CONCLUSIONS p.72
7.1 Summary of Contributions p.72
7.2 Practical Implications for Cooperative Video Chat Interfaces p.72
8. REFERENCES p.75
APPENDIX A: FIGURES p.87
APPENDIX B: TABLES p.94

LIST OF FIGURES

Figure 1: George Du Maurier's Telephonoscope (Maurier, 1878) p.5
Figure 2: Engelbart and colleagues' NLS collaborative video system (Engelbart et al., 1968) p.6
Figure 3: Picturephone Mod II with VU SELF button (Molnar et al., 1969) p.8
Figure 4: CU-SeeMe for Schools (CU Schools, 1997) p.10
Figure 5: Sample Size vs. Power Calculation p.36
Figure 6: Computer + Eye Tracker Setup p.38
Figure 7: Human Factors Lab Environment p.40
Figure 8: Areas of Interest p.45
Figure 9: Self-video Size vs. Self Directed Gaze p.52
Figure 10: Self-video Size vs. Partner Directed Gaze p.53
Figure 11: Self-video Size vs. Tot. Unique Team Improvements p.54
Figure 12: Self-video Size vs. % Grounded Ideas p.55
Figure A1: Statistical Power Sample Size & Power Specification p.87
Figure A2: Randomization Procedure p.88
Figure A3: Experimental Setup Dell 19" (1280 x 1024) Display Monitor p.88
Figure A4: Skype Video Chat Service p.89
Figure A5: Logitech HD Webcam C270 p.90
Figure A6: Transcribed Version of Online Survey p.90

LIST OF TABLES

Table 1: Saccade and Fixation Classifications	p.14
Table 2: A Comparison of Constraints between VC & FTF Conversation	p.26
Table 3: Participant Flow	p.50
Table 4: Primary Hypothesis Testing Summary	p.59
Table 5: Exploratory Hypothesis Testing Summary	p.60
Table B1: Data Acquisition and Track Analysis	p.94
Table B2: Demographic Data	p.96
Table B3: Normality Tests (Actual and Estimated Gaze)	p.97
Table B4: Heteroscedasticity Tests (Actual and Estimated Gaze)	p.98
Table B5: Self Directed Gaze Tests	p.101
Table B6: Partner Directed Gaze Tests	p.102
Table B7: Collaborative Performance Tests	p.105
Table B8: Common Ground Ratio Tests	p.106
Table B9: Environmental Gaze Tests	p.107
Table B10: Self Gaze Awareness Tests	p.108
Table B11: Participant's Gaze Awareness of Partner's Self-Gaze Tests	p.110
Table B12: Participant's Gaze Awareness of Looking at Partner Tests	p.112
Table B13: Gaze Awareness of Partner Looking at Participant	p.114
Table B14: Participant's Gaze Awareness of Looking at Partner's Room	p.116
Table B15: Participant's Gaze Awareness of Looking Offscreen	p.118
Table B16: Self-video Display Evaluation Tests	p.121

CHAPTER 1

LITERATURE REVIEW

1.1 Introduction

Our faces are communicative interfaces: for the majority of human history we have exchanged verbal and non-verbal information, intentions and attention with each other in person, face-to-face (Tomasello, 2010). Whether through body gestures or vocalizations, people have long coordinated cooperative activity in proximity. The recent advent of video chat, videoconference or video call - all of which describe a face-to-face video conversation held over the internet by means of webcams and dedicated software (Oxford Dictionaries, 2016) - has enabled much of the world's population to circumvent the proxemic requirements of face-to-face interaction to converse with virtual representations of each other across time and space through a visual interface. Today, the global availability and use of video chat applications is on the rise: several segments of society - including student and professional populations - are adopting and deploying a variety of readily accessible services such as Skype, Google Hangouts, GotoMeeting and Apple FaceTime to connect synchronously. Compared to the time and resource costs required for two - or more - remote people to travel to a common destination and engage in in-person meetings, video chat affords near instant face-to-face conversation and cooperative work at a distance and can result in reduced energy consumption and greenhouse gas emissions (Rainie & Zickuhr, 2010; Coroama et al., 2012; Dickson & Bowers, 1973). However, rather than replace travel, video chat can also increase the number of meetings between remote team members (Brancatelli, 1985; Elton 1985; Johansen, 1984). Although online video chat applications provide new avenues for tele-presence and synchronous communication, the design of video chat interfaces may affect how conversation partners allocate personal and interpersonal attention and influence the quality of social interaction. This thesis builds on existing research from the disciplines of human factors, computer mediated communication and visual perception to investigate whether the size of the self-video display in video chat interfaces affects the gaze behavior, gaze awareness, collaborative performance, grounding of new ideas, perceived comfort and distraction among pairs of video chat partners during a cooperative brainstorming task. Even though the visual discontinuities of remote video chat compared to in-person communication are well documented in literature, no formal research exists that experimentally examines whether, and to what extent, self-video display size impacts individual and interpersonal gaze behavior of video chat partners during conversation in general and collaborative work in particular.

1.2 Video Chat in 2016

Despite improvements in global access and use, video chat (VC) communication still lags behind other modes of offline and online communication such as texting and calling in terms of daily use. In 2010 Pew Research found that 95% of American adults made or received at least one call, 72% sent or received at least one text, while only 4% participated in at least one video chat on a daily basis (Lenhart, 2010; Rainie & Zickuhr, 2010). Why is VC used less frequently in comparison to other common modes of communication? Where did VC communication come from? How have contemporary VC applications and interfaces been designed to support communication? How do these systems affect our allocation of personal and interpersonal visual attention during dyadic (two-way) conversations? What social opportunities and constraints does VC afford in everyday life and collaborative work specifically? The following chapter unflattens these questions about the origins and context of video-mediated communication, visual attention, eye movements, eye tracking, facial attention, conversational attention, gaze awareness and computer-supported collaborative work and elucidates the research motives and matters of concern underpinning this thesis.

The communication technologies and settings with which people carry out many of their daily interactions have undergone massive changes in the last decade at home and in the workplace. People of all ages and backgrounds spend an increasing part of their day socializing, negotiating and collaborating through online, virtual communication applications that are accessed by means of personal computers, tablets and mobile devices (Fox & Jones, 2009). As of 2010, approximately three quarters of all American adults used the internet, and of these users, nearly a quarter participated in video chats (Rainie & Zickuhr, 2010). On an average day in 2010, roughly 4% of American Internet users participated in a video call. However, since Rainie and Zickuhr's (2010) study, the integration of video chat into Facebook's profile interface and messenger app has exposed the service to over a billion active global users, the introduction of Google Hangouts into Gmail has offered video chat to over 400 million users and the addition of FaceTime to Apple's operating system has extended video chat to millions more. Taken together, these developments have collectively boosted the population of total video chat users around the world. Video chat communication is also proliferating within the workplace, as distributed employees, remote telecommuters and networked organizations increasingly collaborate and operate across large distances, and video conferences, calls and meetings become a regular fixture of cooperative work (Rainie & Zickuhr, 2010). Presaging Rainie and Zickuhr's more general findings, Turner et al.'s (2009) field study of workplace communication tool use at a small West Coast company surveyed 30 members and found that median video chat frequency increased from 'never' to 'less than once a month' between May 2008 and 2009, but that employees used video chat much less than face-to-face (daily), phone (daily), email (daily), text chat (weekly) and social networking sites (weekly). Even though the survey was limited to a single workplace in Silicon Valley, the rise in video chat adoption was echoed by Rainie & Zickuhr's (2010) more general survey (n= 3001 participants) that found daily video chat use by adults increased twofold from Spring 2009 to Summer 2010. However, it should be noted that this reported increase still fell within the Pew survey's +/- 3% margin of error. Further confirming the growth of video chat in the workplace, a 2010 industry survey of communication trends of international business professionals (n = 1800) found that 44% of American respondents reported using video conferencing on a frequent basis, 90% of enterprise employees spend some time working off site, 30% spend a quarter to half of their time out of the office and 91% reported brainstorming as the primary focus of most meetings ("How do we communicate," 2010). However, the validity of this survey is questionable given the potential vested interests of the company that conducted it. Additional impartial research is needed to confirm the reliability of the survey's findings.

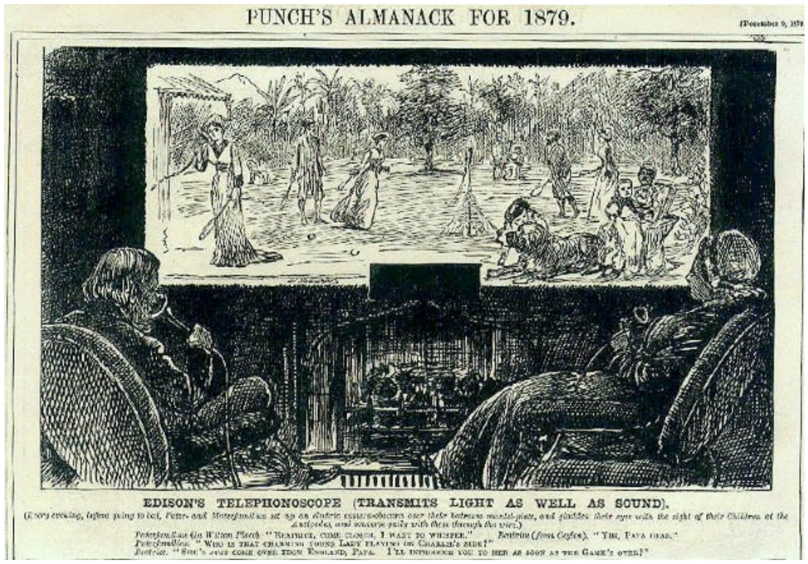
Despite the pervasiveness of video chat applications, current video conferencing interfaces differ from face-to-face communication in several specific ways. Video chat interfaces contain gaze misalignment issues that originate from the spatial separation of the recording camera used to collect video of a user and the imaging screen used to depict it to their partner (Bekaert et al., 2008; Grayson & Monk, 2003; Chen, 2002; Vertegaal, 1999; Agius & Angelides, 1997). Hence, it is impossible for two participants using widespread video chat systems to simultaneously look at their partner's eyes on the screen and look directly into the camera to make eye contact and establish mutual gaze. Popular video chat interfaces also deviate from face-to-face conversation through the inclusion of a self-video display that provides each user with a live mirrored-video feed of themselves below the video of their conversation partner. Thus, when two video chat partners attempt to gaze at the depicted image of one and another in their respective video call windows, they see their corresponding partner gazing downwards along with a miniature, live portrait of themselves, possibly impacting conversation quality.

1.3 Video Chat Before 2016

Over the past century video conferencing has advanced from an obscure newspaper rumor to a common cooperative activity, addressing several - but not all - barriers to synchronous communication along the way. In 1877 the New York Sun published a letter to the editor stating that an acclaimed but unnamed scientist in the city was preparing to unveil the electroscope: an invention “by means of which objects or persons standing or moving in any part of the world may be instantaneously seen anywhere and by anybody” and which might “supersede in a very short time the ordinary methods of telegraphic and telephonic communication.” The anonymous letter included descriptions of how the impending technology would generate new ways for interacting with remote people, environments and objects: “Mothers, husbands and lovers will be enabled to glance at any time at the very persons of their absent children, wives or beloved ones.” Furthermore, it also described how “a combination of the electroscope and the telephone will be made which will permit people, not only actually to converse with each other, no matter how far they are apart, but also to look into each other’s eyes, and watch their every mien, expression, gesture and motion while in the electroscope” (Electrician, 1877). The promise of remote video calling and long distance eye contact was born, sparking a social and scientific project to visually communicate over distance that continues today. The vision of this possible future was first depicted by French cartoonist George du Maurier for the British-produced Punch Magazine (see Figure 1), who portrayed the Telephonoscope as a fictional Edison invention that transmits imagery and audio between two locations - in this particular case parents lounging in the comfort of their home and their daughter in the midst of a badminton match (Maurier, 1878). Shortly afterwards, Scientific American published a description of George Carey’s proposed invention that could transmit images between remote locations using selenium cameras that could allow people to ‘see by electricity’ (Carey, 1880).

Figure 1

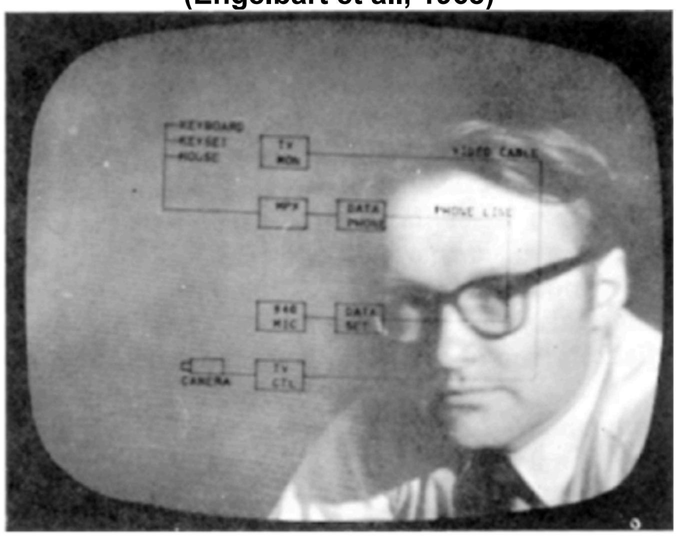
George Du Maurier’s Telephonoscope (Maurier, 1878)



After entering the lexicon and social imagination, French science fiction writer Albert Robida portrayed the possible future of the impending 20th century and described Le téléphonoscope as an encompassing elliptical screen that enables audiovisual communication between an operator and their partner in his work *Le Vingtième siècle. La vie électrique* (Albert, 1893). Writing for Strand magazine in 1898, Arhur Mee spoke of remarkable things to come: “If, as it is said to be not unlikely in the near future, the principle of sight is applied to the telephone as well as that of sound, earth will be in truth a paradise, and distance will lose its enchantment by being abolished altogether” (Mee, 1898).

Following fiction and sparked by research on its predecessors - telegraphy, radio, telephony, television and computation - video chat surfaced against the backdrop of the telecommunication revolution of the late 19th and early 20th centuries. In the late 1920s, researchers at Bell laboratories and administrators at AT&T headquarters in New York City communicated with each other over telephone with an accompanying television feed. A subsequent version of this two way video phone was demonstrated at the Chicago World Fair in 1933-34 (Noll, 1992). In the period spanning from 1936 - 1940, the German Post Office (Reichspostzentralamt) set up a video telephone network across Berlin and other cities consisting of two closed-circuit television systems directly connected via coax cable or remotely via radio (Peters, 1938; Kristiansen, 1991). As post WWII reconstruction began in many of the same German cities, back in the US many of the wartime innovations in telecommunications and computation fueled research and development into how to construct and apply computers to ameliorate human capacities. One such experiment - the oN-Line System (NLS) - became the first operational collaborative video system to involve more than one computer (Engelbart et al., 1968). Assembled by Douglas Engelbart and fellow researchers in the 1960s, the project provided the first working demo of teleconferencing that depicted each participant's face on their partner's display and introduced the concept of screen sharing to 'augment human intellect' and cooperative work (see Figure 2).

Figure 2
Engelbart and colleagues' NLS collaborative video system (Engelbart et al., 1968)



Two-way video communication aimed at the mass market first materialized in the form of video phones, including AT&T's Picturephone. In 1964, attendees of the World's Fair in Queens, New York were able to hold a ten-minute visual conversation using Bell's Mod I Picturephone with a stranger using another Picturephone at a similar installation at Disneyland in California. Given the camera's limited field of view, participants had to pose in a static posture to stay visible for their partner (Fagen et al., 2010). AT&T released its commercial version of the Mod I Picturephone shortly thereafter with a highly promoted call between Lady Bird Johnson and a Picturephone booth in NYC. Early user research by AT&T found that people reacted to this novel mode of communication in unexpected ways:

> "Most people when first confronted with Picturephone seem to imagine that they will use it mainly to display objects or written matter, or they are very much concerned with how they will appear on the screen of the called party. These reactions are only natural, but they also indicate how difficult it is to predict the way people will respond to something new and different.
>
> Those of us who have had the good fortune to use Picturephone regularly in our daily communications find that although it is useful for displaying objects or written matter, its chief value is the face-to-face mode of communication it makes possible. Once the novelty wears off and one can use Picturephone without being self-conscious, he senses in his conversation an enhanced feeling of proximity and intimacy with the other party" (Molnar et al., 1969).

A 1965 case installation at Union Carbide's New York and Chicago offices also revealed that employees initially primped themselves before using the videophone (Schnaars & Wymbs, 2003). The subsequent version of the Picturephone (Mod II) attempted to address these self-conscious concerns with a wider display screen so that participants could see if they were in the camera's range, a toggle switch that allowed a user to swap the video feed of their partner with their own video feed to see how they appeared to their partner (Figure 3), a privacy mode that prevented their partner from seeing their video, a brightness toggle that allowed people to customize how bright their image appeared, an audio dial to adjust volume and a dedicated on/off switch to end the conversation (Schnaars & Wymbs, 2004; Noll, 1992; Diehl, 1969; Diehl, 1972). Based on user feedback from several years of testing the Mod I, Bell scientists also moved the camera from the left of the screen to a new position directly above it:

> "The camera is placed above the picture tube to make the eye contact angle as small as possible. This is significant because, while the camera is looking at the subject, the subject is looking at the picture tube. The apparent "looking away" is annoying to the viewer unless the angle is small. The least annoyance occurs when the subject appears to be looking slightly down which is frequently the case in normal conversation. Locating the camera just above the picture tube creates this effect (Molnar et al., 1969)."

Figure 3

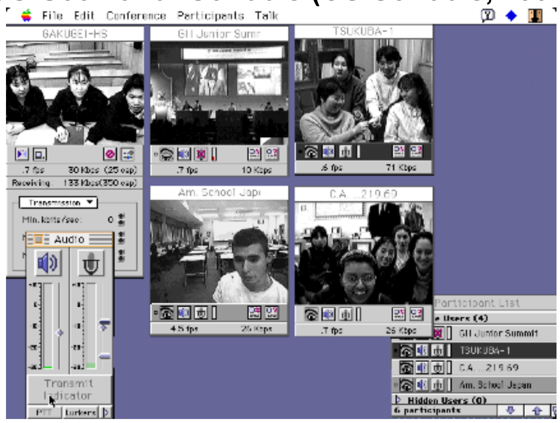
Picturephone Mod II with VU SELF button (Molnar et al., 1969)

The ‘VU Self’ button can be seen as the earliest predecessor to the self-video display, and was added to allow video chat participants to “check their position” and center themselves on the camera before reverting back to their partner’s video feed (Molnar et al., 1969). Despite these changes and it’s promise to “displace today’s means of communication” and “make many of today’s trips unnecessary” the Picturephone and its successors suffered from deficiencies in transmitted image resolution, video compression and image rate (Noll, 1992; Egido, 1988). Even though this new mode of communication claimed to enable intimate conversations with other parties, there were few to be had: mass adoption of video communication technologies remained elusive due to expensive equipment costs, limited interoperability between providers and a lack of a network effect (Schnaars & Wymbs, 2004). Furthermore, previous employees, including AT&T's corporate historian, have also speculated that a primary factor for its failure was that people wanted to see others but not be seen (Noll, 1992; Guernsey, 2000). After multiple failed attempts to reintroduce the video phone to the public and hundreds of millions of dollars in R&D funding, AT&T halted the project by 1974 (Schnaars & Wymbs, 2004; Noll, 1992). In the the 1980s, several Japanese companies including Mitsubishi, Sony and Panasonic attempted to introduce cheaper video phones that worked over existing phone lines but with a fraction of the picture quality (Hawkins, 1988). These entrants were also greeted with muted demand and were discontinued shortly thereafter (Schnaars & Wymbs, 2004). Video conferencing technology targeting the workplace followed a similar trajectory as the Picturephone and also fell short of bullish, prior expectations. By the mid 1980s there were only an estimated 210 videoconferencing installations across 75 large companies (Tyson, 1987). Today's ubiquitous self-video display that appears simultaneously - as opposed to sequentially - alongside a participant's video feed of their partner evolved directly from these early video phone technologies.

Since the 1990s, the rise of the internet and advancements in internet technology have spurred public and private access to video chat. Between 1992 and 1995, a group from Cornell University built and distributed CU-SeeMe the first publicly accessible 'free' video chat software that utilized the web to transport video - and later audio as well - over IP (Dorcey, 1995; Sattler, 1995). Differing from the prior video telephony solutions, Cu-SeeMe and its Unix predecessors Inria and Vic, enabled people to interact with coinciding video streams of both their partner(s) and themselves (see Dorcey, 1995; McCanne et al., 1995; Turletti, 1996). These first self-video displays appeared in windows that were slightly smaller or nearly equivalent in visual extent to the participant's video display of their partner (CU Schools, 1997.; Figure 4). The CU-SeeMe interface gave users extensive control of their self-video and allowed them to adjust their self-video display's visibility (size and location on the screen), orientation (regular or mirrored) and picture quality (brightness and contrast) (Sattler, 1995).

Figure 4

CU-SeeMe for Schools (CU Schools, 1997)



Video and voice over IP (VoIP) rendered the requirement of dedicated lines obsolete, facilitated low-high resolution video transmission over the internet within bandwidth constraints and reduced lag, dropped calls and spotty connections. Coupled with the increasing prevalence of computers with embedded web cams and growing access to free online video chat software, video chat has since suffused through much of society in the early 21st century, finding new use cases along the way. Among other applications, video chat has been adopted and adapted to replace physical travel to remote destinations (Biello, 2009; Denstadli, 2004), to connect leading medical researchers with practicing doctors operating in different locations (Greer et al., 2016), to create portals between astronauts hovering above the earth and the stargazers beneath them (Trotta, 2012), to enable sign language-based telecommunications between deaf and hard-of-hearing people (Fitzgerald, 2003), and to facilitate cooperative work for everyone from young children (Ballagas et al., 2009) to world leaders (Stolberg, 2009). Despite these strides forward, the effects of this medium on individual and interpersonal behavior in general and attention in particular is still largely unknown.

1.4 Visual Attention

When two people video chat, each partner moves their eyes to attend to visual features of the scene located at different positions both on and off their respective screen. The next three sections review predominant theories of attention, the taxonomy of eye movements involved in visual and conversational attention, as well as a brief history of the eye tracking technology researchers employ to record and analyze the ocular movements we cannot see - without.

We are enveloped by vast amounts of stimuli that appeal to our senses but possess only limited information processing capacities to distill what matters. Attention helps us to direct our cognitive resources to integrate small pieces of salient stimuli and construct the complex scenes we are encompassed by. People move their eyes to orient regions of our visual field along a central 'foveal' axis of gaze where the human visual system's sensitivity to fine detail is highest (Duchowski, 2007). Most often, we allocate our attention along this same axis of gaze. Many magicians, film editors and gorillas have exploited this coupling of attention and central gaze to guide their audience's attention to or away from areas of interest (Macknik et al., 2008; Smith & Henderson, 2008; Levin & Simons, 2000; Simons & Chabris, 1999). Indeed, psychologist William James' prescient observation that attention "implies withdrawal from some things in order to deal effectively with others" (James, 1890; James, 1981) still resonates today. But before we direct our focus elsewhere, let's fixate on the origin of attention itself.

Contemporary views of visual attention are predicated on many aspects of earlier theories. Luminaries like Von Helmholtz observed early on that our visual attention is constrained to a small region of our visual field and continuously 'roams' to attend to new areas of interest. Notably, Helmholtz also recognized that we can concentrate on peripheral objects with or without making overt eye movements towards them, but that doing so is necessary when we wish to inspect them in higher detail (Helmholtz, 1925). William James also acknowledged that attention can be willingly and reflexively directed, but favored the view that as we age we increasingly direct our sensorial and intellectual attention to stimuli based on our voluntary expectations and "some remote interest the effort will serve" (James 1984). This dichotomy of visual attention as something which can be actively or passively directed under different conditions was elaborated by Gibson, who demonstrated the extent to which preconceptions - intentions to react in specific ways - can bias how people respond to a stimulus (Gibson, 1941). Over a decade later, Broadbent demonstrated that even when sensory information enters in parallel, it is selectively filtered by sensory channels that possess limited capacities (Broadbent, 1958). Conversely, Deutsch and Deutsch (1963) conjectured that all sensory information is processed at a higher level by central structures in the brain that weight relatively important information over other information. In her unpublished doctoral thesis, Anne Treisman proposed a theory that unified aspects of the explanations contributed by both Broadbent and Deutsch and Deutsch. She described two stages of attention. First, incumbent information passes through an attenuation filter that selectively reduces the strength of certain messages. Second, they are sent to central ‘dictionary units’ that process these attenuated and unattenuated messages on the basis of contextual relevance and importance (Treisman, 1961).

Even though these early theories began to characterize attention, they failed to explain how people are able to conjure a mental picture of the entire scene in front of them despite only being able to attend to ‘narrow’ regions at any one time. Inspired by Dodge’s (1907) early investigation of eye movements at the beginning of the 20th century and empowered by new recording technology, eye movement studies and visualizations began to surface in the mid to late 1960s and illuminate the rapid, sequential nature of human gaze behavior. In Yarbus' (1967) study, participants responded to several different question-based tasks related to the scene depicted in an image. The recorded eye movements showed serial viewing patterns between different portions of the image that were contingent on the question being tasked. Early diagrams of eye movements were at once expressive and frenetic, seemingly contradicting the smooth phenomenology of the human visual experience (Yarbus, 1967). After recording people’s eye movement patterns in response to different images, Noton and Stark (1971a) described these serial patterns as “scan paths.” Notably, they observed that even without leading questions people exposed to an identical image of a square tended to direct their foveal vision to common informative details of the stimulus (the corners) yet exhibited vastly variable scan paths both between and within participants on separate trials (Noton and Stark, 1971b). While these studies highlighted the centrality of foveal vision in attention, Posner et al. (1980) later proposed the concept of a ‘spotlight’ of attention that moves around a scene - independent of foveal vision - and orients where people should look. Under their paradigm, orientation is a covert, mental process and precedes overt detection via gaze. Shortly thereafter, Treisman again advanced an integrative theory that described attention as a bonding agent that stitches together multiple attended features within a scene to resolve an object as a whole (Treisman & Gelade, 1980). Feature integration theory put forth the idea that attention makes use of a feature map to swiftly encode basic information about the scene specifying where feature boundaries are located as well as the size, color, orientation and parallax distance of the various features (Treisman, 1986). This map does not identify what the features are, only where they are located. In summary, visual attention can be seen as a selective perceptual process of voluntarily and involuntarily registering and responding to stimuli in the field of view. Attention is oriented by a combination of low-level features of interest in the peripheral field of view as well as the cognitive processes - tuned by a person’s experience, expectations and the demands summoned by the task at hand - that drive their intent to attend to various features. Finally, early eye tracking methods played a role in tracing the temporal response of people's foveal direction of gaze and provided a new way of visualizing overt visual attention.

1.5 Eye Movements

In order to attend to the changing demands around us in our intermeshed physical and virtual environments, we rely on a small class of eye movements to physically reposition the fovea: vergence movements, saccades, fixations, vestibular movements and physiological nystagmus. To look at other people, screens and objects located varying distances away, a viewer's eyes make vergence movements - their left and right eyes converge or diverge - to help accommodate the object at the necessary depth of field. To gaze at a new location and change the focus of overt attention, people make saccades - voluntary or reflexive movements that reposition the fovea. Temporally, saccadic movements are swift, taking between 10 - 100ms to transition from one position to the next. A gaze velocity threshold of 130 deg/s is often used for saccade detection (Duchowski, 2007). Aside from the sheer speed of each saccadic leap, viewers don't perceive these jerky movements - especially over larger jumps - due to saccadic suppression, which acts to attenuate visual input during saccades (Riggs et al., 1974). Once initiated, saccadic movements are ballistic and may not change course from the intended destination since there appears to be inadequate time for the eye to process and respond to visual feedback mid-movement (Carpenter, 1977). To track a moving target in their environment or on a screen, a viewer can initiate smooth pursuit movements to follow the moving object across their visual field assuming the speed of the target is below the human visual system's upper tracking limit (Leigh & Zee, 1991). To fix our gaze on an object or area of interest, people may initiate fixations - stabilization movements that cause the retina and fovea to hover around the target (Martinez-Conde, 2004). Fixations typically last from 150 ms - 600 ms and are comprised of several smaller movements: micro-saccades, drift and tremor (Yarbus, 1967; Carpenter, 1977; Irwin, 1992). These miniature eye movements cause the area of fixation to fluctuate up to 5 degrees of visual angle (Carpenter, 1977). Micro-saccades are characterized by spatially varying jumps over 1 - 2 minutes of arc that prevent images from becoming completely static on the retina. The significance of this functional noise during fixations cannot be understated: artificially stabilized images on the retina cause the observed image to fade away to a blank state within a few seconds (Riggs et al. 1953; Coppola et al., 1996). Finally, nystagmus is a type of eye movement that attempts to compensate for movement of the head (vestibular nystagmus) or movement of a target across the retina (optokinetic nystagmus) (Carpenter, 1977). Of these eye movements, saccades, fixations and smooth pursuit movements lend themselves to the study of overt visual attention (Duchowski, 2007).

Table 1
Saccade and Fixation Classifications

Source Literature	Saccade Duration	Fixation Duration
Irwin (1992)	10 - 100 ms	150 - 600 ms
Snowden et al. (2012)	25 - 30 ms	200 - 300 ms

1.6 Eye Tracking

Much of our knowledge of eye movements has arisen from the development and deployment of eye tracking technology. Eye movement technologies may measure orientation of the eye in three dimensional space or the eye position relative to a user's head and fall into four primary categories: electrooculography, scleral contact lens, photo or video-oculography and video-based pupil and corneal reflection (Young & Sheena, 1975). This brief review will focus on the evolution of corneal reflection methods given the pervasive application of video-based pupil and corneal reflection technology in performing point of regard analysis for graphical displays and interfaces.

From a historic perspective - with a notable exception - eye measurement techniques have followed a trajectory from highly invasive hardware methods demanding physical contact with the eyes and manually recording results to more remote and unobtrusive technologies that rely on image processing software to enable expedient analyses. In 1898, a highly invasive method involving a plaster of Paris ring mechanically adhered to the cornea and linked to recording pens was used to measure eye movements (Young and Sheena, 1975). A corneal reflection technique for collecting eye measurements was first reported at the turn of the 20th century, but this early method apparently suffered from low accuracy (Robinson, 1968). In the 1950s, contact lens-based solutions were advanced to affix various measurement devices (mirrors, wire coils, etc.) directly to the eye ball allowing for precise measurement of small eye movements at the cost of comfort, mobility and occasionally visibility for the participants wearing them (Duchowski, 2007). In the late 1960s and 1970s, electrooculography emerged as a novel method for recording eye movements by measuring the skin's electric potential levels and differences from electrodes located around both eyes. Meanwhile, photo-oculography and video oculography allowed eye movements to be inferred by measuring distinct features of the eyes (apparent pupil shape, position of the iris-sclera edge and/or corneal reflections of a nearby infrared light source). Besides the physical and ocular discomfort associated with having physical wires and hardware directly on the eyes (scleral contact lens/search coil), around the them (electro-oculography) or in front of them (photooculography or video-oculography), all of these methods can only measure eye movements relative to head position and thus require the head to be fixed in place with a head rest, chin rest, bite bar and/or the use of a separate head tracker to compute the point of gaze in the surrounding environment (Duchowski, 2007).

Given that the human visual system involves head movements and body movements to reposition the field of view about areas of interest (Gibson, 1979), video-based pupil and corneal reflection approaches can accommodate head movements and eye rotations simultaneously and allow gaze-based interactions to be recorded with interactive and graphical displays. By continuously recording both the pupil center and corneal (Purkinje) reflection of a hidden infra-red light source with cameras and image processing software, a participant's point of gaze (also known as point of regard) can be recorded remotely in real time. Since the location of the Purkinje reflection is constant due to the fixed position of the external light source, precise eye movements can be measured by recording the deviations of the pupil center relative to the corneal reflection (Crane, 1994). Table based or head mounted video-based pupil and corneal reflection systems allow for more naturalistic studies of ecologically valid environmental and mobile tasks where the instruments themselves don't become a primary point of discomfort or distraction from the task being studied. Remote corneal reflection eye tracking systems support investigations of personal point of gaze behaviors across external displays and interfaces; however, a single tracker alone is insufficient to study the interpersonal gaze behavior exhibited in social video chats involving at least two separate conversation partners. To record such a social task, each video chat partner needs to be paired with their own identical eye tracking device, calibrated independently and recorded simultaneously throughout the duration of the video conversation.

Today, researchers utilizing eye tracking as a behavioral analysis tool are primarily concerned with how, when, where and why we move our eyes under different visual paradigms (Snowden et al., 2012). Eye tracking technology and analytics tools allow researchers to precisely track a participant's gaze and thus gain insight into a participant's visual attention over the course of a given activity (Duchowski, 2007). A recent survey of the breadth of eye tracking research confirms the wide reach and application for evaluating human computer interactions both on and off screen (Duchowski, 2002).

1.7 Attending to Faces

Faces are among the most biologically and socially salient stimuli we encounter in everyday life and face-to-face conversations in particular (Palermo & Rhodes, 2007). From birth, newborn infants will follow faces with their gaze further into the periphery of their visual field than faces that are scrambled, although this behavior diminishes following the first month of life (Farroni et al., 2002; Johnson et al., 1991; Morton & Johnson, 1991). Older infants also exhibit preferential attention when exposed to more realistic faces (Maurer & Barrera, 1981). The human brain contains a functionally specialized region within the fusiform gyrus that selectively engages when one visually attends to faces but not for other objects (Puce et al., 1995; Puce et al., 1996; Allison et al., 1994). Faces may also draw people's attention to them more than other objects. Ro et al. (2001) exposed participants to a flickering display depicting changing images of faces and five common objects and found that people detected changes to faces significantly faster and more accurately than changes to other objects. Given Simons' (2000) previous finding that people are more likely to notice changes to a visual object or location when they are directly attending to it, faces may recruit superior attentional resources compared to other competing stimuli (Palermo & Rhodes, 2007). Finally, facial detection may be quicker than object detection. Electrophysiological 'ERP' research indicates faces are processed faster, taking around 100ms to signal a response compared to 200ms required for object or word categorization (Pegna et al., 2004).

Beyond the bottom-up salience of human faces, our personal faces are even more effective at captivating our attention compared to other stimuli. Past research by Tong & Nakayama (1999) found that people can detect their own familiar faces significantly faster than unfamiliar faces, even when exposed to an unfamiliar face hundreds of times. Meanwhile, Barton et al. (2006), found that face orientation (regular, upside down) and face familiarity (images of strangers vs. images of famous faces) affected eye gaze behavior, with participants scanning the lower face less when faces were inverted compared to right side up and scanning the upper face less when faces were familiar as opposed to unfamiliar. Dyadic video conferencing presents two sets of visible faces for each participant to attend to, raising the question of whether, and to what extent, a participant's own face competes with their partner's face for their attentional resources, especially when one's partner is unfamiliar. Furthermore, while researchers have posited that people's discomfort with being on camera is one of the primary factors influencing video chat's low adoption when compared to other common modes of communication, it remains unclear if this is due to people's possible discomfort with their partner seeing them and/or to people's own potential aversion to seeing themselves on the self-video display (Guernsey, 2000). Anecdotal evidence suggests that some video chat users may even alter their face through surgery and/or other less invasive means to improve their appearance on video chat (Considine, 2012). A plastic surgeon interviewed by the New York Times in 2012 indicated that "a quarter of the 100 face-lift patients he has a year cited the way they look on webcams as a reason for going under the knife" and that his new procedure "reduced sagging necks but did not leave a scar under the skin - where the camera usually points - as traditional neck-lifts do." Notably, the president of the American Society of Plastic Surgeons confirmed that other plastic surgeons had encountered similar concerns from their patients (Considine, 2012). Additional experimental research is needed to assess the degree of perceived comfort and/or distraction afforded by self-video displays generally as well as Self Video Displays of differing size.

1.8 Conversational Attention

Eye contact during conversation can provide speaking partners with information about their partner's attention and identity, stimulate regions of the brain involved in communication, cue partners to take turns speaking and listening and support task goals (Senju et al., 2009; Kleinke, 1986). In contrast to other primates, humans have developed a depigmented (white) sclera which increases the contrast of the eyes from the surrounding facial skin and may ease the inference of gaze direction. Some researchers have hypothesized that this represents an evolutionary attribute to aid eye contact and facilitate gaze-based social interaction (Kobayashi & Kohshima, 1997). The high luminance contrast between the iris and sclera has also been found to convey gaze direction (Ando, 2002). Evidence from brain imaging studies indicate that infants exhibit cortical activation in response to direct eye contact and can distinguish between direct and averted gaze by the age of four months (Farroni et al., 2002; Farroni et al., 2004). Multiple laboratory studies found that observers were faster at detecting faces exhibiting direct gaze than those with averted eyes when tasked with judging whether the target face was present or absent among a modified series of identical 'distractor' faces with different gaze directions (Senju et al., 2005; Senju et al., 2008). Past research by Senju & Hasegawa (2005) found that it takes longer to detect peripheral targets when fixating on a centrally located face exhibiting direct gaze towards the viewer than when a face exhibits averted gaze. Moreover, multiple studies suggest averted gaze drives attention towards the direction of gaze (Frischen et al., 2007). Taken together, a partner's apparent downward averted gaze during video chat may drive participants' attention towards elements near the bottom of their respective screens, including their self-video displays.

Direct gaze appears to also facilitate identity encoding and gender identification (Hood et al., 2003; Macrae et al., 2002). Several studies suggest that people who gaze directly at their partner on a video recording rate each other more favorably than those who only share a small amount of face-directed gaze (Kleinke et al., 1974; Naiman & Breed, 1974). Cook & Smith found that research participants perceived same sex peers that continuously gazed at them to be more pleasant and less nervous than those who did not gaze at them at all (Cook & Smith, 1975). People that direct greater amounts of gaze at their partner in face-to-face interactions are rated as more attentive than those that gaze at their partners less (Kleinke et al., 1973; Kleinke et al., 1975). Similarly, interviewees that exhibit greater face-directed gaze towards their interviewers are consistently evaluated more positively than when they engage in low levels of gaze (Forbes & Jackson, 1980; Kelly, 1978; Kleinke et al., 1975; Sodikoff et al., 1974). Indeed, the duration of an interviewee's eye contact with an interviewer has been found to be positively correlated with observers' estimates of the interviewee's intelligence (Wheeler et al., 1979).

A recent review of eye contact by Senju et al. (2009) referred to perceived direct gaze as the 'eye contact effect,' whereby eye contact affects neural activity simultaneously and immediately following cognitive processing and behavioral responses. A constellation of neuroimaging studies collectively show perceived eye contact enhances the activation of a network of neural structures involved in social interaction (Adolphs, 2009; Nummenmaa & Calder, 2009) compared to averted gaze including: the fusiform gyrus (Calder et al., 2002; George et al., 2001; Pageler et al., 2003), anterior (Calder et al., 2002; Wicker et al., 2003) and posterior (Conty et al., 2007; Pelphrey et al., 2004; Schilbach et al., 2006) regions of the temporal sulcus, amygdala (Wicker et al., 2003; Kawashima et al., 1999; Sato et al., 2004) as well as the medial prefrontal (Calder et al., 2002; Conty et al., 2007; Schibach et al., 2006; Kampe et al., 2003) and orbitofrontal (Wicker et al., 2003, Conty et al., 2007) cortex. While many of these regions are activated by tasks involving faces and eye gaze, eye gaze has also been found to enhance activation of the intraparietal sulcus, a region that does not typically respond selectively to faces but that is associated with spatial perception and spatially directed attention (Hoffman, 2000). Similar to the stimuli encountered in video chat, most of these studies were conducted with image or video exposures of faces exhibiting averted or direct gaze, but additional research is needed to determine how the task demands of video chat and the simultaneous exposure to two sets of faces exhibiting averted gaze engage this so-called social brain network (Adolphs, 2009; Brothers, 2002).

Past observational studies have also found that gaze is frequently used alongside other nonverbal cues to regulate conversational turn-taking. Face-to-face conversation partners have been found to begin and end an utterance with a partner-directed gaze (Levine and Sutton Smith, 1973; Kendon, 1967). Early observations by Argyle & Cook (1976) and Duncan & Fiske (1977) also suggested that people gaze more at their partner(s) while listening than while speaking; however, these studies relied on observer reports of gaze and the results may be contingent on specific conversational tasks and prior experience of the participants rather than serving as generalizable features of most conversation. Indeed, subsequent studies by Ellyson et al. (1981 & 1980) found that participants discussing issues relevant to their expertise exhibited equivalent amounts of partner directed gaze when speaking as when listening. It is now understood that the role of gaze in turn-taking is highly variable given the context of the conversation, the participants backgrounds and their underlying motives (Nakano, 2003; Beattie, 1979; Lazzerini et al. 1978).

For the past 130 years, inventors have proposed creating a mode of remote communication that allows people to directly gaze into their partner's eyes to approach the richness of face-to-face communication (Electrician, 1877). Despite advancements in video chat technology, current video chat interfaces prevent people from making direct eye contact due to the spatial separation of the video chat camera and screen (Gemmel, 2000; Vertegaal, 1999). Compared to the original Picturephones that were intentionally designed to minimize this parallax through the close coupling of the camera and a small (5" x 5.5") screen, the personal and desktop computers used today have larger screens on average and a greater parallax between the embedded (or external) camera and the portion of the screen where a participant's partner is displayed (Stokes, 1969). Early research at Bell Labs reported that Picturephone users perceived that their partners were gazing directly at them when the 'eye contact angle' between the user, the camera, and their partner's depicted face was less than 5 degrees (Stokes, 1969). Research investigating the role of gaze alignment in video conferencing systems, found that reducing the eye contact angle - the parallax between the camera position and a partner's screen-based video - helped participants to perceive non-verbal signals more so than without gaze correction (Suwita et al., 1997; Mühlbach et al., 1995; Acker & Levitt, 1987). Acker & Levitt (1987), found that participants assigned to eye-contact aligned video chat involving half-silvered mirrors reported significantly increased satisfaction and comfort with the level of eye contact but not with the outcome of a negotiation-based decision made with their partner as compared to participants using a video chat configuration with some parallax between the camera. However, Acker & Levitt (1987) refrained from sharing the extent of this parallax angle, limiting the utility and repeatability of their findings. A later study, also using a half-silvered mirror to overlay a screen and camera, found that participants in the condition with the larger eye contact angle between the camera and their depicted partner (10 degrees of vertical visual angle and 8 degrees of horizontal visual angle) found these configurations reduced apparent eye contact compared to the condition with a zero-degree eye contact angle (Mühlbach et al., 1995). However, the eye contact angles tested by Mühlbach et al. (1995) did not appear to influence satisfaction. These results were confirmed in a subsequent study by Suwita et al. (1997) who added one additional eye contact angle condition where the camera was constantly repositioned behind the half-silvered screen to shadow the position of the depicted participant's head and thus maintain a zero-degree eye contact angle throughout. Even though participants reported perceiving when their partner was looking at them significantly more in both of the zero-degree eye contact conditions compared to the the larger eye contact angle with vertical and horizontal parallax, there was no significant difference between the two zero degree conditions and the condition with the 8 degree horizontal eye contact angle (Suwita et al., 1997). However, both of these studies retrospectively asked participants to indicate whether they were aware of being looked at by reporting whether or not they agreed with directional statements such as: "I never got the feeling of being looked at," which may have led participants to bias their assessments. Bekkering (2004) exposed participants to email messages, voice messages and video messages by an unknown partner taken from three different camera angles: 20 degrees above the partner's central line of gaze; 20 degrees to the side of their gaze and 0 degrees directly facing their gaze. The sender was trusted less when depicted from above or to the side compared to the 0 degree video. Voicemail also generated significantly higher trust perceptions than the videos with gaze parallax between the camera and the onlooker's direction of gaze, suggesting that audio can carry equivalent information influencing perceived trust (Bekkering, 2004; Bekkering & Shim, 2006). Although significant, the effect sizes for these differences were quite small and it is not clear whether a similar trend would occur in bidirectional video communications. The parallax angles selected for testing in these early eye contact studies were largely arbitrary and did not consider smaller and off-axis intervals.

Under most circumstances, people cannot distinguish whether another person is gazing directly at their eyes or at their surrounding face but can discern when gaze is averted away from their face (Argyle & Cook, 1976; Harper et al. 1978; Frischen et al., 2007). Eye contact is not exclusively experienced along a fine line emanating from an observed person's central axis of gaze, but rather occurs within the intersection of a narrow cone surrounding the central axis of gaze that varies in relation to several characteristics of the observer, perceiver and the medium in between them (Gamer & Hecht, 2007). Recent research suggests that observers can still perceive eye contact when the onlooker gazes slightly below the observer's line of sight. Chen (2002) found that from a distance of 2.4m, participants became aware of their partner's deviations in eye contact that traveled beyond 1 degree above, to the left or to the right of their eyes, but only perceived an equivalent break in eye contact when their partner gazed 5 degrees below the participant's eyes in face-to-face viewing conditions. Compared to the face-to-face condition, observers in the videoconferencing group were even less sensitive to eye contact deviations below the axis of the web camera, with observers beginning to perceive a loss in eye contact when their partner looked 8 degrees below the camera. Chen (2002) hypothesizes that this may have been influenced by the low image resolution of the depicted video feed and may represent an inherent bias towards perceiving eye contact in situations where direct gaze is uncertain such as videoconferencing. However, it is equally possible that this represents an affirmative reporting bias linked to the verbal protocol used by the experimenter to ask whether the observer perceived direct eye contact. Another possibility is that participants may have learned to infer a partner's face-directed gaze differently through previous video chat experience. Grayson & Monk (2003) showed that even though true mutual gaze cannot be achieved with conventional, low cost video chat systems, it is possible for users to learn to infer their partner's gaze direction effectively enough to perceive when their partner is looking at the participant ranging from 55% when the video window is place to the far left side of the screen opposite from the camera all of the way to 87% when the video window is located directly beneath the center camera. However, they only tested whether participants could discern when their partner was looking at them in situations where the video window was reduced to a medium size and located near the top of their partner's screen. This deviates from common video chat setups in at least three ways. First, people commonly video chat with their partner's video occupying the full screen, centered in the middle or placed lower down. Second, most widespread video conferencing software includes a self-video display at the bottom of the screen which was notably absent in all of Grayson & Monk's (2003) experimental conditions. Third, as the authors note, their setup did not present participants with a realistic video chat task.

To correct this gaze angle, researchers have proposed utilizing half silvered mirrors to position a camera directly behind an image array (Acker & Levitt, 1987; Mühlbach et al., 1995), embedding a pinhole camera into a screen (Rose & Clarke, 1995), using software to warp images of a participant so they appear to be viewing their partner (Gemmell et al., 2000), invoking avatars to replace people with virtual syndicates (Ohya et al., 1993) utilizing small screens that minimize the parallax with the camera (Molnar et al., 1969) and/or placing the screen and camera at a greater distance away from the viewer to achieve the same effect (Sellen, 1995; Hunter, 1980; Acker & Levitt, 1987). Vertegaal et al. (1999) recommended that future video chat systems focus on conveying nonverbal cues through relaying gaze direction via integrated eye tracking system and visible metaphors such as video windows that appear to shift their orientation on the basis of where participants are looking. However, eye tracking systems still remain cumbersome, expensive and absent from widespread video chat software and hardware. However, most of these proposed systems are cost prohibitive, burdensome to setup or result in disturbing facial distortions. So far, none of these solutions are evident in widespread video conferencing applications. The proliferation of smaller mobile devices with smaller parallaxes between the camera and video screen may help people transmit and receive direct gaze compared to their larger desktop ancestors. However, as Bekkering and Shim (2006) recently noted, people often position phones closer to the face which may increase parallax relative to the viewer and negate the advantages of the smaller screen size. In summary, most desktop and laptop video chat configurations provide some visual cues of a partner's gaze, but can create the false appearance that video chat participants are looking away from each other when they are actually making virtual eye contact with their partner's image on their screen.

Compared to in-the-moment gaze awareness - discriminating where a person is looking at a specific instance - retrospective gaze awareness describes discriminating where and/or how long people looked after the task is complete so as to not disrupt or detract from the natural interaction being examined. Guan et al. (2006) found that people can retrospectively provide an accurate account of the sequence of what they attended to during computer-based tasks while watching a video of their previous task interactions. However, people's verbal AOIs did not correspond to 38% of eye movement AOIs for simple tasks and 56% for complex tasks, indicating that people omit many AOIs during this verbal protocol, possibly due to forgetting where they previously directed their attention and/or failing to verbally match the ballistic nature and rapid rate of their earlier eye movements. A subsequent study by Johansen and Hansen (2016) tasked participants with exploring a web page while their eye movements were recorded and then asked them to repeat their eye movements again on the same website immediately following their first viewing. Participants fixated 70% of the website's same elements, but revisited certain elements more than others ranging from 77% (photos) to 30% (logo element). However, participants' overt gaze recollection may have been limited by the act of trying to cognitively recall their previous eye movements.

There are fewer studies looking at people's ability to retrospectively self-report overall gaze duration for screen-based elements following computer-based tasks. Previous research by Albert and Tedesco (2010) used eye tracking to confirm that participants exposed to web pages for short (7 second) durations reliably estimated which elements they did see compared to which ones they did not see as well as which elements they gazed at for 'a significant amount of time' compared to elements that they spent 'little or no time' looking at. However, this research used a simple five-point scale ranging from 'did not see' to 'definitely saw' to gauge a participant's awareness and did not evaluate the reliability of people's estimated gaze duration compared to their recorded gaze duration along a continuous interval registered by an eye tracking apparatus. Additionally, the research primed participants ahead of time by telling them in advance that they would be asked about what they saw on each of the web pages. As the authors note, additional research should investigate realistic task-driven behaviors, investigate people's awareness of gaze directed to other salient visual elements, including images of faces, and examine gaze awareness following longer tasks (Albert & Tedesco, 2010). Furthermore, Kleinke et al. (1986) posited that conversational attention research should expand to measure each participant's awareness of their own gaze as well as the gaze that they receive from others.

1.9 Computer-Supported Cooperative Work Via Video Conferencing

With the advent of distributed and highly connected national and international teams both within and between organizations, more people are engaging in computer-supported collaborative activities over video chat (Rainie & Zickuhr, 2010; Turner et al., 2009; Agius & Angelides, 1997; "How do we communicate," 2010). When two (or more) people convene to brainstorm ideas to improve a situation, every participant offers a combination of distinct and overlapping knowledge, expertise, skills and perspectives. One-way people in a conversation establish mutual ideas, knowledge and assumptions is through grounding - a mechanism first described by Clark (1989) and later elaborated by Clark and Brennan (1991). Clark and Brennan's grounding in communication theory (1991) describes common ground as mutual knowledge, mutual beliefs and mutual assumptions necessary to coordinate the content and process of a conversation between two people. During communication common ground is constantly updated moment by moment through a process Clark & Brennan call grounding (1991). Grounding is shaped by both "purpose - what the two people are trying to accomplish in their communication" and "the medium of communication - the techniques available in the medium for accomplishing that purpose and what it costs to use them." Continued attention through eye gaze is one of the forms of evidence that each conversation member provides to - and acquires from - their partner to ensure that they are understood (Clark and Brennan, 1991). For instance, when a person looks away from their partner in face to face communication their partner may feel unattended to and lose interest or seek to regain their attention (Goodwin, 1981).

Aside from the purpose of a conversation, the medium through which two people connect also can affect the specialized techniques and collaborative effort required to understand each other sufficiently for the task. Different mediums and their associated interface(s) can color participants' collaboration, impose unique communicative constraints on their exchanges and afford particular behaviors. Early studies on video mediated communication in the workplace sought to understand the differences between face-to-face communication and new collaborative channels. In a classic series of human factors studies, Chapanis et al. (1972; 1977) found that pairs of participants were able to problem solve just as well with voice communication as with face-to-face communication, but that conversation pairs that could see each other used more words than pairs that were limited to just voice communication. Shortly after the first iteration of this study, Ochsman and Chapanis (1974) expanded their analysis to ten different communication channels and three different cooperative tasks and found no difference in the collaborative outcomes of solution time, behavioral activity and linguistic measures between pairs communicating through video and audio as compared to audio alone, concluding that audio carries more influence on collaborative outcomes than video. These early studies are notable both for their ambition as well as their susceptibility to multiple testing without adequate statistical controls to limit test-wise false positives (Ochsman and Chapanis's 1974 analysis of 10 different communication modes only appointed two teams (n=4 participants) to each of the 30 task by communication mode combinations and carried out 198 comparisons. Assuming a significance level of alpha = .05 for each comparison, there is a 990% chance Ochsman and Chapanis's reported results include at least one false positive without even accounting for the insufficiently small sample size for each condition. However, an influential review of video conferencing applications found that most studies that followed Ochsman and Chapanis's (1974) work have confirmed that video-mediated communication provides little to no improvement over audio-mediated communication in terms of performance for many different cooperative tasks (Sellen, 1995). In Short et al's (1976) study, people subjectively rated video communication as providing a greater sense of presence than multi speaker audio but still less than face-to-face interaction. Compared to audio alone, video transmits more non-verbal cues including some general components of gaze, gestures, body movements, mouth movements involved in speech, facial expressions, posture, proximity and appearance but shares many redundant verbal cues. Teoh et al. (2010) found that participants video chatting with a wider angle camera and unrestricted view of their partner gave higher ratings of social presence compared to those in the setting with a traditional, restricted camera view limited to the head and upper torso. They also identified the need for more research that seeks to compare different variations of the same mode in terms of quantitative as well as qualitative outcomes.

Dyadic brainstorming describes a type of conversation task between two people with the purpose of generating ideas and sharing knowledge centered around one or more goals. To compare the constraints and opportunities afforded by video chat relative to face-to-face conversation during a brainstorming task, the table below (Table 2) was adapted from Clarke and Brennan's (1991) original list of medium constraints on grounding. The table has been updated for this comparison by breaking visibility into two parts: visibility of a participant's partner and visibility of the participant's own body.

Table 2
A Comparison of Constraints between Video Chat and Face-to-Face Conversation Mediums

Medium Constraints	Video Chat	Face-to-Face Conversation
Copresence	No* • Does not allow conversation partner A & B to occupy the same physical environment • Partners view a cropped window of their partner's surroundings • Can hear their partner but only see what their partner's upper body is doing • Cannot see what their partner is looking at	Yes • Partners share the same physical environment • Can see each other's surroundings • Can see and hear each other • Can see what each other is doing • Can see what their partner is looking at
Visibility (Partner)	Yes* • Participants can see a 2D cropped view of their partner's head and upper body that may appear to be located close or far away from the participant depending on both participants' cameras (field of view, embedded angles) and each participants proximity to their camera. • Apparent gaze direction differs from actual gaze direction • Insufficient bandwidth and other networking problems can occasionally cause the video to freeze or lag behind a partner's utterances	Yes • Participants can typically see all or most of their partner's head and body in their field of view, in full stereo, with the visual angle determined by their proximity to their partner. • Participants can directly observe a partner's apparent gaze direction

Table 2
A Comparison of Constraints between Video Chat and Face-to-Face Conversation Mediums

Medium Constraints	Video Chat	Face-to-Face Conversation
Visibility (Self)	Yes* • Participants can see a cropped and mirrored view of their own head, upper body and immediate background from the perspective of the web camera • Participants can see their own upper and lower body by tilting their head	No* • Participants cannot typically see their own head and neck without the use of a reflective surface, mirror or camera • Participants can see their own upper and lower body by tilting their head
Audibility	Yes* • Participants can hear a compressed, stereo version of each other and take note of timing and intonation • Highly reverberant environments can occasionally cause echo and create audible feedback loops that persist • Insufficient bandwidth can also cause audio delay	Yes • Participants can hear each other and take note of timing and intonation. • Reverberant environments can create echoes that diminish
Cotemporality	Yes* • Partner B receives a signal at roughly the same time as partner A produces • Depending on the quality of the connection between partners, some mild delay can occur	Yes • Partner B receives a signal at roughly the same time as partner A produces • There is no audible delay

Table 2
A Comparison of Constraints between Video Chat and Face-to-Face Conversation Mediums

Medium Constraints	Video Chat	Face-to-Face Conversation
Simultaneity	Yes • Partner A & B can send and receive a communication simultaneously • Delays can occur	Yes • Partner A & B can send and receive a communication simultaneously
Sequentiality	Yes • Typically, partner A and B's turns cannot get out of sequence • If severe delay is present, the turns between partners may be punctuated by other activities as they wait for the connection to restore so they can hear each other's utterances.	Yes • Partner A and B's turns cannot get out of sequence
Reviewability	No* • Participants cannot review each other's video and verbal communications after they are uttered unless they were sent using text-based chat through the video chat interface during the video conference	No • Participants cannot review the utterances spoken by their partner

Table 2
A Comparison of Constraints between Video Chat and Face-to-Face Conversation Mediums

Medium Constraints	**Video Chat**	**Face-to-Face Conversation**
Revisability	No* • Participants cannot revise video based communications after initiating them • Participants can verbally repair or re-phrase their previous statement	No* • Participants cannot revise spoken communications after uttering them • Participants can verbally repair or re-phrase their previous statement

Compared to in-person, face-to-face conversation, video chat supports but impresses new constraints on the visibility, audibility, cotemporality, simultaneity and sequentiality of dyadic conversation. Notably, VC lacks the copresence provided by face-to-face interaction in which partners can share the same surroundings and monitor each other's actions and where they are gazing. Conversely, VC affords each participant the capacity to visually attend to their own body, head and face from the perspective of their web camera in the midst of conversation with their partner. Since speech fades quickly, both VC and FTF conversation mediums don't enable participants to review and revise each other's' communications after the conversation.

CHAPTER 2

OBJECTIVES & HYPOTHESES

2.1 Research Questions

Video chat is a medium of communication that enables people to extend their gaze beyond their immediate personal surroundings to attend to and communicate with a remote partner as well as with themselves. Given that mirrors are seldom used in face-to-face conversation, each partner's self-video display (SVD) may impose unique communication constraints on interpersonal conversations over video chat. Unfortunately, no previous research has examined whether the salience of a participant's SVD influences their interpersonal visual attention, gaze awareness, collaborative performance, grounding of new ideas, perceived comfort and distraction during a video chat. Based on the literature reviewed, this thesis aims to address the following questions:

1) Does the size of a participant's SVD affect their actual and estimated allocation of visual attention towards their own SVD and their partner's image during a collaborative video chat?
2) Does the size of participants' SVDs influence their collaborative performance and grounding during remote, cooperative work?
3) Does SVD size affect how comforting and/or how distracting participants find their SVDs?
4) How likely are participants to prefer video chat with an SVD versus without an SVD?

2.2 Confirmatory Hypotheses

Of all of the visual stimuli that humans encounter during their lives, faces are among the most salient (Palermo & Rhodes, 2007). As opposed to dyadic face-to-face chats, in virtual video chats we contend with four sets of faces on the Skype interface: our face, our self(ie) video display, our partner and our partner's self-video display. Only one of these is directly visible to us during face-to-face conversation (our partner's face), but two of them are immediately visible during video chat. New methods are needed to measure and analyze the simultaneous gaze behavior of two video chat partners to understand how knowledge and the presence of these four faces influence participants' interpersonal and self-directed gaze. Given the paradox that a video chat participant's direct gaze towards their partner's image results in their partner seeing the participant averting their gaze down and away from them (Chen, 2002; Suwita et al., 1997), the previous finding that people are faster at detecting peripheral targets when fixating on centrally located faces exhibiting averted gaze than direct gaze (Senju & Hasegawa, 2005), that a partner's downward averted gaze can drive a partner's gaze towards their direction of attention (Frischen et al., 2007), the abundance of evidence suggesting people's preferential attention towards faces in general and our own countenances in particular (Tong & Nakayama, 1999), as well as the omnipresent salience of a participant's own peripheral self-video display, this study examines the hypotheses that video chat UIs with larger self-video displays will increase participants' self-directed gaze duration and decrease their partner-directed gaze duration compared to smaller self-video displays.

H1: VC participants in the large SVD condition will spend a greater percentage of the conversation gazing at their SVDs compared to participants in the small SVD condition.
H2: After controlling for partner AOI size, VC participants with small SVDs will spend a greater percentage of the conversation gazing at their partners than participants with large SVDs.

Brainstorming can be viewed as a collaborative form of conversation to accomplish both information generation and sharing in the pursuit of one or more joint goals. Sellen (1995) suggests that video chat adds little to collaborative performance beyond what can be achieved with just audio, but others have called for more research comparing variations within existing modes as opposed to between different modes (Teoh et al, 2010). I hypothesize that varying self-video display size will modulate collaborative outcomes by competing with a participant's partner's video for their attention and possibly limiting the transmission of verbal and non-verbal information.

H3: VC partners' collaborative performance on the brainstorming task will differ between the large and small SVD conditions.

Eye gaze and other nonverbal behaviors constitute an important form of evidence that conversation partners exchange to confirm understanding, especially in cooperative work (Clark & Brennan, 1991). Given that the disparity between the camera position and a video chat participant's depicted partner distorts apparent gaze direction and prevents partners from being able to establish mutual gaze, I predict that there will be a negative relationship between the amount of time partners stare at their own self videos and the percentage of unique ideas generated by conversation pairs that they recall in common following the video chat. A study by Grayson & Monk (2003) found that even though true mutual gaze cannot be achieved with video chat systems, it is possible for users to learn to infer their partner's gaze direction sufficiently enough to perceive when their video is being looked at. Thus I also hypothesize that there will be a positive relationship between the duration of gaze participants allocate to the area of the screen depicting their partner and the grounding of new ideas that conversation pairs achieve.

H4: The percentage of the total ideas generated by a conversation pair that both VC partners report in common immediately after the brainstorming task will differ between the large and small SVD conditions.
H5: There will be a negative relationship between the percentage of the conversation that VC participants look at their own SVD and their ability to ground ideas that they developed together with their partner.
H6: There will be a positive relationship between the percentage of the conversation that VC participants look at their partner and the percentage of the total ideas generated by a conversation pair that both VC partners report in common immediately after the video chat.

2.3 Exploratory Hypotheses

By reducing the size of the self-video display, I believe that participants will distribute their attention differently to their partner's room on the screen as well as to their own room off-screen.

E1: The percentage of time that VC partners gaze at their partner's room will differ between the large and small SVD conditions.
E2: The percentage of time that VC partners gaze off-screen at their own room will differ between the large and small SVD conditions.

Following the recommendations from past reviews on eye contact, this study will assess each participant's awareness of their own gaze as well as the gaze that they perceived to receive from others (Kleinke, 1986). While several authors have long pointed out that the position of the video camera above the display causes people to appear as though they are looking down (Stokes, 1969; Sallio et al., 1982; Mühlbach et al., 1985; Flohrer & Weikinnes, 1988; Suwita et al., 1997), no one has investigated whether this affects how long participants think their partner(s) stare at their own self-video display located below the axis of the camera. Grayson & Monk (2003) suggest that participants may be able to learn whether their partners are looking at the participant during video chat, but did not consider: how self-video displays that are ubiquitous in widespread video chat interfaces may influence these estimates, whether people respond differently to full screen video chat windows that increase the parallax between the video chat camera and a partner's depicted face, and whether people's ability to learn and discern where their partner is looking may differ during realistic video chat exchanges. Given the disparity between the position of video chat cameras above the screen and partners' faces below the camera, I hypothesize that people will misestimate how much gaze their partners allocated to them, their self-video displays, their rooms and off-screen immediately following a task.

E3: VC participants will misestimate the percentage of the conversation that they gaze at their own SVD.
E4: VC participants will misestimate the percentage of the conversation that their partners gaze at their corresponding SVDs.
E5: VC participants will misestimate the percentage of the conversation that they gaze at their partner.
E6: VC participants will misestimate the percentage of the conversation that their partners gaze at them.
E7: VC participants will misestimate the percentage of the conversation that they gaze at their partner's room.
E8: VC participants will misestimate the percentage of the conversation that they gaze off-screen at their own room.

There is a dearth of prior evidence assessing people's subjective assessment of how people appraise their self-video displays. Human factors and HCI researchers have long discussed and researched possible solutions to the aforementioned parallax problem but have yet to investigate publicly voiced concerns about the characteristics of self-video displays. For instance, one public thread on Skype's forum has twenty-five postings of different people requesting the option to remove their self-video from view (Jeffreys, 2014). The thread has now been viewed 17,419 times. Currently, it is unclear just how many people would prefer to chat with or without their self-video display present. I hypothesize that self-video display size will influence participants' ratings of perceived comfort and distraction and affect whether or not they prefer to video chat with their SVD visible.

E9: VC participants' SVD comfort ratings will differ between the small and large SVD groups.

E10: VC participants' SVD distraction ratings will differ between the small and large SVD groups.

E11: VC participants with large SVDs will be more likely to prefer video chatting without an SVD compared to participants with small SVDs.

CHAPTER 3

METHODS

3.1 Participants

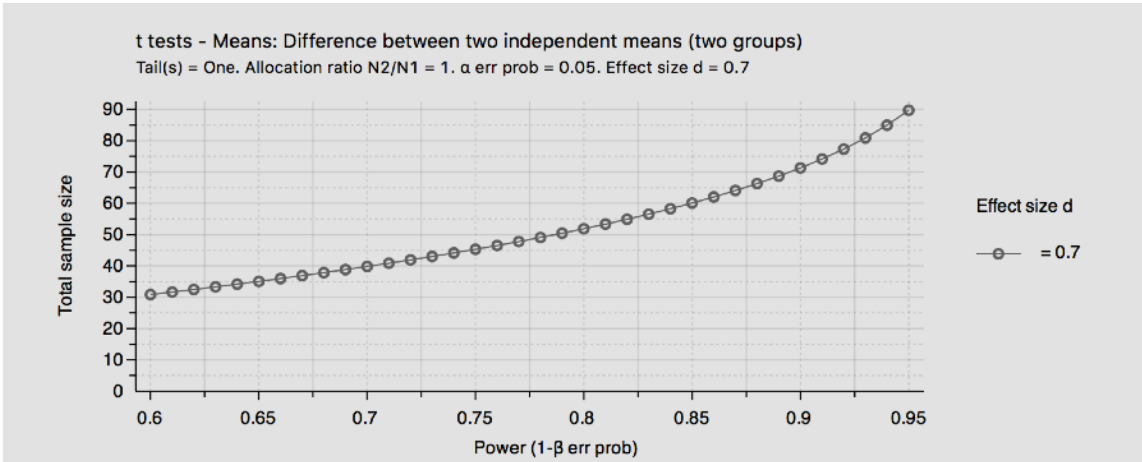
20 pairs of non-related female college students (n=40, age range: 18 - 25, M=20.4, SD=1.74) were recruited from a WEIRD (western, educated industrialized, rich, democratic) population to voluntarily participate in a 15-minute brainstorming video chat (Henrich et al., 2010). Participants were recruited using a combination of flyers distributed around the Cornell campus (posted on the information boards in buildings belonging to each of Cornell’s seven undergraduate colleges) and using the Department of Psychology's SUSAN online recruitment service. Participants were active U.S. (n=38) or Canadian (n=2) citizens. All participants were healthy and confirmed that they did not wear glasses or contacts. Everyone completed signed consent forms and were compensated with either $15 or extra credit via the SUSAN service following the completion of the study. The experimental protocol was approved in advance by the Cornell University Institutional Review Board for Human Participants. Male participants, participants not enrolled at Cornell (i.e. Cornell employees), people outside the target age range of 18 – 25 and people wearing glasses or contacts were all excluded from this study.

3.2 Statistical Power and Sample Size

Sample size was estimated using the G*power calculation for the (first) one-tailed hypothesis (H1). I expected the effect size to be ‘medium - large’ (d=.7) and sought a statistical power (T) level of =.7, given an alpha significance level of .05. The requisite sample size for this experimental design with k=2 parallel groups was calculated to be n(tot)=40 participants, or 10 pairs of partners (n=20 participants) for each self-video display size condition (see Figure 5).

Figure 5

Sample Size vs. Power Calculation



3.3 Randomization Procedure

Given past observations that individuals with ongoing relationships together may need fewer cues to communicate successfully, this study controlled for prior relationship to constrain partners to form common ground through the course of the video chat without contamination from previous experience working together or other relationship-based biases (Acker, Clift and Branco, 1985). Pairs of conversation partners with no prior direct relationship were formed through random assignment. Interested participants were contacted individually via email and assigned IDs 1- 40 once they confirmed their intent to participate in the study (ID numbers were allocated sequentially from earliest to latest response). R, the open source statistical package, was used to generate a randomly ordered list of these IDs (see Appendix A). Random pairs of participants were generated by splitting every other number in this sequence into one of two columns with each row acting as one random conversation pair. To ensure that all participants were paired with people they had not met previously, prior relation was checked using the post-conversation survey. Each conversation pair was then randomly assigned to one of two parallel conditions, either the small self-video display condition or the medium self-video display condition.

3.4 Experimental Setup

All of the collaborative video chat sessions were scheduled at a mutually convenient time for both partners via independent email correspondences. One participant from each pair was instructed to arrive ten minutes before the other participant to prevent both partners from meeting beforehand in the hallway or entry room. Before each study session began, each workstation was set up with Skype open on both desktop monitors, the Skype window maximized to full screen and the self-video display size adjusted to the assigned condition.

Each partner within a pair was assigned to one of two similar rooms in the Human Ecology Building Human Factors Laboratory. For the first ten minutes of the session the first participant was led into the furthest of the two testing rooms. Once inside, the participant was told to sit down and make themselves comfortable in front of a desktop computer while they read over the consent form. While this first participant reviewed the form, the second partner was led to an adjacent room in the human factors lab and instructed to also review the consent form. The experimenter then met with each video chat partner independently to present an overview of their role in the study, and encouraged them to ask any questions. Each participant was individually informed that they could quit the study at any time without any penalty and that the experimenter would be standing by outside the room if any issues arose. Following this, both participants spent two minutes filling out a paper form with their background information (see Appendix A). After collecting the background form, the experimenter spent five minutes calibrating the eye tracking camera systems in each room.

Equipment

Eye Tracking System:

Two remote Mirametrix S2 eye trackers were selected for this study based on their satisfactory technical specifications and unobtrusive appearance which supported the dynamic task at hand without requiring participants to wear any apparatus. Each Mirametrix S2 eye tracker offered binocular bright pupil tracking, was accurate to <1 degree of visual angle, allowed remote gaze data capture at 60 Hz and supported head motion within a 25cm (width) x 11cm (height) x 30cm (depth) window of each participant’s initial calibration position 65cm away from the tracker (Mirametrix, 2011). The intensity of the infrared light source (<1 mW per cm^2) used for bright pupil tracking satisfied IEC 60825 as well as the NIOSH safety requirements for infrared exposure levels (Mirametrix, 2011). Each eye tracker was discretely placed under a desktop screen so that the top of the eye tracker bar was centered and flush with the lower edge of the screen frame.

Figure 6

Computer + Eye Tracker Setup

Computer Display:

Each partner used a 19” (1280 x 1024 pixels) desktop screen connected to a Dell (Dell Optiplex 320 Core2Duo, 1.80GHz, 2GB RAM, 300GB HDD) computer running the Windows 7 operating system (see Appendix A).

Video Camera:

A clip-on Logitech HD Web Cam C270 was affixed to the top of the desktop screen used for this study with the camera angled downwards 10 degrees from horizontal. The web camera enabled HD (1280 x 1024 pixels) video chatting via a USB 2.0 connection (see Appendix A).

Eye Tracking Software

Mirametrix Tracker:

The Tracker software was used on both partners' computers to record gaze behavior and determine where they were looking on the screen during each video chat (see Appendix A).

Mirametrix Viewer:
The Viewer was used in tandem with the Tracker to initiate both gaze and screen recording for all of the video chat sessions and to save these for further viewing and analysis.

Video Chat Service:
Skype version 6.3.60 (c) 2003 - 2013 was selected as the video chat software on the basis of its prevalence and familiarity to most university students and working professionals (see Appendix A). Both experimental conditions were conducted in full screen mode. In the large self-video display condition the self-portrait subtended a vertical visual angle of 5.5 degrees and a horizontal angle of 7.3 degrees from the participant sitting 65cm away from the screen. The small self-video display subtended a vertical visual angle of .8 degrees and a horizontal angle of 1.43 degrees.

Environment

Human Factors Lab:
Each 12' (length) x 8' (width) x 10' (height) testing room contained matching white walls, mirrored layouts, an identical fixed office chair, 29" office desk, desktop computer (Windows 7 operating system) and remote eye tracking setups (Mirametrix). All ceiling lighting in the room was diffuse, indirect and out of the eye tracker's field of view. Participants were seated on opposite sides of their respective rooms and the door between them was shut during the video chat conversation to attenuate sound leakage between rooms and eliminate echo (Figure 7).

Figure 7
Human Factors Lab Environment

3.5 Calibration and Partial Blinding

To calibrate the eye tracker, each participant was asked to assume a comfortable seated position with their back resting directly against the back of the chair. The chair was then adjusted so that their eyes were approximately 65cm away from the eye tracker. This distance - measured from the participants' pupils to the midpoint of the Mirametrix eye tracker - was confirmed using the depth estimate bar in the Mirametrix Tracker software. Since eye movements are known to be task dependent (Duchowski, 2007), participants were told that the computers were installed with a system for measuring posture and eye movements and that they should try to keep a consistent, relaxed posture during the video chat. After the experimenter ensured that the system registered each participant's left and right eyes and was centered on their pupils, participants were then told to look at a sequence of nine blue dots on the screen to calibrate the measurement system. This exercise was repeated until the calibration achieved an average accuracy of < .5 degrees of visual angle across the entire screen. To further prevent the experimental hypotheses from being inferred, participants were also not informed whether or not they were assigned to the medium or the small self-video display condition and were kept unaware of their partner's condition.

3.6 Video Chat Brainstorming Task

The brainstorming task was constructed to address several aspects of collaborative work in distributed teams:

1) The task was representative of a common activity conducted by most companies: developing ideas for how to improve an organization

2) The prompt participants responded to was open ended and demanded multiple possible ideas as opposed to a single ‘correct’ solution
3) The purpose of the prompt was pertinent to participants’ everyday life and relationship to the organization that they were members of
4) The task did not require specific domain skills other than the capacity to generate new ideas
5) Participants needed to cooperatively work together as a team to develop new ideas

After calibrating the eye tracking setups, the experimenter used the Mirametrix Viewer software to start recording each partner’s gaze, screen video and time interval. Partner 1 was prompted to start a fifteen minute Skype video chat with partner 2 discussing the provided question: How might Cornell University improve student life on campus? Participants were informed that their goal was to generate as many ideas as they could together over the 15-minute session. Each participant’s gaze was tracked through the course of the entire video chat session starting from the moment the experimenter closed the door behind the two participants. Following the video chat, each conversation partner independently filled out a brief self-report questionnaire (see Appendix A) containing additional questions about their gaze estimates during the conversation and personal evaluations of the self-video display.

3.7 Visual Attention Measures

Throughout the video chat, each participant’s gaze points both on and off the screen were recorded sixty times per second to create a record of where they were looking (gaze points) and what they were looking at (AOI). Following the study, each gaze point was labeled with the appropriate area of interest (AOI) that it fell within. Both gaze points and AOIs were used to calculate the following objective measures of the proportion of a participant’s gaze directed at their SVD, their partner, their partner’s room and their own room over the duration of the video chat.

Personal Visual Attention

1) Self Directed Gaze:

The % of a participant’s total gaze directed at their own self-video display.

Interpersonal Visual Attention

2) Partner Directed Gaze:

The % of a participant’s total gaze directed at their chat partner.

Environmental Visual Attention

3) Partner Room Directed Gaze:

The % of a participant's total gaze directed at the environment behind their chat partner.

4) Personal Room Directed Gaze:

The % of a participant's total gaze directed off-screen at the environment surrounding their display monitor.

At the end of the video chat, the participants independently completed an online questionnaire (see Appendix A for a transcribed version) and were asked to estimate the percentage of the video chat duration that they looked at themselves, their partner, their partner's room and off-screen. The sequence of these questions was randomized for each participant.

Self-Reported Visual Attention

5) Est. Self-Directed Gaze:

The estimated % of the video chat that a participant looked at their own self-video display.

6) Est. Partner Directed Gaze:

The estimated % of the video chat that a participant looked at their conversation partner.

7) Est. Partner's Self Directed Gaze:

The estimated % of the video chat that a participant thought their partner looked at their own self-portrait window.

8) Est. Partner's Gaze Directed at the Participant:

The estimated % of the video chat that a participant (A) thought her partner (B) looked at the participant (A).

9) Est. Partner Room Directed Gaze

The estimated % of the video chat that a participant looked at their partner's room.

10) Est. Personal Room Directed Gaze

The estimated % of the video chat that a participant looked (off screen) at their own room.

Gaze Awareness

11) Gaze Awareness:

The difference between measured visual attention and estimated visual attention for each region of interest.

After completing these estimates, participants were asked to list all of the improvements that they identified with their partner during the video chat. These were used to calculate measures of each conversation pair's collaborative performance and common ground at the end of the brainstorm session.

Collaborative Outcomes

12) Collaborative Performance:

The total # of unique ideas generated by a pair of partners

13) Common Ground:

The percentage of the total # of unique ideas generated by a pair that were reported in common by both partners immediately following the video chat.

Next, participants completed a subjective evaluation of how comforting and distracting they perceived their self-video window to be by marking their rating on a 4 point Likert scale. Finally, they were asked to indicate their preference whether they would rather video chat with or without a self-video display window.

Self-video Display Evaluation

SVD Comfort Rating:

A participant's subjective rating of the perceived comfort provided by their SVD (1 = not comforting, 2 = slightly comforting, 3 = comforting, 4= very comforting)

SVD Distraction Rating:

A participant's subjective rating of the perceived distraction provided by their SVD (1 = not distracting, 2 = slightly distracting, 3 = distracting, 4 = very distracting)

SVD Presence Preference:

A participant's preference to video chat with or without a self-video display present (1 = would rather chat with an SVD, 2 = would rather chat without an SVD)

3.8 Data Treatment and Track Analysis

All eye tracking data was exported as CSV tables using the Mirametrix Viewer software. The x and y coordinates for all of the gaze points were originally recorded as percentages of the screen size and later converted into pixels by multiplying the percentages by the screen pixel dimensions (1280 x 1024) in SPSS. For X coordinates, 0 was the left most position while 1280 was the right most position of the screen. For Y coordinates, 0 was the top-most position and 1024 was the bottom-most position.

Gaze Points

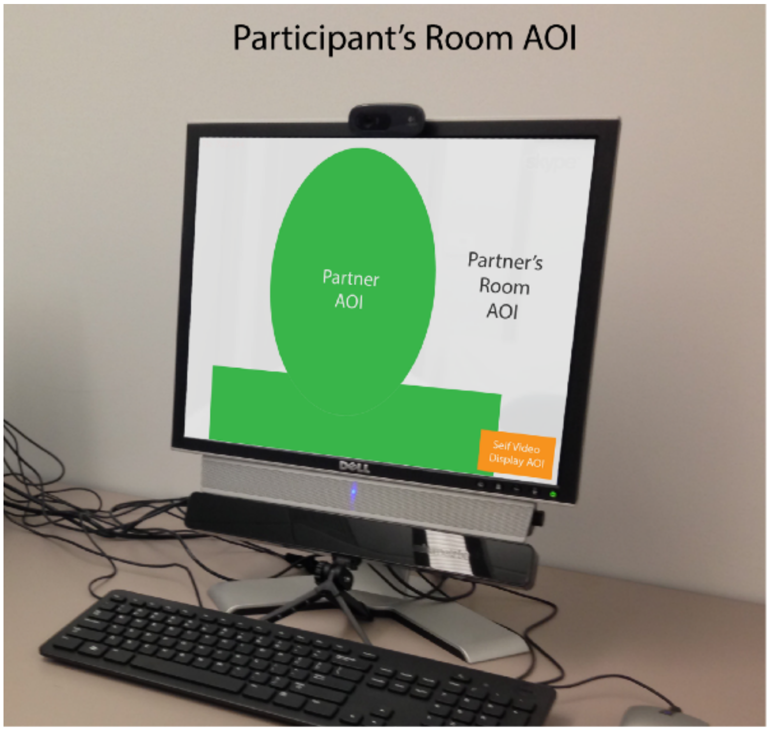
Each partner's gaze points were recorded using Mirametrix software for both their left and right eyes. If both the left (LXg, LYg) and right (RXg, RYg) gaze coordinates were valid for a particular instance, the gaze point (Xg, Yg) was computed as the average of both of the left and right coordinates. If either the left or right gaze coordinates were invalid, then the gaze point (Xg, Yg) equaled the remaining valid gaze coordinates. If both the left and right gaze coordinates were invalid for a particular time stamp, then the gaze point for that instance was also marked as invalid. Gaze points that occurred during blinks, saccades and periods of tracker loss were all marked as invalid by Mirametrix's internal algorithm.

Areas of Interest (AOI)

In order to study the influence of SVD size on personal, interpersonal and environmental gaze behavior, raw eye tracking data for each participant was grouped into areas of interest (AOI). Screen AOIs were constructed around the pixels depicting each participant's self-video display, their partner and their partner's room. Additionally, an AOI was computed for the participant's room comprising the offscreen area surrounding the display monitor. Gaze counts and percentages of total gaze were calculated for each AOI.

Figure 8

Areas of Interest



AOIs were constructed around the following VC interface elements for each participant: a) the portion of the screen displaying their partner b) the area of the screen showing the participant's SVD c) the region of the screen depicting their partner's room d) the offscreen area of the participant's visual field covering their own room. Partner, Self-video Display, Partner's Room and Participant's Room AOIs were calculated for each participant by 1) taking four full screen shots of each recorded video chat session at 5 minute increments (0:00, 5:00, 10:00, 15:00); 2) fitting each screen shot to a pixel grid matching the recorded desktop screen dimensions (1280 x 1024) in Adobe Illustrator; 3) fitting an ellipse around the boundary of their partner's head & neck; 4) averaging the center coordinates, minor radius and major radius of each set of four ellipses to determine the Partner Head AOI location and dimensions; 5) fitting a rectangle around the visible portion of their partner's upper torso; 6) averaging center x coordinates, width and length measurements of each set of four rectangles placed over the partner's upper body minus the intersection with the Partner Head AOI to determine the mean Partner Torso AOI location and dimensions; 7) taking the union of the Partner Head AOI and Partner Torso AOI to determine the corresponding Partner AOI; 8) fitting a rectangle over the participant's own SVD; 9) taking the center x coordinate, length and width of the rectangle overlaying their self-video display to determine the self-video display AOI; 10) confirming that all SVDs were consistently sized and placed within each of the two different experimental SVD conditions; 11) subtracting the intersection of the total screen area with the union of the Partner AOI and Self-video AOI to determine the Partner's Room AOI; 12) subtracting the total screen area from the total recorded area to determine the Participant's Room AOI.

All gaze data points were labelled with the AOI corresponding to where each participant was looking at the time. To determine whether each gaze point was bounded by a specified AOI, x and y coordinates for the corresponding gaze point were inserted into the following equations and labelled accordingly.

Partner Head AOI:
A gaze point (xg, yg) was bounded by the Partner Head AOI when the following was satisfied:
Where xg refers to the horizontal pixel coordinate of the gaze point on the computer screen, h refers to the horizontal coordinate of the ellipse (head) center point on the screen, y refers to the vertical pixel coordinate of the gaze point on the screen, k refers to the vertical coordinate of the ellipse (head) center point on the screen, rx represents the semi minor horizontal radius of the ellipse and ry represents the semi major vertical radius of the ellipse. If the outcome was less than or equal to 1 the gaze point was labelled as being within the Partner Head AOI.

$$((xg - h)^2)/(rx^2) + ((yg - k)^2)/(ry^2) \leq 1$$

Partner Torso AOI:
A gaze point (xg, yg) was bounded by the Partner Torso AOI when the following was satisfied:

$$(x1 \leq xg \leq x2), (y1 \leq yg \leq y2)$$

AND

$$((xg - h)^2)/(rx^2) + ((yg - k)^2)/(ry^2) > 1$$

Where xg is the horizontal pixel coordinate of the gaze point on the computer screen, x1 is the x-value of the upper left corner (x1,y1) of the rectangle and x2 is the x-value of the lower right corner (x2,y2).

If xg was between x1 and x2, yg was less than or equal to y2 and (xg, yg) was not contained within the partner's head AOI then the gaze point was labelled as residing within the Torso AOI.

Partner AOI:
A gaze point (xg, yg) was bounded by the Partner AOI when the following was satisfied:

$$(x1 \leq xg \leq x2), (yg > y1)$$

OR

$$((xg - h)^2)/(rx^2) + ((yg - k)^2)/(ry^2) \leq 1$$

If a gaze point (xg, yg) was contained within the partner's head AOI) or the partner AOI then the gaze point was labelled as being within the partner AOI.

Self-video Display AOI:
A gaze point (xg, yg) was bounded by the Self-video Display AOI when the following was satisfied:

$$(xi \leq xg \leq xii)$$

AND

$$(yi \leq yg \leq yii)$$

Where xg is the horizontal pixel coordinate of the gaze point on the display screen, xi is the x-value of the left edge of the SVD, xii is the x-value of the right edge of the SVD, yi is the y-value of the top edge of the SVD and yii describes the y-value of the bottom edge of the SVD.

If (xg was between xi and xii) and (yg was between yi and yii), then the gaze point was coded within the Self-video Display AOI.

Partner's Room AOI:
A gaze point (xg, yg) was bounded by the Partner's Room AOI when the following was satisfied:

$$((x1 > xg \text{ OR } xg > x2), (yg \leq y1) \text{ OR } ((xg - h)^2)/(rx^2) + ((yg - k)^2)/(ry^2) > 1))$$

AND

$$((xi > xg \text{ OR } xg > xii) \text{ AND } (yii < yg \text{ OR } yg < yi))$$

AND

$$(0 < xg < 1280, 0 < yg < 1024)$$

Where x=1280 is the rightmost edge of the screen and y=1024 is the bottommost edge of the screen.

If a gaze point (xg, yg) was not contained within the Partner AOI, Self Display AOI and was not located off screen, then the gaze point was coded within the Partner's Room AOI.

Participant's Room (Offscreen) AOI:
A gaze point (xg, yg) was bounded within the Participant's Room AOI when the following was satisfied:

$$xg > 1280$$

OR

$$xg < 0$$

OR

$$yg > 1024$$

OR

$$yg < 0$$

If a gaze point (xg, yg) was not contained within the participant's screen, then the gaze point was coded within the Participant's Room AOI.

After all gaze points were labelled with their corresponding AOI, objective measures of visual attention were computed by dividing the total # of gaze points each participant allocated to each AOI by the total # of gaze points for the entire duration of their 15-minute video chat.

Partner-directed gaze was calculated by dividing the total # of valid gaze points a participant directed at their partner's AOI divided by the total # of valid gaze points for the entire duration of the 15-minute video chat. Similarly, self-directed gaze was computed by dividing the total # of valid gaze points a participant allocated to their self-video display divided by the total # of valid gaze points for the video chat. Partner's room-directed gaze was calculated by dividing the total # of gaze points aimed at the portion of the screen depicting their partner's room divided by the total # of valid gaze points for the video chat. Lastly, participant's room gaze was computed by dividing the total # of gaze points a participant directed off screen by the total # of valid gaze points for the video chat. All background data was collected using paper surveys and all self-report data was recorded through online forms. The Statistical Package for Social Sciences (SPSS version 24) was used to aggregate all of the experimental objective and subjective data and to compute all of the statistical analyses.

Track loss

For each participant, all gaze points that occurred during blinks were excluded (where xg = 0, yg = 0) from the analysis. Additionally, all invalid points of gaze were removed (all cases where there were no valid gaze points for both the left and right eye). All of the remaining points of gaze within the 15minute duration of each video chat were included in the analysis (see Appendix for a table of track loss data). At least 50% of a participant's tracks needed to be valid for their eye tracking data to be included in this study.

CHAPTER 4

RESULTS

4.1 Demographics

Forty undergraduate female participants (age range: 18 - 25, M=20.4, SD=1.74) participated in this study. Approximately one third (32.5%) of participants reported using video chat one or more times per week and all participants reported using video chat at least once per month on average (Median use = 1.5 sessions/month, IQR = 1 - 4 sessions/month).

4.2 Participant Flow

From the 40 scheduled participants, 22 (11 pairs) were randomly assigned to the medium SVD condition and 18 (9 pairs) were assigned to the small SVD condition. Six participants' gaze data were excluded from the primary analysis as a result of eye tracking equipment errors. All participants' self-report data were included in the exploratory gaze estimation, collaborative performance and subjective evaluation analyses.

Table 3

Participant Flow

	Medium SVD Condition	Small SVD Condition
Participants Randomly Assigned	22	18
Participants' Data Excluded From Primary Analysis	4	2
Participants' Data Available for Primary Analysis	18	16

4.3 Analysis Procedure

For all personal, interpersonal and environmental visual attention outcomes, independent 2 sample t tests were used to compare the between groups gaze behaviors of video chat partners with large SVDs vs. the gaze behaviors of video chat partners with small SVDs. Likewise, independent 2 sample t tests were computed to compare the collaborative outcomes of partners with large SVDs and partners with small SVDs. Shapiro Wilk tests were used to assess normality and Levene's tests were used to assess homogeneity of variance. Skewed data and/or heteroscedastic data were described with medians, interquartile ranges (with lower and upper bounds) and analyzed with non-parametric Mann Whitney tests. Spearman's rank order correlations were computed to assess possible relationships between gaze behavior and grounding. Pearson's Chi-Square tests were used to compare self-video display evaluation questions between the two SVD conditions. Bonferroni corrections were used to adjust for multiple comparisons within the primary analysis.

For all gaze awareness measures, Wilcoxon signed-rank tests were used to compare participant's estimates of their gaze behavior with actual measures of their gaze obtained from the eye tracking setup. Wilcoxon signed-ranks were also utilized to compare participants' estimates of their partner's gaze behavior with actual measures of their partner's gaze.

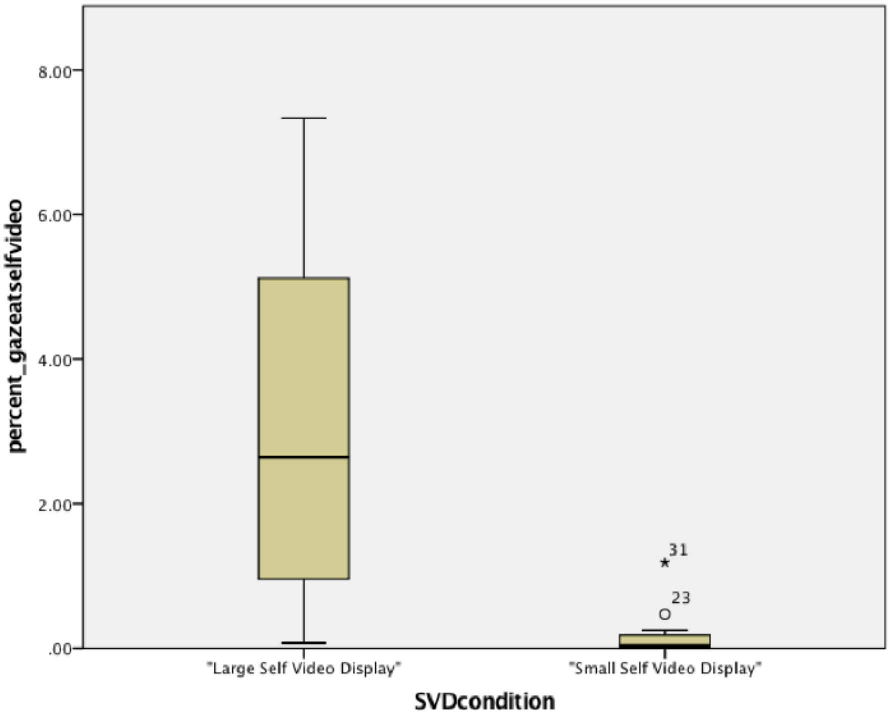
Participants' subjective SVD evaluations were compared between the large and small SVD conditions with cross tabulated frequencies and Chi-Square tests. In order to fulfill the assumptions of the Chi-Square analysis, perceived comfort and perceived distraction were converted into binary variables. Different positive levels of perceived comfort (slightly comforting, moderately comforting, very comforting) were collapsed into 'comforting' while the alternative 'not comforting' remained the same. Similarly, the different levels of perceived distraction (slightly distracting, moderately distracting, very distracting) were collapsed into 'distracting' while the alternative 'not distracting' remained the same. All tests were pre-specified.

4.4 Personal Visual Attention

Self-Directed Gaze

H1: There was an effect of SVD size on self-directed gaze. VC participants with large SVDs (Mdn = 2.89%, IQR = .92 - 6.33%) spent significantly more of the conversation gazing at their self-video displays compared to participants with small SVDs (Mdn = .03%, IQR = 0 - .19%) $U = 15$, $Z = -4.454$, $p < .001$, $r = -.75$).

Figure 9
Self-Video Size vs. Self-Directed Gaze



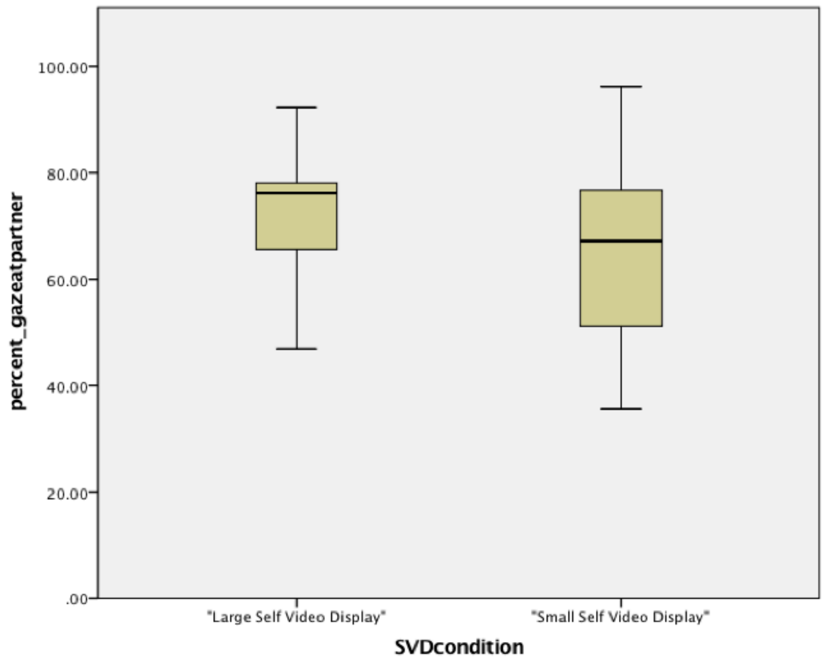
4.5 Interpersonal Visual Attention

Partner Directed Gaze

H2: The covariate, partner AOI size, was not significantly related to the percentage of the conversation participants spent gazing at their partners F(1, 31) = 1.036, p=.317. After controlling for AOI size, there was no significant effect of SVD size on the percentage of the conversation participants spent gazing at their partners, F(2, 31), p = .163, partial eta squared = .062): there was no significant difference detected between VC participants with large SVDs (Mean = 72.68%, CI = 65.44 - 79.91%) and small SVDs (Mean = 65.27%, CI = 57.60% - 72.94%).

Figure 10

Self-Video Size vs. Partner-Directed Gaze



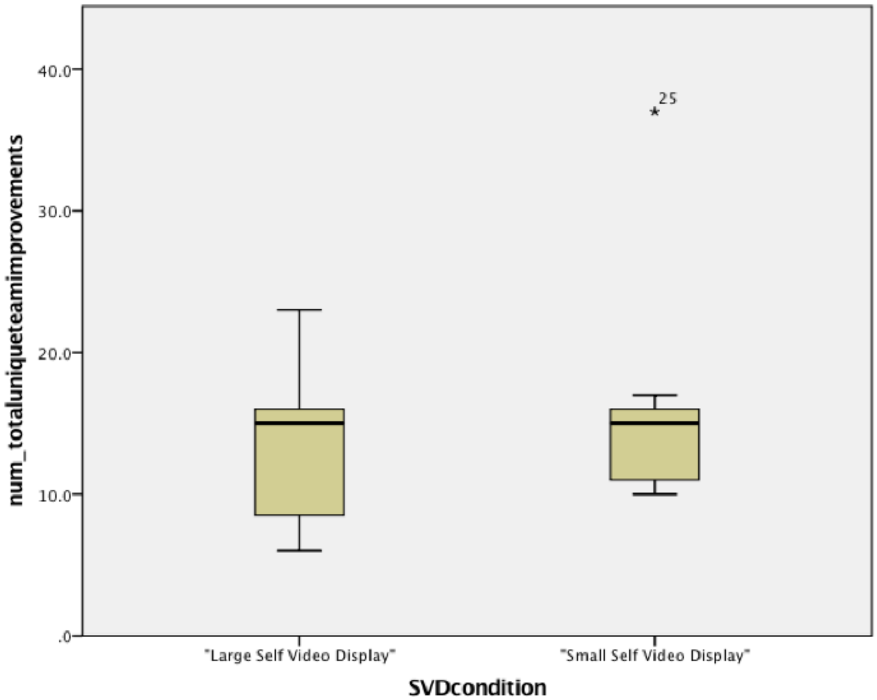
4.6 Collaborative Outcomes

Collaborative Performance

H3: There was no effect of SVD size on the collaborative performance of VC conversation pairs. The total number of unique ideas generated by VC conversation pairs with large SVDs (Mdn = 15, IQR = 7 - 17) did not significantly differ from pairs with small SVDs (Mdn = 15, IQR = 10.5 - 16.5) U = 55.5, Z = .46, p = .65, r= .07).

Figure 11

Self-Video Size vs. Tot. Unique Team Improvements

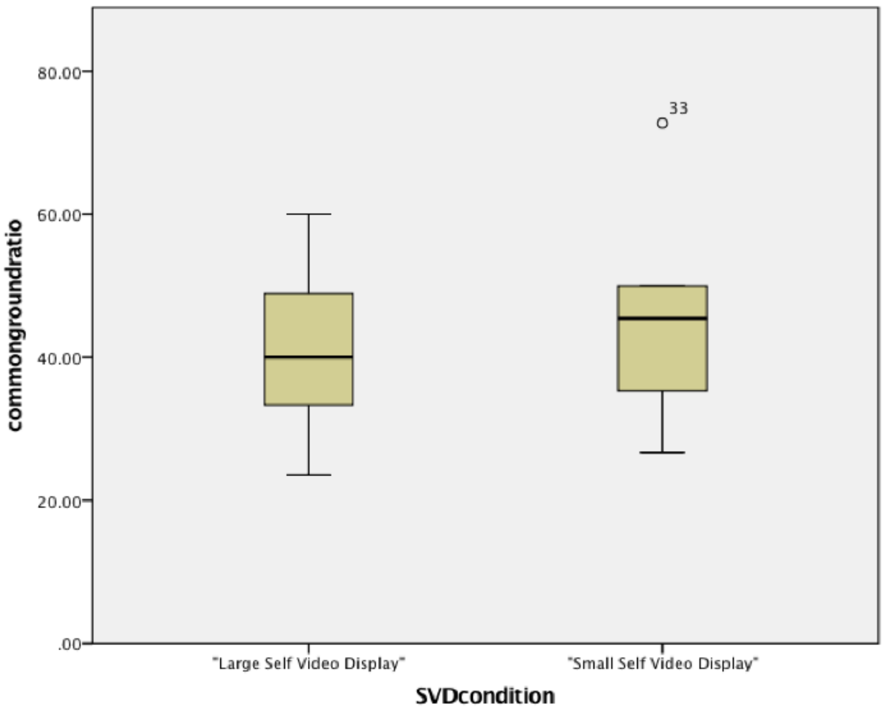


Common Ground

H4: The percentage of the total ideas generated by a conversation pair that both VC partners report in common immediately after the video chat will differ between the large and small SVD conditions.

On average, there was no effect of SVD size on the common ground conversation partners demonstrated after the brainstorming video chat. The percentage of the total ideas generated by conversation pairs that both VC partners reported in common immediately after the video chat did not significantly differ (Mean difference = -2.52%, 95% CI (-14.52, 9.48), $t(18) = -.442$, $p=.664$) between pairs that chatted with large SVDs ($M = 41.65\%$, $SE = 3.48$) and pairs that chatted with small SVDs ($M = 44.18\%$, $SE = 4.67$).

Figure 12
Self-video Size vs. % Grounded Ideas



H5: For both SVD groups, there was a non-significant, negative relationship (Large SVD: r_s = -.50, p = .125 | Small SVD: r_s = -.479, p = .230) between the percentage of the conversation that participants gazed at their SVD and the common ground they established (the percentage of the total ideas generated with their partner that they reported in common after video chatting).

H6: For both SVD groups, there was a non-significant positive relationship (Large SVD: r_s = .39, p = .265 | Small SVD: r_s = .371, p = .365 between the percentage of the conversation that participants gazed at their partner and the common ground they established (the percentage of the total ideas generated with their partner that they reported in common after video chatting).

4.7 Environmental Visual Attention

Partner's room directed gaze

E1: There was no effect of SVD size on the percentage of time that VC partners gazed at their partner's room. VC participants with large SVDs (Mdn = 4.78, IQR = 2.04 -

6.67%) did not spend a significantly different proportion of the conversation gazing at their partner's room compared to participants with small SVDs (Mdn = 5.07%, IQR = 3.64 - 12.71%) U = 187, Z = 1.48, p=.138, r=.23).

Personal room directed gaze

E2: On average, there was no effect of SVD size on the percentage of gaze participants allocated off-screen to their own personal room. The percentage of time that VC partners gazed at their own room did not significantly differ (Mean Difference: -8.28, 95% CI (-18.09 - 1.53), t(32), p = .073) between participants with large SVDs (M = 18.05%, SE = 2.65) and participants with small SVDs (M = 26.33%, SE = 4.15).

4.8 Retrospective Gaze Awareness

Participant's Gaze Awareness of Looking at Their Own SVD

E3: VC participants in both the large and small SVD groups overestimated the percentage of the conversation that they were staring at their own SVD. For VC participants with large SVDs, their estimates of self-directed gaze (Mdn = 20%, IQR = 5 - 31.25%) were significantly higher compared to their recorded self-directed gaze (Mdn = 2.89%, IQR = .92 - 6.33%), T = 160, p < .001, r = .54. Similarly, VC participants with small SVDs gave estimates of self-directed gaze (Mdn = 7.5%, IQR = 2.5 - 12.5%) that were also significantly higher compared to their recorded self-directed gaze (Mdn = .03%, IQR = 0 - .19%), T = 135, p<.001, r = .61. In each SVD size group, all of the participants except for 1 overestimated how long they gazed at their own SVD.

Participant's Gaze Awareness of Partner Looking at Their Own SVD

E4: VC participants in both the large and small SVD groups overestimated the percentage of the conversation that their partners were gazing at their corresponding SVD. For VC participants with large SVDs, their estimates of how long their partner's gazed at their corresponding SVD (Mdn = 10%, IQR = 5 - 30%) were significantly higher compared to their partners' measured gaze directed at their corresponding SVD (Mdn = 2.89%, IQR = .92 - 6.33%), T = 168, p < .001 , r =.60. Similarly, VC participants with small SVDs provided estimates of how long their partners gazed at their corresponding SVD (Mdn = 10%, IQR = 4.5 - 26.25%) that were also significantly higher compared to their partners' recorded self-directed gaze (Mdn = .03%, IQR = 0 - .19%), T = 119, p < .001, r =.59. In each SVD group, all of the participants except for 1 overestimated how long their partner gazed at their corresponding SVD.

Participant's Gaze Awareness of Looking at Their Partner

E5: Participants' estimates of how long they gazed at their partner did not significantly differ from their measured gaze at their partner, regardless of SVD size. For participants with large SVDs, their estimates of the percentage of the VC conversation they gazed at their partner (Mdn = 65%, IQR = 48.75 - 90%) did not significantly differ from the measured gaze they actually directed at their partner (Mdn = 76.14%, IQR = 63.89 - 78.75%) T = 55, p = .18, r = .22. Likewise, for participants with small SVDs, their estimates of the percentage of the VC conversation they gazed at their partner (Mdn = 62.5%, IQR = 43.75 - 95%) did not significantly differ from the measured gaze they actually directed at their partner (Mdn = 67.12%, IQR = 50.43 - 77.41%) T =. 71, p =.88 , r =.03.

Participant's Gaze Awareness of Partner Looking at the Participant

E6: Participants' estimates of how long their partners gazed at them did not significantly differ from their partner's measured gaze directed at the participant. For participants with large SVDs, their estimates of the percentage of the video chat that their partners gazed at them (Mdn = 75%, IQR = 50 - 90%) did not significantly differ from the measured percentage of the conversation that their partner gazed at them (Mdn = 76.14%, IQR = 63.89 - 78.75%) T = 67, p = .420, r = .13. Similarly, for participants with small SVDs, their estimates of the percentage of the video chat that their partners gazed at them (Mdn = 70%, IQR = 37.5 - 95%) did not significantly differ from the measured percentage of the conversation that their partner gazed at them (Mdn = 67.13%, IQR = 50.43 - 77.41%) T = 68, p = 1.00, r = .001.

Participant's Gaze Awareness of Looking at Their Partner's Room

E7: Participant's estimates of how long they gazed at their partner's room did not significantly differ from the measured gaze they directed at their partner's room. For participants with large SVDs, their estimates of the percentage of the video chat that they gazed at their partner's room (Mdn = 5%, IQR = 0 - 10%) did not significantly differ from the measured percentage of the conversation that they gazed at their partner's room (Mdn = 4.78%, IQR = 2.04 - 6.67%), T = 127, p = .07, r = .3. Likewise, for participants with small SVDs, their estimates of the percentage of the video chat that they gazed at their partner's room (Mdn = 7.5, IQR = 0 - 15%) did not significantly differ from the measured percentage of the video chat that they actually gazed at their partner's room (Mdn = 5.07%, IQR = 3.64 - 12.71%) T = 67, p = .96, r = .01.

Participant's Gaze Awareness of Looking at Their Own Room

E8: Participants underestimated how long they gazed offscreen at their own room compared to the measured gaze they actually directed at their own room. For participants with large SVDs, their estimates of the percentage of the video chat that they gazed offscreen at their own room (Mdn = 0%, IQR = 0 - 5%) were significantly lower than the measured percentage of the conversation that they gazed at their own room (Mdn = 14.88%, IQR = 12.02 - 22.51%), T = 15, p = .002, r = .51. Likewise, for participants with small SVDs, their estimates of the percentage of the video chat that they gazed offscreen at their own room (Mdn = 3.5%, IQR = 0 - 5%) were significantly lower than the measured percentage of the video chat that they actually gazed at their own room (Mdn = 21.62%, IQR = 15.74 - 40.68%) T = 2, p < .001, r = .60. In both groups, all of the participants except for 1 underestimated how long they gazed off screen.

4.9 Self-Video Display Evaluation

Perceived Comfort

E9: There was no significant association detected between self-video display size and whether or not participants found their SVD comforting X^2 (1) = .135 , p = 1.00. 73% of participants in the large self-video display condition found their SVD comforting as compared to 78% of participants in the small display condition.

Perceived Distraction

E10: There was a significant association between participants' self-video display size and whether or not participants found their SVD distracting X^2 (1) = 4.55, p = (.033). Based on the odds ratio, the odds of participants with large SVDs finding their SVD distracting were 1.74 times higher than participants with small SVDs. 77% of participants with large SVDs found their SVDs distracting. By comparison, 44% of participants with small SVDs found their SVDs distracting.

SVD Preference

E11: There was no significant association detected between SVD size and whether or not participants prefer video chatting with or without an SVD X^2 (1) = .750, p = .298. Approximately one third (35%) of all participants would prefer to video chat without an SVD present (41% of participants with large SVDs and 28% of participants with small SVDs).

Table 4

Primary Hypothesis Testing Summary

Primary Hypotheses	Outcome
H1: VC participants in the large SVD condition will spend a greater percentage of the conversation gazing at their SVDs compared to participants in the small SVD condition.	Not Rejected
H2: After controlling for partner AOI size, VC participants with small SVDs will spend a greater percentage of the conversation gazing at their partners' images than participants with large SVDs.	Rejected
H3: VC partners' collaborative performance on the brainstorming task will differ between the large and small SVD conditions.	Rejected
H4: The percentage of the total ideas generated by a conversation pair that both VC partners report in common immediately after the brainstorming task will differ between the large and small SVD conditions.	Rejected
H5: There will be a negative relationship between the percentage of the conversation that VC participants look at their own SVD and their ability to ground ideas that they developed together with their partner.	Rejected
H6: There will be a positive relationship between the percentage of the conversation that VC participants look at their partner's image and the percentage of the total ideas generated by a conversation pair that both VC partners report in common immediately after the video chat.	Rejected

Table 5
Exploratory Hypothesis Testing Summary

Exploratory Hypotheses	Outcome
E1: The percentage of time that VC partners gaze at their partner's room will differ between the large and small SVD conditions.	Rejected
E2: The percentage of time that VC partners gaze at their own room will differ between the large and small SVD conditions.	Rejected
E3: VC participants will misestimate the percentage of the conversation that they gaze at their own SVD.	Not Rejected
E4: VC participants will misestimate the percentage of the conversation that their partners gaze at their corresponding SVDs.	Not Rejected
E5: VC participants will misestimate the percentage of the conversation that they gaze at their partner.	Rejected
E6: VC participants will misestimate the percentage of the conversation that their partners gaze at them.	Rejected
E7: VC participants will misestimate the percentage of the conversation that they gaze at their partner's room.	Rejected
E8: VC participants will misestimate the percentage of the conversation that they gaze offscreen at their own room.	Not Rejected
E9: VC participants' SVD comfort ratings will differ between the small and large SVD groups.	Rejected
E10: VC participants' SVD distraction ratings will differ between the small and large SVD groups.	Not Rejected
E11: VC participants with large SVDs will be more likely to prefer video chatting without an SVD compared to participants with small SVDs.	Rejected

CHAPTER 5

DISCUSSION

The primary goal of this thesis was to investigate whether the size of peoples' self-video displays during a cooperative video chat task influences how they allocate attention, how well they generate and ground new ideas together with their partners and how comforting and distracting people perceive their SVDs to be. Furthermore, this study also investigated participant's awareness of their own gaze as well as the gaze of their partner during a video chat. Table 5 presents an overview of the outcomes of all the primary and exploratory hypotheses in this study.

At the time of Rainie & Zickuhr's (2010) survey, they found that the proportion of Americans who had used the internet for video calling was increasing (from 20% in the spring of 2009 to 23% in the summer of 2010) along with the number of people using it on a daily basis (increase from 2% to 4% of the population during the same period) and that use was highest for 18 - 29 year olds (29%), college students (24-30%), those with a household income over $75,000 (34%) and males (26%) compared to females (20%). By comparison, 100% of the college aged female participants in this study had participated in a video chat and nearly a third used it one or more times a week, suggesting that use may continue to be rising across the western, educated, industrialized, rich and democratic (WEIRD) demographic surveyed in this study (Henrich et al., 2010).

5.1 Visual Attention

Across both SVD conditions, participants spent a relatively small amount of time gazing at their own self-video display (2%), compared to their partner's video (69%), offscreen (22%), and their partner's room (6%). Compared to VC participants with small SVDs, VC participants with large SVDs gazed at their SVDs longer but did not spend a significantly different percentage of the conversation gazing at their partner, their partner's room or offscreen at their own room. This suggests that self-video display size influences participants' self-directed gaze but does not detract from their overall overt attention directed to other areas of interest both on and off the screen. On average participants with large SVDs spent nearly 3% (27 seconds) of the 15-minute conversation gazing at themselves compared to those with smaller SVDs who spent nearly none of the conversation overtly attending to their self-video display. Those with small SVDs appeared to distribute the additional 3% gaze evenly between their partner's room and offscreen at their own room. There are at least three possible explanations for why participants with larger self-video displays may have gazed at their SVD longer. First, the human visual system is highly sensitive to faces. Previous visual perception studies suggest that we attend to faces faster than competing stimuli and are far more likely to exhibit preferential visual search strategies for looking at our own faces (Palermo & Rhodes, 2007; Tong & Nakayama, 1999). The larger self-video display transmits a larger retinal image that may increase the salience of the participant's own face in their field of view and provide a larger area for a participant to attend. Second, gaze can be involuntarily driven in the direction that another person appears to be looking. People respond faster to stimuli that appear at the edge of a screen when a central face appears to be looking in the same direction, even when participants are verbally informed that the stimuli are more likely to appear on the opposite side of the screen from the face's direction of gaze (Friesen & Kingstone, 1998; Friesen et al., 2004). Hence, it is possible that the apparent downward gaze of a video chat participant's partner cues them to attend to their own self-video displays longer. Third, the human visual system is attracted to morphologically similar and nearby items. Snowden et al. (2012) suggest that patterns of fixations are not random; we tend to fixate on similar elements that look like the target we are engaging with and prefer to fixate on proximate entities located close to where we were previously looking. The visual proximity and morphological similarity between the participant's self-video and the video of their partner may also contribute to self-directed gaze. Since the second and third factors were held relatively constant across SVD size conditions, it seems likely that the increased size and salience of a participant's own face as well as its closer proximity to their partner's video was the primary driver of their self-directed gaze.

However, self-gaze appears to be highly idiosyncratic at the level of the individual, ranging from 29% - 0%. Indeed, other personal and interpersonal factors may play an important role in people's allocation of self-directed gaze. Norton & Pettegrew (1979) found a positive relationship between an individual's self-rating of attentiveness and their gaze behavior directed at other people. Future investigations should systematically explore or control for the possible relationships between individual characteristics (such as personality, introversion/extroversion, attentiveness), group characteristics (familiarity, rapport and prior collaborative experience) and self-gaze behavior. A repeated measures experimental design could address many of these.

5.2 Collaborative Outcomes

There was no difference in the number of unique ideas developed by pairs with small SVDs compared to pairs with large SVDs and no difference in the percentage of common ideas that pairs were able to ground between each SVD group. The finding that SVD size does not influence collaborative performance is unsurprising given the large variation in cooperative dynamics between partners as well as pairs. Past research investigating creative tasks across different forms of video chat and other communication channels suggest that modifying the visual channel provides little advantage over audio when the task is constrained to generating ideas (Teoh et al., 2010). Sellen (1995) hypothesized that this is due to the redundancy of the visual and auditory channels when processing verbal information. Furthermore, (Reinig & Briggs, 2008) have pointed out that the quantity of ideas generated is a limited measure of collaborative performance and often results in diminishing returns with additional ideas.

Clark and Brennan's grounding and communication theory states that a person's unbroken attention towards their partner (and vis versa) affirms that they are being listened to and understood. The process by which participants establish parts of a conversation as mutually shared information is known as grounding (Clarke and Brennan, 1991). Partners' performance on the grounding task was negatively, but not significantly, related to how long participants gazed at their SVD and positively, but not significantly, related to how long they gazed at their partner. Given the large effect size for the negative relationship between how long participants looked at their SVD and the number of the ideas they were able to independently recall following the task, self-gaze may play a role in partners' ability to store ideas generated during the chat and process them afterwards. For economical reasons, the desired power and associated sample size for this study were computed at the level of the conversation participant rather than for pairs, which boosts the likelihood of type II errors for all hypothesis tests testing paired outcomes. Given the large effect sizes recorded for the correlations above, future research investigating the relationship between video chat participant's gaze behavior and grounding should calculate sample size in terms of pairs rather than participants to properly investigate these possible relationships while minimizing the chance for false negatives.

5.3 Gaze Awareness

Regardless of SVD size, participants in both groups overestimated the percentage of the conversation that they gazed at their own SVD as well as the percentage of the conversation that they thought their partner gazed at their corresponding SVD. As for the former finding, it is possible that participants actually were covertly attending to their self-video display for a longer duration than their overt gaze indicated or alternatively, they may have biased their estimates due to the relative salience of their own face in their memory. One possible explanation for this deviation may be that even though video chat participants spend a proportionally small percentage of the chat directly gazing at their self-video, they may still be covertly attending to it. A core assumption underlying all eye tracking research is the premise that foveal gaze continuously corresponds with attention when this is not invariably the case. Indeed, eye trackers "cannot track the covert movements of visual attention (Duchowski, 2007)." As Posner and others originally pointed out, people can willingly dissociate their attention from their foveal direction of gaze. Indeed, astronomers in the night often covertly divert their gaze from center to see dim stars and constellations more clearly by projecting their images onto the parafoveal region of their eyes with the highest density of dim light sensitive rod photoreceptors (Posner et al. 1980; Duchowski, 2002). Wojciulik et al. (1998) found that the fusiform face area - an area of the human brain associated with processing facial stimuli - activates when people voluntarily direct their covert attention towards a face located in the periphery even while they are fixating their gaze on a blank central target and decreases in signal intensity when people covertly attend to other objects. Hence, it is possible that participants' estimates of their overt gaze were actually biased by their covert attention towards their own face on their self-video display. A second possible explanation is that VC participants may cognitively bias their estimate. Past research suggests that facial familiarity can facilitate stronger internal representations of these faces in the observer's memory compared to unfamiliar faces (Althoff and Cohen, 1999; Rizzo et al., 1987). Hence, observer's exposure to their own self-video display may elevate their cognitive recall of their self-directed gaze. These hypothetical mechanisms are not exhaustive and further research is needed to determine which of them - if any - are explanatory and reliable.

Participant's also thought that their partners looked at themselves far more than they actually did, suggesting that video chat configurations may limit people's gaze awareness of their partner's self-monitoring behavior. Participants' overestimation of their partner's self-directed gaze may arise from the parallax between the video camera that records a participant and the screen that depicts their partner. Embedded video cameras are almost universally located above the screen areas that depict a partner's face as well as the participant's own self-video display. The dominance of this configuration is tied to recent and early research that show observers are markedly less sensitive to slight downward deviations in eye contact than when onlookers gaze above or to the side of the observer's eyes (Chen, 2002; Stokes, 1969; Molnar et al., 1969). Despite facilitating the impression of direct gaze when a partner stares within 5 degrees below their camera, the superior placement of the camera can also contribute to the visible impression that participants are avoiding eye contact and possibly looking down at their own SVD when, according to the results above, they actually are more likely to gaze downwards at each other's respective image on their screen below the camera. As discussed above, this downward gaze may further cue participants to follow the depicted direction of their partner's gaze and attend to their own self-video display. Several laboratory studies indicate that perceived averted gaze drives an involuntary shift of covert or overt attention towards the gazed at location (Frischen et al., 2007), although it is still unclear to what extent this factor drives attention during longer computer mediated conversations.

Across both SVD conditions, participants' estimates of how long they looked at their partner and their partner's room did not significantly differ from their measured gaze directed towards both of these areas of interest. This suggests that participants possessed some level of awareness about relatively how long they gazed at their partner and their room compared to the other competing stimuli both on and off screen. Additionally, participants' estimates of how long their partner looked at them also did not significantly differ from their partner's corresponding gaze directed at the participant, regardless of the size of their SVDs. This result is surprising given that the parallax between a participant's depicted partner and the participant's own camera distorts the apparent direction of a partner's gaze. Video chat partners appear to actually possess a level of gaze awareness to estimate relative *how long* their partner looked at their video of the participant. This complements Grayson & Monk's (2003) finding that it is possible for users to learn to infer where their VC partner is looking enough to perceive *when* they are looking at the participant with reasonable accuracy. This accuracy was highest when the video camera was positioned above a participant's video of their partner, a similar orientation to the setup found in the present study.

VC participants significantly and consistently underestimated how long they gazed off screen at their own room in both SVD size conditions. Participant's apparent lack of awareness of how long they looked at their own room confirms that people's retrospective gaze awareness varies depending on the area of interest under consideration. Their low estimates may be linked to the centrally focused video chat task as well as the sparse and unremarkable white walls of the laboratory environment.

5.4 Self-video Display Evaluation

A majority of VC participants found their SVD to be comforting, regardless of its size. Conversely, participants with large SVDs were more likely than participants with small SVDs to rate their SVD as distracting: more than three quarters of all participants with large SVDs perceived their SVD to be distracting compared to slightly less than half of participants with small SVDs. Beyond the bathroom sink, many people seldom converse with their collaborators in face to face settings while simultaneously seeing their own reflected face. Humans are intrinsically fascinated with faces in general and especially their own countenances and may be both comforted by it’s familiarity yet distracted by it’s temporary visibility in video chat. An alternative possible explanation is that by partially obstructing and competing with a participant’s video of their partner, larger SVDs may also impede perceived social presence - the extent to which remote partners feel like they are together. According to social presence theory, different communication media provide participants with a sense of being co-located the closer they are to approaching genuine face-to-face communication.

Although there was no association between the size of participants’ SVD and whether or not they preferred to video chat with their SVD present, over a third of all video chat participants indicated that they would rather video chat without their SVD present. At the time this study was conducted no video chat interfaces allowed you to remove the self-video display from view and at the time of writing this, only one of the primary video chat services provides such an option. Future qualitative research can inquire into why people would prefer to video chat with or without their SVD present.

CHAPTER 6

LIMITATIONS

6.1 Limitations and Future Research

Generalizability

This study only surveyed one specific slice of the global population of video chat users, and an exceptionally WEIRD one at that (Henrich et al., 2010). Future research should consider other conversation partner relationship configurations, additional segments of society, video conference tasks, environmental settings and new video chat technologies. This study controlled for existing relationships by only looking at the cooperative and gaze behavior of strangers with no prior affiliation or direct association - aside from their shared social context as members of the same overarching organization (students at Cornell University). Further research should also investigate how the relationship between video chat partners affects the distribution of individual and interpersonal attention. In office and private settings, the author presumes that much, but certainly not all, of video chat conversations take place between people with already established relationships who have experience cooperating together. Past research suggests that novel faces require more visual attention compared to familiar faces (Althoff and Cohen, 1999). Hence, strangers may draw more of participants' attention than familiar partners, which could mediate how long participant's allocate their gaze across different features of the scene. However, Althoff and Cohen's (1999) study utilized famous faces instead of personally familiar faces and exposure times were brief compared to longer video chat exchanges.

Prospective research can also involve distributed members from existing organizations and companies in future studies. Specifically, it would be fascinating to know whether distributed team members who belong to the same 'in-group' and engage in video conferences on a regular basis, exhibit similar or distinct gaze behavior compared to distributed team members collaborating for the first time. In a brainstorming exercise the cost of a participant's partner not appearing to gaze - and thereby attend - to the participant's face on achieving the goal of the conversation may be less than in other activities. Future studies can consider a variety of VC tasks, circumstances and goals where the cost of perceived inattentiveness may be more consequential, such as interviewing remote candidates. For instance, in a previous study using photos, Amaltfitano & Kalt (1977) found that research participants appraised images of people gazing below the axis of the camera as less deserving of a job than images depicting people staring straight at the camera. As of 2016, most computers' embedded web cams are located above video of one's partner, with the implication that people primarily look below the camera during video chat conversations. Hence, researchers may wish to examine whether applicants that are instructed to look directly into their web camera are rated more or less favorably in terms of attentiveness, credibility and/or intelligence than other interviewees that stare at the depicted image of their interviewer on the screen below.

Video chat dynamics and behaviors between close colleagues, friends, family members and couples may also vary. For instance, existing studies suggest that couples exhibit different video chat use behaviors and characteristics (Rintel, 2013). Beyond one-on-one video chats, one-to-many or many-to-many video chat configurations should also be considered. As Vertegaal (1999) and Suwita et al. (1997) have pointed out, conversational dynamics and structure become more complicated with a greater number of participants and the use of gaze as a turn taking cue is limited by the inability for people to overtly direct their gaze to a particular partner. For instance, looking directly into the camera results in all partners perceiving that the participant is gazing directly at them, while looking at each partner's corresponding image results in horizontal parallax in addition to the vertical parallax found in dyadic video chats (Suwita et al., 1997). Researchers can also investigate whether partner and self-directed gaze behaviors change in multiparty video chats.

Past research indicates that video chat use is lower for females compared to males and declines significantly for those over 50 (Rainie & Zickuhr, 2010). Additional research is needed to investigate whether these findings generalize across the spectrums of gender, age and ethnicity, beyond the female, western, college educated, industrialized, rich and democratic participants involved in this study. Such information can help to ensure that future video chat interfaces are inclusive of the wide spectrum of people interacting with them and help to broaden the user base of video chat.
Given the WEIRD frame of this study, a multi-institutional study could provide a better basis for generalization by recruiting a wider sample of students. Furthermore, such a study could be conducted in multiple countries (with a standardized procedure) to test whether these initial results generalize beyond WEIRD societies. As a parallel pursuit, it would be interesting to study whether partner and self-monitoring behavior during video chat is consistent across different developmental stages and groups including babies, children, adolescents, adults and intergenerational dyads. For instance, despite the American Academy of Pediatrics past recommendation for parents to discourage media exposure for children under two years old (Brown, 2011), a recent survey on video chat

found that 85% of parents surveyed had used video chat with their 6-24 months old children and that over a third of them used it at least once per week (McClure et al., 2015). Given babies' sensitivity to faces, eye contact and gestures - all of which are modified or compromised by video chat - future research can also study to what extent self-video displays affect their allocation of attention when communicating with loved ones.

Scope of Trials

Separate research should also elaborate on the psychophysical approaches of Chen (2002) or Grayson & Monk (2003) to systematically assess whether, and to what extent, people can discriminate when their partner is looking at them, their room, their own self-video display or off-screen. Future investigations of self-video displays in video mediated communication should also consider additional qualitative dimensions of people's subjective appraisal of their own displays and the design dimensions that may influence their interpretations. Additionally, researchers should solicit personal feedback about how people feel when they perceive their partners averting their gaze and looking down in different task scenarios to understand the potential costs associated with not being able to infer where and what one's partner is looking at.

Although this study was conducted over a secure Wi-Fi connection with limited disruptions due to lag, there were a few isolated instances where the video chat image quality, audio quality and/or connection speed faltered. Delays can disrupt the transmission of information between participants and potentially interfere with the process of grounding new ideas and sharing of abstract knowledge (Tscholl, 2005). Chen (2002) found that lower video quality can actually increase the acceptable angular deviation where an observer perceives direct eye contact from their partner. Future studies should record these anomalies by (for instance) plotting connection speed, frames per second and/or other measures of image quality over time. Subgroup sensitivity analyses comparing data during periods of high, medium and low quality may be conducted to scan for possible effects on user gaze behavior. Furthermore, researchers may wish to precisely manipulate the audio and video quality, alignment and connection speed to explore how these different dimensions may modulate video chat quality and possibly affect conversant behavior and collaborative outcomes. In this case, conversation partners can be stratified according to different levels of video quality.

The video chat interface dimensions of window size and self-video display presence should also be considered in future studies. As Grayson & Monk (2003) point out, some people video chat with a reduced video size to minimize bandwidth and or to enable multi-tasking while video chatting. Future investigations should characterize how varying the size of a participant's video of their partner in relation to the participant's self-video display influence partner directed and self-directed gaze. While the VC service (Skype) used in this study allowed users to manually select from one of two SVD size options, it also prevented users from being able to remove the SVD altogether. Researchers may wish to utilize other video chat services that allow for this condition to be evaluated. They can also consider the ever expanding list of other video chat services, operating systems and devices, screen sizes and orientations all of which present variations to the classic video chat interface. The world continues to change since this study was conducted: more people have access to video chat via personal computers and mobile devices, often communicating with each other via video while on the go. Mobile video services like Apple FaceTime and Facebook Messenger are growing in global use and exposure with the advent of smart phones. Future studies can experimentally examine whether proximity to the camera, camera field of view, embedded camera angle and subtended visual face angle impact individual and inter-personal gaze behavior. Smaller screen sizes (laptops, tablets, mobile phones and smart watches) especially need to be studied given that the visual angles between the camera, a partner's face on the screen and the self-portrait is likely smaller (Bekkering and Shim, 2006).

Environmental setting characteristics may also influence social and individual video chat dynamics and gaze behavior. While office personalization and background setting characteristics were controlled for in this study, future studies may inquire whether ambient background dimensions such as noise, lighting, temperature, wall color, occupancy level may influence gaze allocation and collaborative performance.

Imprecision

Efforts can be taken to improve the precision of measurement. Notably, the Mirametrix equipment used in this study has diminished gaze recording accuracy towards the periphery of the screen. This resulted in reduced gaze resolution over the self-video display area which may have become worse over the course of each 15-minute conversation. This study used a static eye tracking system that requires users to maintain a consistent posture and proximity range within the device's effective field of view. Participants may have moved after calibration while their partners were being calibrated or shifted during the discussion. Future studies should ensure that there is not a posture by video chat interface interaction on individual and social gaze behavior. One alternative is to use a mobile eye tracking system in future research to allow people to adjust to relaxed and dynamic postures during the course of conversation. However, researchers need to balance the improved precision of measurement offered by some mobile eye tracking systems with the likely influence such technologies may exert on individual and interpersonal gaze behavior (Duchowski, 2007). Humans are remarkably sensitive to changes in their facial appearance, so the headgear of numerous mainstream mobile eye tracking systems will introduce their own distractions and bias gaze allocation.

Potential Bias

Due to ethical guidelines preventing a truly blind experimental design, participants in this study were informed that measures about their posture and eye position were being collected and thus may have become aware of the eye tracking apparatus. Others have found that the presence of eye tracking equipment does not influence participants self-reported measurements of gaze awareness (Albert & Tedesco, 2010). For this study we assumed that both estimated and recorded measurements of gaze were similarly unaffected by the participant’s awareness of the eye tracking apparatus. However, it is possible that knowledge of the eye tracking apparatus in conjunction with the video chat task may have modified participants’ motivation and subsequent overt gaze behavior given the known influence of the instructed goals and tasks in an experiment on gaze patterns (Althoff & Cohen, 1999). This study used a fixed chair and static desk. Future research may benefit from controlling for head location by using an adjustable desk to ensure that the eye height and head location is consistent as captured by the video camera and represented on a partner’s screen. Alternatively, the resulting variation in parallax can be included as a predictor for different objective and subjective outcomes. This study controlled for prior relationship and gender, but did not control for sexual orientation. Future studies may wish to include sexual orientation as a separate covariate. Beyond just controlling for use of eye contacts and/or glasses, future studies may benefit from assessing participant’s eye sight. For instance, it is conceivable that the smaller self-video display may have posed problems for anyone with even slightly reduced visual acuity. Although Likert scales measuring different subjective constructs and dimensions are common in human factors and UX research, the subjective evaluation questions for perceived comfort, distraction and preference that were used in this study were not previously validated and thus only provide a preliminary comparison of the two self-video display sizes in terms of perceived comfort and distraction (Gliem & Gliem, 2003). Further validation of these questions is needed.

CHAPTER 7

CONCLUSIONS

7.1 Summary of Contributions

This thesis provides preliminary evidence that the size of the self-video display can influence video chat participants' visual attention and gaze awareness and modulate their self-directed gaze. Even though researchers in visual perception, human factors, human computer interaction and user experience typically employ eye tracking technology to measure individual gaze, this thesis is also the first such study to employ two eye tracking setups simultaneously to measure interpersonal visual attention between more than one person during computer mediated communication. Rather than relying on users to adjust their behavior or modify their faces to accommodate our video chat interfaces (see Considine, 2012), video chat UIs can be designed to promote user comfort and satisfaction, while also seeking to limit personal and interpersonal distractions imposed by the self-video display. This study suggests that minimizing the visual size of self-video can help to reduce, but not eliminate, self-directed gaze as well as perceived distraction. However, participants appear to lack gaze awareness of how long their partners stare at their corresponding self-video display regardless of self-video display size and participants think their partners gaze at themselves significantly more than they actually do. The author hypothesizes that this is likely linked to the spatial configuration of most video chat systems, where video chat cameras are placed on the top of the screen and self-video displays are positioned on the bottom. This arrangement results in the deceptive appearance that one's partner is looking down and averting their gaze from the participant regardless of whether the partner is actually looking at the participant on their screen, staring at their own self-video display or looking below the axis of the camera on or off the screen.

7.2 Practical Implications for Cooperative Video Chat Interfaces

Comparably more evidence exists for guiding the design of video chat hardware than for video chat interfaces. Multiple studies have suggested that the parallax between the position of the videoconferencing camera and depiction of a participant's partner should be minimized to improve the transmission of eye contact and other non-verbal gaze cues (Suwita, 1997; Mühlbach et al. 1995). Echoing Stokes (1969), Chen (2002) recommends setting the visual angle between the video chat camera and the depicted eyes to be less than 5 degrees, assuming the video window depicting one's partner is placed directly below the camera. Nevertheless, these rules are frequently violated especially on desktops and laptops with large visual angles.

This is the first study to experimentally examine the influence of the self-video display on non-verbal communication and shows that people overestimate how much they, and their partners, look at their self-video display following a cooperative task. People with large video displays only gazed at themselves for a small (3%) fraction of the conversation duration and predominantly found their displays to be distracting. Video chat designers and researchers should critically evaluate and explore other SVD interface dimensions beyond size, including but not limited to placement and persistence as well as combinations of these dimensions to accommodate participant's intentional and attentional demands over the video chat. For instance, placing the SVD at the top of the screen, closer to the camera, can foster the appearance of direct gaze but may further confuse participants' gaze awareness of whether their partners are looking at the participant or themselves. Past research indicates that partner-directed gaze is not equal over the course of a conversation and that people tend to gaze directly at their partner at the beginning and end of an exchange (Levine and Sutton Smith, 1973; Kendon, 1967). Similarly, self-gaze may support user's personal and interpersonal demands to varying degrees during a video chat: participants may wish to check in with themselves more or less at different stages of the conversation. Fish et al.'s (1990) study on a large video conferencing system with no self-video display found that people frequently assumed that if they could view their partner(s) on the video screen their partner(s) could see them, whether or not they actually visible to the camera. Hence, the period preceding and at the beginning of a video call may be an important time for participants to view their self-video display to adjust their web camera, frame their image, attend to their appearance or calibrate their connection but may not afford as much utility later on. Looking backwards to move forwards, UI designers may also derive inspiration from the original stated purpose of the 'Vu Self' button on the original Picturephone Mod II which enabled participants to manually switch their self-video on or off to confirm that they were adequately centered in the camera or monitor their appearance before or during the conversation (Noll, 1992). Indeed, over a third of people surveyed in this study indicated that they would prefer to video chat without a self-video display present. Designers should allow people to enhance or reduce the saliency of their self-video display by either enabling them to remove their SVD, reduce the size of it, change the position or change the opacity. However, a cursory review of today's video chat technologies suggests that the ability to manually toggle the self-video display off and on is still the exception and not the norm: the majority of video chat services don't allow people to remove their self-video displays. Finally, even though self-video display size does not appear to significantly affect the capacity of two partners to ground and recall each other's ideas after the fact, it is clear that participants' ability to do so is quite limited irrespective of self-video size. Given that current video chat tools do not support recording, reviewing or revising verbal communication messages, future collaborative video chat tools may explore adding automated transcription services to build common ground around the ideas they generated during the conversation.

REFERENCES

Acker, S., Clift, C., & Branco, S. (1985). A two-way telecourse, interactive in both audio and video channels: A model for inter-institutional cooperation. Issues in Higher Education, 14, 16-23.

Acker, S. R., & Levitt, S. R. (1987). Designing videoconference facilities for improved eye contact. Journal of Broadcasting & Electronic Media, 31(2), 181-191.

Adolphs, R. (2009). The social brain: neural basis of social knowledge. Annual review of psychology, 60, 693.

Agius, H., & Angelides, M. (1997). Desktop video conferencing in the organisation. Information & management, 31(6), 291-302.

Albert, R. (1893). Le Vingtième Siècle: La Vie Électrique/Albert Robida. Paris: Librairie illustrée.

Albert, W., & Tedesco, D. (2010). Reliability of self-reported awareness measures based on eye tracking. Journal of Usability Studies, 5(2), 50-64.

Allison, T., Ginter, H., McCarthy, G., Nobre, A., Puce, A., Luby, M., & Spencer, D. (1994). Face recognition in human extrastriate cortex. Journal of neurophysiology, 71(2), 821-825.

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: a reprocessing effect in face perception. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25(4), 997.

Amalfitano, J. G., & Kalt, N. C. (1977). Effects of eye contact on the evaluation of job applicants. Journal of Employment Counseling, 14(1), 46-48.

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze. *Perception*, *31*(6), 657-674.

Argyle, M., & Cook, M. (1976). Gaze and mutual gaze.

Ballagas, R., Kaye, J., Ames, M., Go, J., & Raffle, H. (2009). Family communication: phone conversations with children. In Proceedings of the 8th international Conference on Interaction Design and Children (pp. 321-324). ACM.

Barton, J. J., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. Perception, 35(8), 1089-1105.

Beattie, G. W. (1979). Contextual constraints on the floor-apportionment function of speaker-gaze in dyadic conversations. British Journal of Social & Clinical Psychology.

Bekaert P., Dumont M., Maesen S., & Rogmans S. (2008). A prototype for practical eye-gaze corrected video chat on graphics hardware. *SIGMAP*, 236–243.

Bekkering, T. J. E. (2004). Visual angle in videoconferencing: The issue of trust (Doctoral dissertation, Mississippi State University. Management and Information Systems Department.).

Bekkering, E., & Shim, J. P. (2006). Trust in videoconferencing. Communications of the ACM, 49(7), 103-107.

Biello, D. (2009). Can videoconferencing replace travel? Scientific American.

Brancatelli, J. (1985). The problem with teleconferencing. Frequent Flier Magazine, January, 61-64.

Broadbent, D. E. (1958). The effects of noise on behaviour.

Brothers, L. (2002). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. Foundations in social neuroscience, 367-385.

Brown, A. (2011). Media use by children younger than 2 years. Pediatrics, 128(5), 1040-1045.

Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., & Young, A. W. (2002). Reading the mind from eye gaze. *Neuropsychologia*, *40*(8), 1129-1138.

Carey, G. R. (1880, June 5). Seeing by Electricity. Scientific American. Retrieved November 1, 2016, from http://histv2.free.fr/carey/carey2.htm

Carpenter, R. (1988). Movements of the eyes (2nd rev. Pion Limited.

Carpenter, R. H. S. (1977). Movements of the Eyes. London: Pion.

Chapanis, A., Parrish, R. N., Ochsman, R. B., & Weeks, G. D. (1977). Studies in interactive communication: II. The effects of four communication modes on the linguistic performance of teams during cooperative problem solving. Human Factors: The Journal of the Human Factors and Ergonomics Society, 19(2), 101-126.

Chapanis, A., Ochsman, R. B., Parrish, R. N., & Weeks, G. D. (1972). Studies in interactive communication: I. The effects of four communication modes on the behavior of teams during cooperative problem-solving. Human Factors: The Journal of the Human Factors and Ergonomics Society, 14(6), 487-509.

Chen, M. (2002, April). Leveraging the asymmetric sensitivity of eye contact for videoconference. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 49-56). ACM.

Clark, H., & Brennan, S. (1991). Grounding in communication. Perspectives on socially shared cognition, 13(1991), 127-149.

Clark, H., & Schaefer, E. (1989). Contributing to discourse. Cognitive science, 13(2), 259-294.

Cook, M., & SMITH, J. (1975). The role of gaze in impression formation. British Journal of Social and Clinical Psychology, 14(1), 19-25.

Considine, A. (2012, April 17). Ready for My Video Chat Close-Up. New York Times. Retrieved October 04, 2015, from http://www.nytimes.com/2012/04/19/fashion/ready-for-my-video-chat-close-up.html

Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia*, *45*(13), 3024-3037.

Coppola, D., & Purves, D. (1996). The extraordinarily rapid disappearance of entopic images. Proceedings of the National Academy of Sciences, 93(15), 8001-8004.

Coroama, V., Hilty, L., & Birtel, M. (2012). Effects of Internet-based multiple-site conferences on greenhouse gas emissions. Telematics and Informatics, 29(4), 362-374.

Crane, H. D. (1994). The Purkinje image eyetracker, image stabilization, and related forms of stimulus manipulation. Visual science and engineering: Models and applications, 15-89.

CU-Schools [Photograph found in Wikimedia]. (1997). In Wikimedia Commons. Retrieved November 1, 2016, from https://upload.wikimedia.org/wikipedia/en/c/cf/CU-Schools.GIF

Denstadli, J. (2004). Impacts of videoconferencing on business travel: the Norwegian experience. Journal of Air Transport Management, 10(6), 371-376.

Deutsch, J. A., & Deutsch, D. (1963). Attention: some theoretical considerations. Psychological review, 70(1), 80.

Dickson, E., & Bowers, R. (1973). The video telephone, a new era in telecommunications: a preliminary technology assessment. Program on Science, Technology and Society, Cornell University.

Diehl, R. (1972). AT&T 'Picturephone' Mod II of 1972-73, -Control Pad, close-up view [Photograph found in Wikimedia Commons]. In Wikimedia Commons. Retrieved November 1, 2016, from https://commons.wikimedia.org/wiki/File:MyPicturephone_031_-AT&T_Picturephone_Mod_II_Control_Pad-.jpg (Originally photographed 1972)

Diehl R. (1969). Postwar American Television: AT&T Model 2 Picturephone. Retrieved November 1, 2016, from http://www.earlytelevision.org/picturephone.html

Dodge, R. (1907). An experimental study of visual fixation. The Psychological Review: Monograph Supplements, 8(4), i.

Dorcey, T. (1995). CU-SeeMe desktop videoconferencing software. Connexions, 9(3), 42-45.

Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. Behavior Research Methods, Instruments, & Computers, 34(4), 455-470.

Duchowski, A. (2007). *Eye Tracking Methodology Theory and Practice*. New York: Springer. 2007.

Duncan, S., & Fiske, D. W. (2015). Face-to-face interaction: Research, methods, and theory (Vol. 3). Routledge.

Electrician. (1877, March 29). The Electroscope: Is this Marvel to be the Next Achievement of Modern Science? New York Sun. Retrieved November 1, 2016 from http://histv2.free.fr/19/electroscope.htm

Ellyson, S., Dovidio, J. F., & Corson, R. L. (1981). Visual behavior differences in females as a function of self-perceived expertise. Journal of Nonverbal Behavior, 5(3), 164-171.

Ellyson, S. L., Dovidio, J. F., Corson, R. L., & Vinicur, D. L. (1980). Visual dominance behavior in female dyads: Situational and personality factors. Social Psychology Quarterly, 328-336.

Elton, M. C. (1985). Visual communication systems: trials and experiences. Proceedings of the IEEE, 73(4), 700-705.

Egido, C. (1988). Video conferencing as a technology to support group work: a review of its failures. In Proceedings of the 1988 ACM conference on Computer Supported Cooperative Work (pp. 13-24). ACM.

Engelbart, D., & English, W. (1968). A research center for augmenting human intellect. In Proceedings of the December 9-11, 1968, fall joint computer conference, part I (pp. 395-410). ACM.

Fagen, M. D., Joel, A. E., O'Neil, E. F., & Schindler, G. E. (1975). A History of Engineering and Science in the Bell System: Transmission Technology (1925-1975) (Vol. 7). The Laboratories.

Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience*, *16*(8), 1320-1326.

Farroni, T., et al. (2002) Eye contact detection in humans from birth. Proc. Natl. Acad. Sci. U. S. A. 99, 9602-9605

Fish, R. S., Kraut, R. E., & Chalfonte, B. L. (1990, September). The VideoWindow system in informal communication. In Proceedings of the 1990 ACM conference on Computer-supported cooperative work (pp. 1-11). ACM.

Fitzgerald, T. (2003). For the Deaf, Communication without the Wait. New York Times.

Forbes, R. J., & Jackson, P. R. (1980). Non-verbal behaviour and the outcome of selection interviews. Journal of Occupational Psychology, 53(1), 65-72.

Fox S., & Jones S. (2009). *Generations online in 2009*. Pew Internet and American Life Project. Retrieved December 12, 2011, from http://www.pewinternet.org/Reports/2009/Generations-Online-in-2009.aspx/

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic bulletin & review*, *5*(3), 490-495.

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 319.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin*, *133*(4), 694.

Gamer, M., & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 705.

Gemmell, J., Toyama, K., Zitnick, C., Kang, T., and Seitz, S. (2000). Gaze awareness for videoconferencing: A software approach. IEEE MultiMedia, 7(4), 26–35.

George, N., Driver, J., & Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *Neuroimage*, *13*(6), 1102-1112.

Goodwin, C. (1981). Conversational organization: Interaction between speakers and hearers. Academic Press.

Gibson, J. (1979). The ecological approach to visual perception.

Gibson, J. (1941). A critical review of the concept of set in contemporary experimental psychology. Psychological bulletin, 38(9), 781.

Gliem, R. R., & Gliem, J. A. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education.

Grayson, D., & Monk, A. (2003). Are you looking at me? Eye contact and desktop video conferencing. ACM Transactions on Computer-Human Interaction (TOCHI), 10(3), 221-243.

Greer, J., & Regueiro, M. (2016). Teleconferences to Facilitate Multidisciplinary Care and Education in IBD. In Telemanagement of Inflammatory Bowel Disease (pp. 101-116). Springer International Publishing.

Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006, April). The validity of the stimulated retrospective think-aloud method as measured by eye tracking. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 1253-1262). ACM.

Guernsey, L. (2000, April 13). Cautionary Tale; The Perpetual Next Big Thing. New York Times. Retrieved October 25, 2016, from http://www.nytimes.com/2000/04/13/technology/cautionary-tale-the-perpetual-next-big-thing.html

Hawkins, W. (1988). For-your-home video phones. Popular Science, 232.

Helmholtz, H. (1925). Physiological optics. Optical Society of America, 3, 318.

Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world?. Behavioral and brain sciences, 33(2-3), 61-83.

Hoffman, E., & Haxby, J. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nature neuroscience, 3(1), 80-84.

Hood, B. M., Macrae, C. N., Cole-Davies, V., & Dias, M. (2003). Eye remember you: The effects of gaze direction on face recognition in children and adults. Developmental Science, 6(1), 67-71.

How do we communicate in business today? (2010). Retrieved October 26, 2016, from http://www.plantronics.com/us/howwework/

Hunter, G. (1980). Teleconference in Virtual Space. In IFIP Congress (pp. 1045-1048).

Irwin, D. (1992). Visual memory within and across fixations. In Eye movements and visual cognition (pp. 146-165). Springer New York.

James, W. (1984). Psychology, briefer course (Vol. 14). Harvard University Press.

James, W. (1981). The Principles of Psychology (Vol. I). Cambridge, MA: Harvard University Press.

James, W. (1890). The Principles of Psychology. New York: H. Holt and Co.

Jeffreys, H. (2014, July 22). Hide Self-video on Video Call. Skype Community Discussions. Retreived November 4th, 2016, from https://community.skype.com/t5/Android-archive/Hide-self-video-on-video-call/td-p/3380721

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of faces and its subsequent decline.Cognition, 40, 1–19.

Johansen, S. A., & Hansen, J. P. (2006). Do we need eye trackers to tell where people look?. In CHI'06 Extended Abstracts on Human Factors in Computing Systems (pp. 923-928). ACM.

Johansen, R. (1984). Teleconferencing and Beyond: Communications in the Office of the Future. McGraw-Hill, Inc.

Kampe, K. K., Frith, C. D., & Frith, U. (2003). "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *The Journal of neuroscience*, *23*(12), 5258-5263.

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., ... & Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring. *Brain*, *122*(4), 779-783.

Kelly, E. W. (1978). Effects of counselor's eye contact on student–clients' perceptions. Perceptual and motor skills.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. Acta psychologica, 26, 22-63.

Kleinke, C. L. (1986). Gaze and eye contact: a research review. Psychological bulletin, 100(1), 78.

Kleinke, C. L., Meeker, F. B., & La Fong, C. (1974). Effects of gaze, touch, and use of name on evaluation of “engaged” couples. Journal of Research in Personality, 7(4), 368-373.

Kleinke, C. L., Bustos, A. A., Meeker, F. B., & Staneski, R. A. (1973). Effects of self-attributed and other-attributed gaze on interpersonal evaluations between males and females. Journal of experimental social Psychology, 9(2), 154-163.

Kleinke, C. L., Staneski, R. A., & Pipp, S. L. (1975). Effects of gaze, distance, and attractiveness on males' first impressions of females. Representative Research in Social Psychology.

Kleinke, C. L., Staneski, R. A., & Berger, D. E. (1975). Evaluation of an interviewer as a function of interviewer gaze, reinforcement of subject gaze, and interviewer attractiveness. Journal of Personality and Social Psychology, 31(1), 115.

Kobayashi, H., and Kohshima, S. (1997) Unique morphology of the human eye. Nature 387, 767-768

Kristiansen, T. (Ed.). (1991). A Window to the Future: The Videotelephone Experience in Norway. Norwegian Telecom Research Department.

Kuster, C., Popa, T., Bazin, J. C., Gotsman, C., & Gross, M. (2012). Gaze correction for home video conferencing. ACM Transactions on Graphics (TOG), 31(6), 174.

Lazzerini, A. J., Stephenson, G. M., & Neave, H. (1978). Eye-contact in dyads: A test of the Independence Hypothesis. British Journal of Social and Clinical Psychology, 17(3), 227-229.

Leigh, R., & Zee, D. (1991). Smooth pursuit and visual fixation. The neurology of eye movements, Ed, 2, 139-166.

Lenhart, A. (2010). Cell phones and American adults: they make just as many calls, but text less often than teens. Washington, DC: Pew Research Center< http:// pewinternet. org/Reports/2010/Cell-Phones-and-American-Adults. aspx.

Levin, D. T., & Simons, D. J. (2000). Perceiving stability in a changing world: combining shots and intergrating views in motion pictures and the real world. Media Psychology, 2(4), 357-380.

Levine, M. H., & Sutton-Smith, B. (1973). Effects of age, sex, and task on visual behavior during dyadic interaction. Developmental Psychology, 9(3), 400.

Macknik, S. L., King, M., Randi, J., Robbins, A., Thompson, J., & Martinez-Conde, S. (2008). Attention and awareness in stage magic: turning tricks into research. *Nature Reviews Neuroscience*, *9*(11), 871-879.

Macrae, C.N., et al. (2002) Are you looking at me? Eye gaze and person perception. Psychological Science. 13, 460-464

Martinez-Conde, S., Macknik, S., & Hubel, D. (2004). The role of fixational eye movements in visual perception. Nature Reviews Neuroscience, 5(3), 229-240.

Maurer, D., & Barrera, M. (1981). Infants' perception of natural and distorted arrangements of a schematic face. Child Development, 47, 523–527.

Maurier, G. (1878). Punch magazine. December 9th.

McCanne, S., & Jacobson, V. (1995). vic: A flexible framework for packet video. In Proceedings of the third ACM international conference on Multimedia (pp. 511-522). ACM.

McCarthy, G., Puce, A., Gore, J., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. Journal of cognitive neuroscience, 9(5), 605-610.

McClure, E., Chentsova-Dutton, Y., Barr, R., Holochwost, S., & Parrott, W. (2015). "Facetime doesn't count": Video chat as an exception to media restrictions for infants and toddlers. International Journal of Child-Computer Interaction, 6, 1-6.

Mee, A. (1898). The pleasure telephone. The Strand Magazine, 16(44), 339-345.

Mirametrix Research. (2010). Application Programming Interface. Retrieved July 15, 2016 from www.mirametrix.com

Molnar, J., Dorros, I., Davis, C., Harris J., Williams, R., Korn, F., Ritchie, A., Nast, D., Welber, I., Andrews, F., Hardaway, H., Ekstrand, S., Spencer, A. (1969). Picturephone service-a new way of communicating. Bell Laboratories Record, 47(5), 134.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two process theory of infant face recognition. Psychological Review, 98, 164–181.

Mühlbach, L., Bocker, M., & Prussog, A. (1995). Telepresence in videocommunications: A study on stereoscopy and individual eye contact. Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(2), 290-305.

Naiman, T. H., & Breed, G. (1974). Gaze duration as a cue for judging conversational tone. Representative Research in Social Psychology.

Nakano, Y. I., Reinstein, G., Stocky, T., & Cassell, J. (2003). Towards a model of face-to-face grounding. In Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1 (pp. 553-561). Association for Computational Linguistics.

Noll, A. M. (1992). Videophone: A flop that won't die. New York Times, 12.

Noll, A. M. (1992). Anatomy of a failure: Picturephone revisited. Telecommunications policy, 16(4), 307-316.

Norton, R. W., & Pettegrew, L. S. (1979). Attentiveness as a style of communication: A structural analysis. Communications Monographs, 46(1), 13-26.

Noton, D., & Stark, L. (1971a). Eye movements and visual perception. Scientific American.

Noton, D., & Stark, L. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. Vision research, 11(9), 929-IN8.

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in cognitive sciences*, *13*(3), 135-143.

Ochsman, R. B., & Chapanis, A. (1974). The effects of 10 communication modes on the behavior of teams during co-operative problem-solving. International Journal of Man-Machine Studies, 6(5), 579-619.

Ohya, J., Kitamura, Y., Takemura, H., Kishino, F., & Terashima, N. (1993). Real-time reproduction of 3D human images in virtual space teleconferencing. In Virtual Reality Annual International Symposium, 1993. IEEE (pp. 408-414). IEEE.

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. Neuropsychologia, 45(1), 75-92.

Pernice, K., & Nielsen, J. (2009). Eyetracking methodology: How to conduct and evaluate usability studies using eyetracking. Nielsen Norman Group Technical Report.

Peters C. (1938). Talks on ‘See-Phone’: television applied to German telephones enables speakers to see each other. The New York Times. Sept 18.

Pageler, N. M., Menon, V., Merin, N. M., Eliez, S., Brown, W. E., & Reiss, A. L. (2003). Effect of head orientation on gaze processing in fusiform gyrus and superior temporal sulcus. *Neuroimage*, *20*(1), 318-329.

Pegna, A. J., Khateb, A., Michel, C. M., & Landis, T. (2004). Visual recognition of faces, objects, and words using degraded stimuli: where and when it occurs. Human Brain Mapping, 22(4), 300-311.

Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When strangers pass processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological science*, *15*(9), 598-603.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. Journal of experimental psychology: General, 109(2), 160.

Puce, A., Allison, T., Asgari, M., Gore, J., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. The Journal of Neuroscience, 16(16), 5205-5215.

Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. Journal of neurophysiology, 74(3), 1192-1199.

Rainie, L., & Zickuhr K. (2010). Video Calling and Video Chat. Pew Internet and American Life Project. Retrieved December 20, 2011, from http://pewinternet.org/Reports/2010/Video-chat.aspx

Reinig, B., & Briggs, R. (2008). On the relationship between idea-quantity and idea-quality during ideation. Group Decision and Negotiation, 17(5), 403-420.

Riggs, L., Merton, P., & Morton, H. (1974). Suppression of visual phosphenes during saccadic eye movements. Vision research, 14(10), 997-1011.

Riggs, L., Ratliff, F., Cornsweet, J., & Cornsweet, T. (1953). The disappearance of steadily fixated visual test objects. JOSA, 43(6), 495-501.

Rintel, S. (2013). Video calling in long-distance relationships: The opportunistic use of audio/video distortions as a relational resource. The Electronic Journal of Communication/La Revue Electronic de Communication (Special issue on Videoconferencing in practice: 21st century challenges), 23(1).

Rizzo, M., Hurtig, R., & Damasio, A. R. (1987). The role of scanpaths in facial recognition and learning. Annals of neurology, 22(1), 41-45.

Robinson, D. A. (1968). The oculomotor control system: A review. Proceedings of the IEEE, 56(6), 1032-1049.

Rose, D., & Clarke, P. (1995). A review of eye-to-eye videoconferencing techniques. BT technology journal, 13(4), 127-131.

Sallio, P., Boudeville, J. P., and Cabon, P. (1982). Visual ergonomics of a future videoconference service. In Proceedings of the 10th International Symposium on Human Factors in Telecommunications (pp. 10-16).

Sattler, M. (1995). Internet TV with CU-See Me. Sams.

Sato, W., Yoshikawa, S., Kochiyama, T., & Matsumura, M. (2004). The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. *Neuroimage*, *22*(2), 1006-1013.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718-730.

Schnaars, S., & Wymbs, C. (2004). On the persistence of lackluster demand—the history of the video telephone. Technological Forecasting and Social Change, 71(3), 197-216.

Sellen, A. J. (1995). Remote conversations: The effects of mediating talk with technology. Human-computer interaction, 10(4), 401-444.

Senju, Atsushi and Johnson, Mark H. (2009) The eye contact effect: mechanisms and development. Trends in Cognitive Sciences 13 (3), pp. 127-134. ISSN 1364-6613.

Senju, A., et al. (2008) Is anyone looking at me? Direct gaze detection in children with and without autism. Brain Cogn. 67, 127-139

Senju, A., and Hasegawa, T. (2005) Direct gaze captures visuospatial attention. Vis. Cogn. 12, 127 - 144

Senju, A., et al. (2005) Does perceived direct gaze boost detection in adults and children 9 with and without autism? The stare-in-the-crowd effect revisited. Vis. Cogn. 12, 1474 - 1496

Short, J., Williams, E., & Christie, B. (1976). The social psychology of telecommunications.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception, 28(9), 1059-1074.

Simons, D.J. (2000). Current approaches to change blindness. Visual Cognition, 7, 1–15.

Smith, T. J., & Henderson, J. M. (2008). Edit Blindness: The relationship between attention and global change blindness in dynamic scenes. Journal of Eye Movement Research, 2(2).

Snowden, R., Snowden, R., Thompson, P., & Troscianko, T. (2012). Basic vision: an introduction to visual perception. Oxford University Press.

Sodikoff, C., Firestone, I. J., & Kaplan, K. J. (1974). Subject self-disclosure and attitude change as a function of interviewer self-disclosure and eye-contact. Personality and Social Psychology Bulletin, 1, 243-246.

Stokes, R. (1969). Human Factors and Appearance Design Considerations of the Mod II PICTUREPHONE&# 174; Station Set. IEEE Transactions on Communication Technology, 17(2), 318-323.

Stolberg, S. (2009). Obama Makes History in Live Internet Video Chat. New York Times, 26.

Suwita, A., Böcker, M., Mühlbach, L., & Runde, D. (1997). Overcoming human factors deficiencies of videocommunications systems by means of advanced image technologies. Displays, 17(2), 75-88.

Teoh, C., Regenbrecht, H., & O'Hare, D. (2010, November). Investigating factors influencing trust in video-mediated communication. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction* (pp. 312-319). ACM.

Tomasello, M. (2010). Origins of human communication. MIT press.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: Evidence from visualsearch. Journal of Experimental Psychology: Human Perception and Performance, 25, 1–20.

Treisman, A. (1986). Features and objects in visual processing. Scientific American, 255(5), 114-125.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive psychology, 12(1), 97-136.

Treisman, A. (1961). Attention and speech. Unpublished doctoral thesis, Oxford University.

Trotta, A. (2012, November 19). Earth-to-Space Video Chat Inspires Students to Reach for the Stars! Retrieved October 18, 2016, from https://www.nasa.gov/audience/foreducators/earth-to-space-video-chat.html

Tscholl, M., McCarthy, J., & Scholl, J. (2005, May). The effect of video-augmented chat on collaborative learning with cases. In Proceedings of th 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years! (pp. 682-686). International Society of the Learning Sciences.

Turletti, T., & Huitema, C. (1996). Videoconferencing on the Internet. IEEE/ACM Transactions on networking, 4(3), 340-351.

Tyson, J. (1987). Video teleconferencing is cutting costs while increasing the productivity of many companies. Communication8 New8, 24.

Vertegaal, R. (1999, May). The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 294-301). ACM.

Video chat [Def. 1]. (n.d.). Oxford Dictionaries, Retrieved July 17th, 2016, from http://www.oxforddictionaries.com/us/definition/american_english/video-chat.

Virilio P. (1994). The Vision Machine. London: British Film Institute, 63.

Wheeler, R. W., Baron, J. C., Michell, S., & Ginsburg, H. J. (1979). Eye contact and the perception of intelligence. Bulletin of the Psychonomic Society, 13(2), 101-102.

Wicker, B., Perrett, D. I., Baron-Cohen, S., & Decety, J. (2003). Being the target of another's emotion: a PET study. *Neuropsychologia*, *41*(2), 139-146.

Yarbus, A. (1967). Eye movements and vision. 1967. New York.

Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. Behavior research methods & instrumentation, 7(5), 397-429.

APPENDIX A: FIGURES

Figure A1.

Statistical Power

G*power Sample Size Calculation and Power Specification Output

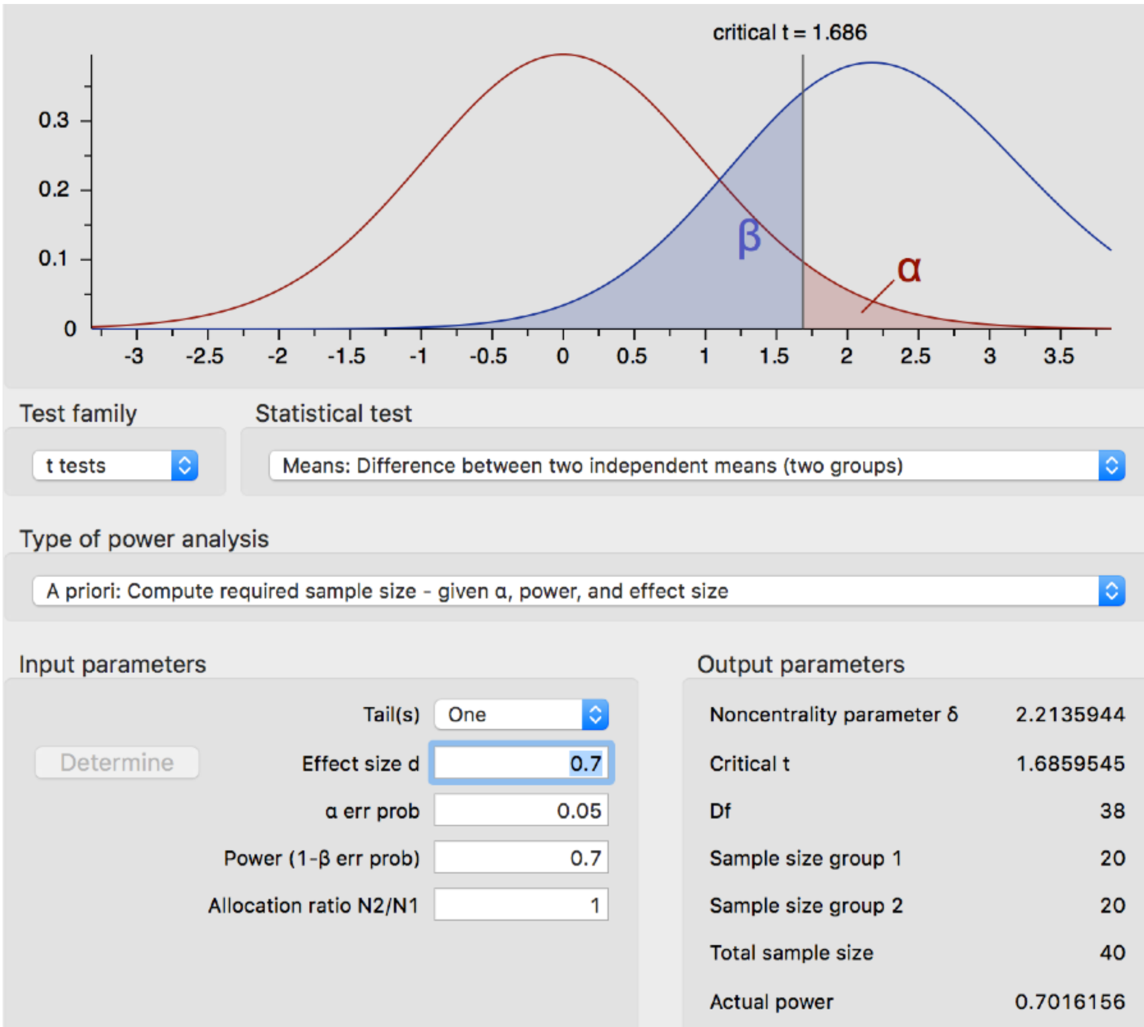


Figure A2

Randomization Procedure

R code for creating random participant IDs

```
pairs<-sample(1:40)
```

R code for forming random pairs of participants (where each row in the matrix is one pair):

```
randompairs<- matrix(pairs, ncol=2, byrow=TRUE)
```

R code for randomly assigning each pair to either a medium or small VC self-video display size condition:

```
SVDassignment<- sample(1:20)
```

Figure A3

Experimental Setup

Dell 19" (1280 x 1024) Display Monitor

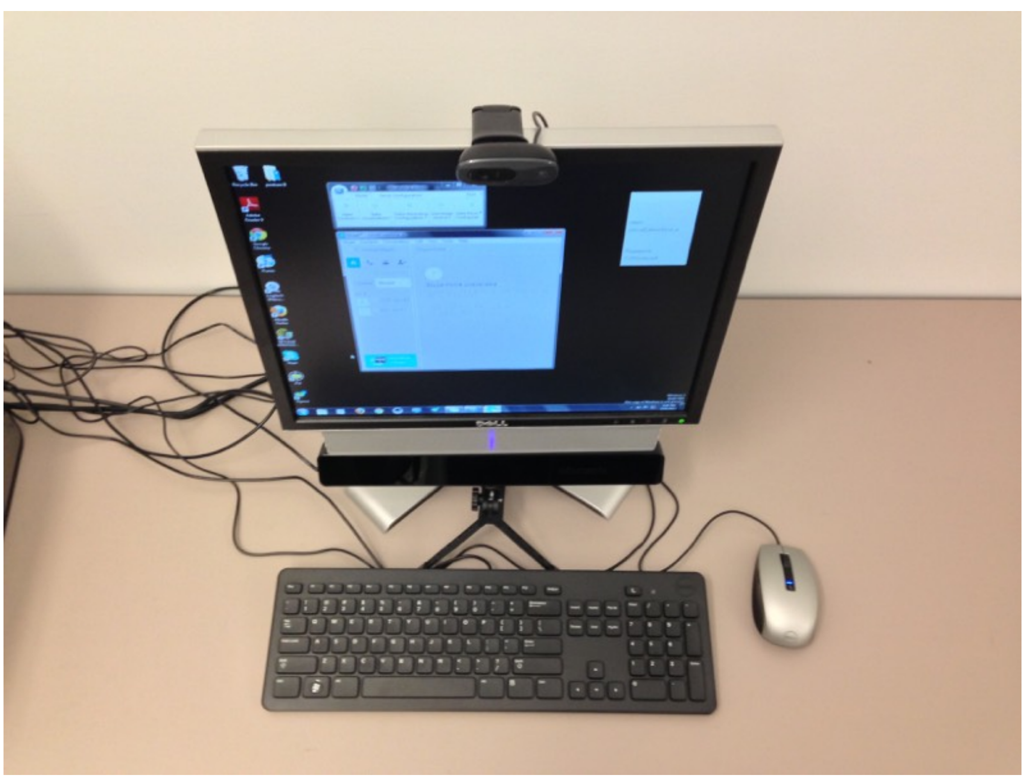


Figure A4

Skype Video Chat Service

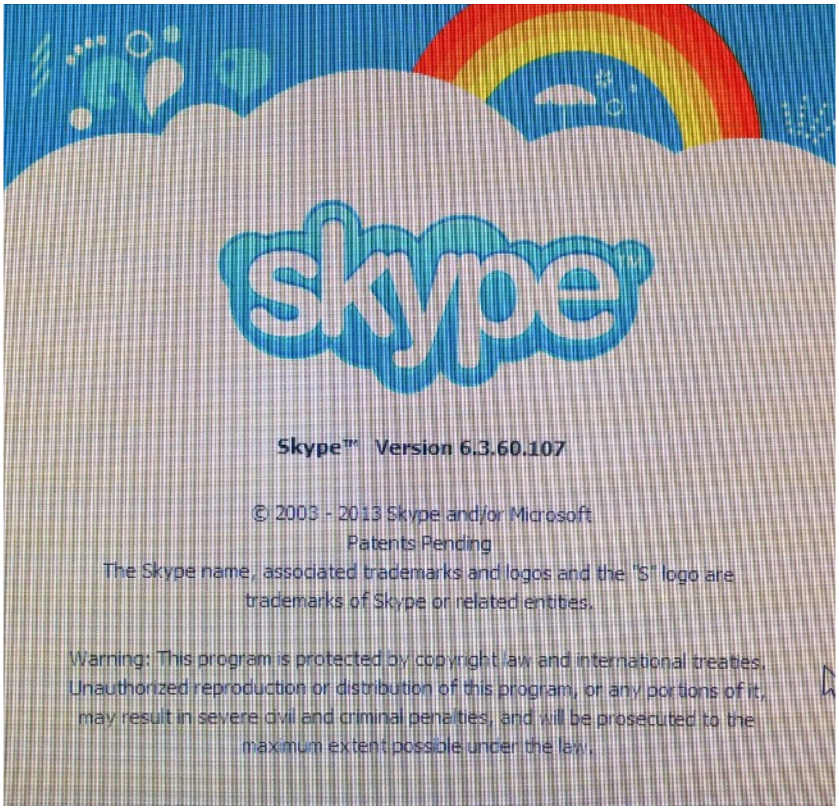


Figure A5

Logitech HD Webcam C270

Figure A6

Transcribed Version of Online Survey

See Chapter 2 for an overview about which sequences of questions were randomized

Background Information Survey

When you finish, please turn the survey over and remain seated until the experimenter enters to pick up the survey. Please take your time to answer each question carefully and honestly.

1) Are you a student at Cornell?

2) What is your netID?

3) Gender (circle one):

Female Male

4) Age:

5) Citizenship(s) (circle all that apply):

- United States

- Canada
- China
- South Korea
- India
- Singapore
- Taiwan
- Turkey
- Mexico
- Thailand
- Germany
- Other (please specify)

(If you are not a US citizen) How many total years have you lived in the US?

6) Native Language(s) (circle all that apply)
- English
- Mandarin
- Cantonese
- Taiwanese
- Korean
- Spanish
- French
- German
- Hindi
- Turkish
- Thai
- Other (please specify)

7) Country of Birth (circle all that apply)
- United States
- Canada
- China
- South Korea
- India
- Singapore
- Taiwan
- Turkey
- Mexico
- Thailand
- Germany
- Other (please specify)

8) Ethnicity (circle all that apply)

- American Indian/Alaskan Native
- Asian
- African-American
- Native Hawaiian or Other Pacific Islander
- White/Caucasian
- Hispanic/Latino

9) Occupation (circle)
- Undergraduate
- Graduate
- Other (please specify)

10) Are you wearing contacts?

a) Yes   b) no

11) What video chat service do you use the most often (circle your answer)?

a) Skype   b) Google Hangouts   c) Apple Facetime   d) Other: ______________________

12) How many years have you been using video chat?

13) On average, how many times do you use video chat per week?

14) On average, how many times do you use video chat per month?

15) Do you have a history of having back pain?

Video Chat Survey

I. Please answer each of the following questions about the video chat conversation that you just completed with your chat partner:

1) What percentage of the total video chat do you think you were gazing down at your self-display on the screen?

2) What percentage of the total video chat do you think that you were gazing at your partner on the screen?

3) What percentage of the total video chat do you think that you were gazing at your partner's room on the screen?

4) What percentage of the total video chat do you think you were gazing off screen?

5) What percentage of the total video chat do you think that your partner was gazing at herself?

6) What percentage of the total video chat do you think that your partner was gazing at you on their screen?

7) Recall the improvements to Cornell student life that you and your partner identified during the video chat. List as many of these improvements as you can remember below. Only list improvements that you discussed with your partner during the conversation. Please number each improvement:

8) Do you know your partner outside of this study?

II. Self-Display Description

1. Could you tell when your chat partner was looking at his/her self-video display?
   (yes, no)

2. How distracting did you find your own self-video display window (circle a number)?

Not Distracting	Slightly Distracting	Moderately Distracting	Very Distracting
1	2	3	4

3. How comforting did you find your own self-video display window (circle a number)?

Not Comforting	Slightly Comforting	Moderately Comforting	Very Comforting
1	2	3	4

4. If given the choice, would you rather
   1. have a self-video display window visible during video chats
   2. have no self-video display window visible during video chats

APPENDIX B: TABLES

Table B1
Data Acquisition and Track Analysis

Track Loss

Participant	Track Loss Ratio
1	NA
2	0.22
3	0.16
4	0.28
5	0.14
6	0.46
7	0.17
8	NA
9	0.24
10	0.32
11	0.31
12	0.31
13	0.19
14	0.24
15	0.19
16	0.46
17	0.41
18	0.17
19	0.19
20	NA
21	0.33
22	NA
23	0.36

Track Loss

Participant	Track Loss Ratio
24	0.25
25	0.34
26	NA
27	0.37
28	0.30
29	0.38
30	0.20
31	0.27
32	0.36
33	0.18
34	0.28
35	NA
36	0.27
37	0.49
38	0.46
39	0.16
40	0.44

Table B2

Demographic Data

Age of Participants

	N	Range	Minimum	Maximum	Mean		Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
age	40	7.00	18.00	25.00	20.4250	.27477	1.73778
Valid N (listwise)	40						

Average Video Chat Use Per Month

N	Valid	40
	Missing	0
Mean		3.8625
Median		1.5000
Std. Deviation		6.01599
Percentiles	5	.0500
	25	1.0000
	50	1.5000
	75	4.0000
	95	24.5500

Table B3

Normality Tests (Actual and Estimated Gaze)

Tests of Normality[c,d,e]

	SVDcondition	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
percent_gazeatpartner	"Large Self Video Display"	.202	18	.050	.926	18	.167
	"Small Self Video Display"	.122	16	.200*	.975	16	.906
est_percentgazeatpartner	"Large Self Video Display"	.132	22	.200*	.953	22	.358
	"Small Self Video Display"	.154	18	.200*	.931	18	.204
percent_gazeatselfvideo	"Large Self Video Display"	.278	18	.001	.597	18	.000
	"Small Self Video Display"	.304	16	.000	.570	16	.000
est_percentgazeatselfvideo	"Large Self Video Display"	.148	22	.200*	.877	22	.011
	"Small Self Video Display"	.223	18	.018	.876	18	.022
percent_gazeatpartneroom	"Large Self Video Display"	.183	18	.112	.922	18	.143
	"Small Self Video Display"	.224	16	.030	.825	16	.006
est_percentgazeatpartneroom	"Large Self Video Display"	.174	22	.081	.851	22	.003
	"Small Self Video Display"	.286	18	.000	.724	18	.000
percent_gazeoffscreen	"Large Self Video Display"	.181	18	.123	.910	18	.085
	"Small Self Video Display"	.196	16	.103	.943	16	.388
est_percentgazeoffscreen	"Large Self Video Display"	.396	22	.000	.361	22	.000
	"Small Self Video Display"	.408	18	.000	.456	18	.000

Normality Tests (Estimated Partner Gaze)

Tests of Normality

	SVDcondition	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
est_percentpartnergaze datyou	"Large Self Video Display"	.171	22	.093	.899	22	.029
	"Small Self Video Display"	.146	18	.200*	.894	18	.046
est_percentpartnergaze datselfvideo	"Large Self Video Display"	.263	22	.000	.852	22	.004
	"Small Self Video Display"	.233	18	.011	.889	18	.037

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Normality Tests (Collaborative Performance)

Tests of Normality

	SVDcondition	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
commongroundratio	"Large Self Video Display"	.128	11	.200*	.973	11	.915
	"Small Self Video Display"	.228	9	.197	.908	9	.300
num_totaluniqueteamim provements	"Large Self Video Display"	.190	11	.200*	.924	11	.355
	"Small Self Video Display"	.331	9	.005	.675	9	.001

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table B4
Heteroscedasticity Tests (Actual and Estimated Gaze)

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
percent_gazeatpartner	Based on Mean	2.114	1	32	.156
	Based on Median	2.305	1	32	.139
	Based on Median and with adjusted df	2.305	1	31.764	.139
	Based on trimmed mean	2.191	1	32	.149
est_percentgazeatpartner	Based on Mean	1.175	1	38	.285
	Based on Median	1.169	1	38	.287
	Based on Median and with adjusted df	1.169	1	36.798	.287
	Based on trimmed mean	1.168	1	38	.287
percent_gazeatselfvideo	Based on Mean	7.488	1	32	.010
	Based on Median	5.566	1	32	.025
	Based on Median and with adjusted df	5.566	1	17.077	.030
	Based on trimmed mean	5.893	1	32	.021
est_percentgazeatselfvideo	Based on Mean	9.752	1	38	.003
	Based on Median	7.477	1	38	.009
	Based on Median and with adjusted df	7.477	1	24.984	.011
	Based on trimmed mean	8.373	1	38	.006

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
percent_gazeatpartneroom	Based on Mean	11.636	1	32	.002
	Based on Median	3.999	1	32	.054
	Based on Median and with adjusted df	3.999	1	16.629	.062
	Based on trimmed mean	10.515	1	32	.003
est_percentgazeatpartneroom	Based on Mean	6.582	1	38	.014
	Based on Median	4.595	1	38	.039
	Based on Median and with adjusted df	4.595	1	20.034	.045
	Based on trimmed mean	5.172	1	38	.029
percent_gazeoffscreen	Based on Mean	3.426	1	32	.073
	Based on Median	2.860	1	32	.101
	Based on Median and with adjusted df	2.860	1	30.657	.101
	Based on trimmed mean	3.315	1	32	.078
est_percentgazeoffscreen	Based on Mean	.055	1	38	.815
	Based on Median	.209	1	38	.650
	Based on Median and with adjusted df	.209	1	37.970	.650
	Based on trimmed mean	.132	1	38	.719

Heteroscedasticity Tests (Estimated Partner Gaze)

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
est_percentpartnergazedatyou	Based on Mean	.353	1	38	.556
	Based on Median	.223	1	38	.639
	Based on Median and with adjusted df	.223	1	37.987	.639
	Based on trimmed mean	.399	1	38	.531
est_percentpartnergazedatselfvideo	Based on Mean	.714	1	38	.403
	Based on Median	.334	1	38	.567
	Based on Median and with adjusted df	.334	1	33.842	.567
	Based on trimmed mean	.721	1	38	.401

Heteroscedasticity Tests (Collaborative Performance)

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
num_totaluniqueteamimprovements	Based on Mean	.091	1	18	.766
	Based on Median	.082	1	18	.777
	Based on Median and with adjusted df	.082	1	13.858	.778
	Based on trimmed mean	.027	1	18	.871
commongroundratio	Based on Mean	.018	1	18	.895
	Based on Median	.017	1	18	.897
	Based on Median and with adjusted df	.017	1	15.838	.898
	Based on trimmed mean	.028	1	18	.869

Table B5
Self Directed Gaze Tests

SVDcondition = "Small Self Video Display"

Statistics[a]

percent_gazeatselfvideo

N	Valid	16
	Missing	2
Median		.0320
Skewness		2.994
Std. Error of Skewness		.564
Kurtosis		9.724
Std. Error of Kurtosis		1.091
Percentiles	25	.0015
	50	.0320
	75	.1882

a. SVDcondition = "Small Self Video Display"

SVDcondition = "Large Self Video Display"

Statistics[a]

percent_gazeatselfvideo

N	Valid	18
	Missing	4
Median		2.8888
Skewness		3.269
Std. Error of Skewness		.536
Kurtosis		12.227
Std. Error of Kurtosis		1.038
Percentiles	25	.9164
	50	2.8888
	75	6.3307

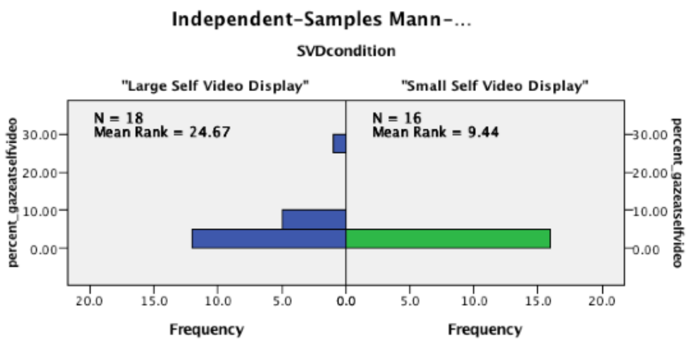
a. SVDcondition = "Large Self Video Display"

Table B6
Partner Directed Gaze Tests

Descriptive Statistics

Dependent Variable: percent_gazeatpartner

SVDcondition	Mean	Std. Deviation	N
"Large Self Video Display"	72.4944	12.47498	18
"Small Self Video Display"	65.4762	17.49055	16
Total	69.1917	15.22720	34



Total N	34
Mann-Whitney U	15.000
Wilcoxon W	151.000
Test Statistic	15.000
Standard Error	28.961
Standardized Test Statistic	-4.454
Asymptotic Sig. (2-sided test)	.000
Exact Sig. (2-sided test)	.000

Estimated Marginal Means

SVDcondition

Estimates

Dependent Variable: percent_gazeatpartner

SVDcondition	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
"Large Self Video Display"	72.678[a]	3.547	65.444	79.911
"Small Self Video Display"	65.270[a]	3.762	57.597	72.943

a. Covariates appearing in the model are evaluated at the following values: partnerAOIsize = 520604.53.

Tests of Between-Subjects Effects

Dependent Variable: percent_gazeatpartner

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	651.073[a]	2	325.536	1.442	.252	.085
Intercept	7148.936	1	7148.936	31.657	.000	.505
partnerAOIsize	233.857	1	233.857	1.036	.317	.032
SVDcondition	462.251	1	462.251	2.047	.163	.062
Error	7000.561	31	225.825			
Total	170426.439	34				
Corrected Total	7651.634	33				

a. R Squared = .085 (Adjusted R Squared = .026)

Parameter Estimates

Dependent Variable: percent_gazeatpartner

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared
					Lower Bound	Upper Bound	
Intercept	54.995	10.963	5.016	.000	32.635	77.355	.448
partnerAOIsize	1.974E-5	1.939E-5	1.018	.317	-1.982E-5	5.929E-5	.032
[SVDcondition=1]	7.407	5.177	1.431	.163	-3.152	17.967	.062
[SVDcondition=2]	0[a]	.	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Table B7
Collaborative Performance Tests

SVDcondition = "Large Self Video Display"

Statistics[a]

num_totaluniqueteamimprovements

N	Valid	11
	Missing	11
Median		15.000
Percentiles	25	7.000
	50	15.000
	75	17.000

a. SVDcondition = "Large Self Video Display"

SVDcondition = "Small Self Video Display"

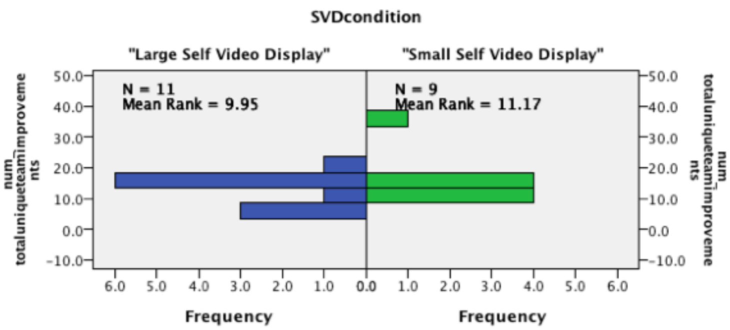
Statistics[a]

num_totaluniqueteamimprovements

N	Valid	9
	Missing	9
Median		15.000
Percentiles	25	10.500
	50	15.000
	75	16.500

a. SVDcondition = "Small Self Video Display"

Independent-Samples Mann-Whitney U Test



Total N	20
Mann-Whitney U	55.500
Wilcoxon W	100.500
Test Statistic	55.500
Standard Error	13.028
Standardized Test Statistic	.461
Asymptotic Sig. (2-sided test)	.645
Exact Sig. (2-sided test)	.656

Table B8
Common Ground Ratio Tests

T-Test

Group Statistics

	SVDcondition	N	Mean	Std. Deviation	Std. Error Mean
commongroundratio	"Large Self Video Display"	11	41.6528	11.55686	3.48452
	"Small Self Video Display"	9	44.1763	14.01153	4.67051

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
commongroundratio	Equal variances assumed	.018	.895	-.442	18	.664	-2.52345	5.71115	-14.52213	9.47523
	Equal variances not assumed			-.433	15.534	.671	-2.52345	5.82714	-14.90661	9.85972

Table B9
Environmental Gaze Tests

SVDcondition = "Large Self Video Display"

Statistics[a]

percent_gazeatpartneroom

N	Valid	18
	Missing	4
Median		4.7757
Percentiles	25	2.0360
	50	4.7757
	75	6.6668

a. SVDcondition = "Large Self Video Display"

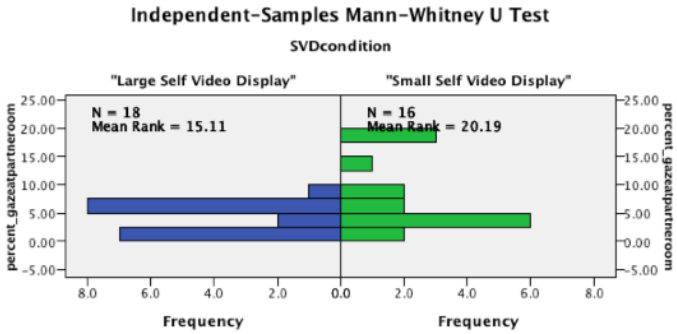
SVDcondition = "Small Self Video Display"

Statistics[a]

percent_gazeatpartneroom

N	Valid	16
	Missing	2
Median		5.0680
Percentiles	25	3.6412
	50	5.0680
	75	12.7101

a. SVDcondition = "Small Self Video Display"



Total N	34
Mann-Whitney U	187.000
Wilcoxon W	323.000
Test Statistic	187.000
Standard Error	28.983
Standardized Test Statistic	1.484
Asymptotic Sig. (2-sided test)	.138
Exact Sig. (2-sided test)	.144

Table B10
Self Gaze Awareness Tests

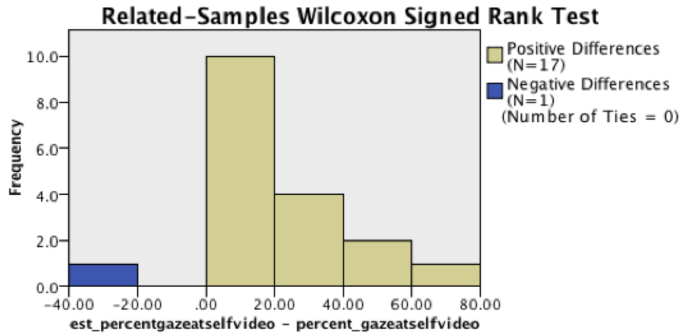
Participant's gaze awareness of looking at their own SVD

SVDcondition = "Large Self Video Display"

Statistics[a]

		percent_gazeatselfvideo	est_percentgazeatselfvideo
N	Valid	18	22
	Missing	4	0
Median		2.8888	20.0000
Percentiles	25	.9164	5.0000
	50	2.8888	20.0000
	75	6.3307	31.2500

a. SVDcondition = "Large Self Video Display"



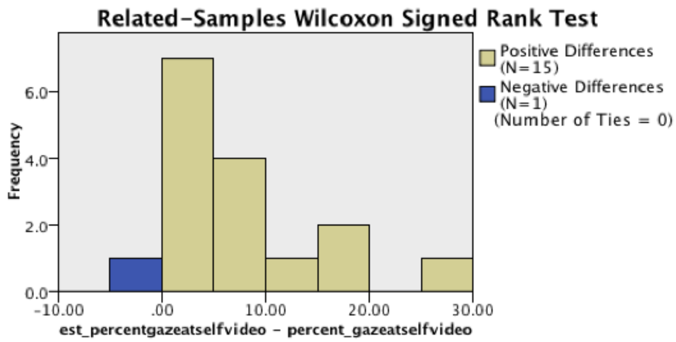
Total N	18
Test Statistic	160.000
Standard Error	22.962
Standardized Test Statistic	3.245
Asymptotic Sig. (2-sided test)	.001

SVDcondition = "Small Self Video Display"

Statistics[a]

		percent_gazeatselfvideo	est_percentgazeatselfvideo
N	Valid	16	18
	Missing	2	0
Median		.0320	7.5000
Percentiles	25	.0015	2.5000
	50	.0320	7.5000
	75	.1882	12.5000

a. SVDcondition = "Small Self Video Display"



Total N	16
Test Statistic	135.000
Standard Error	19.339
Standardized Test Statistic	3.464
Asymptotic Sig. (2-sided test)	.001

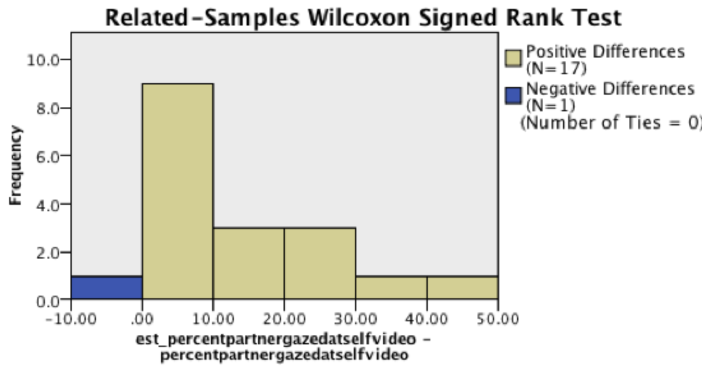
Table B11
Participant's Gaze Awareness of Partner's Self-Gaze Tests

SVDcondition = "Large Self Video Display"

Statistics[a]

		percentpartnergazedatselfvideo	est_percentpartnergazedatselfvideo
N	Valid	18	22
	Missing	4	0
Median		2.8900	10.0000
Percentiles	25	.9150	5.0000
	50	2.8900	10.0000
	75	6.3275	30.0000

a. SVDcondition = "Large Self Video Display"



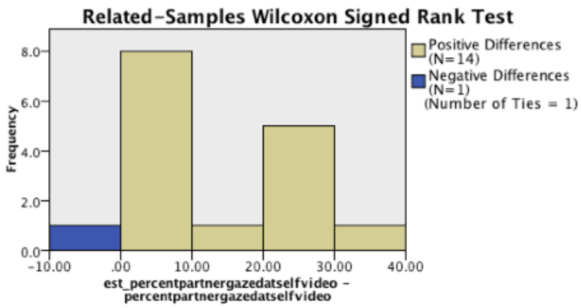
Total N	18
Test Statistic	168.000
Standard Error	22.962
Standardized Test Statistic	3.593
Asymptotic Sig. (2-sided test)	.000

SVDcondition = "Small Self Video Display"

Statistics[a]

		percentpartnergazedatselfvideo	est_percentpartnergazedatselfvideo
N	Valid	16	18
	Missing	2	0
Median		.0300	10.0000
Percentiles	25	.0025	4.5000
	50	.0300	10.0000
	75	.1900	26.2500

a. SVDcondition = "Small Self Video Display"



Total N	16
Test Statistic	119.000
Standard Error	17.607
Standardized Test Statistic	3.351
Asymptotic Sig. (2-sided test)	.001

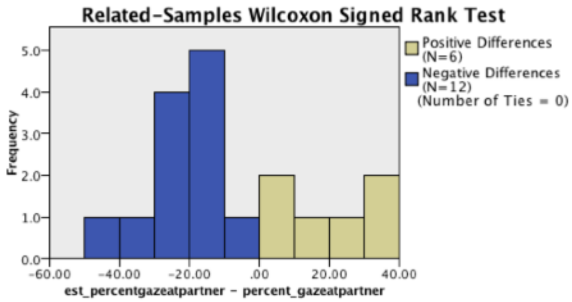
Table B12
Participant's Gaze Awareness of Looking at their Partner Tests

SVDcondition = "Large Self Video Display"

Statistics[a]

		percent_gazeatpartner	est_percentgazeatpartner
N	Valid	18	22
	Missing	4	0
Median		76.1358	65.0000
Percentiles	25	63.8877	48.7500
	50	76.1358	65.0000
	75	78.7450	90.0000

a. SVDcondition = "Large Self Video Display"



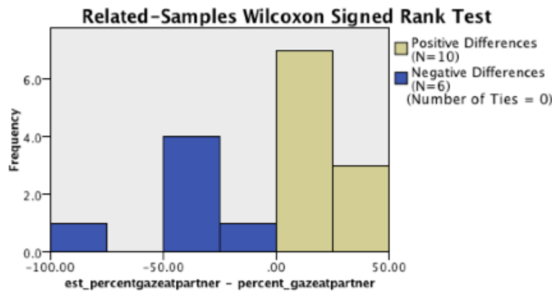
Total N	18
Test Statistic	55.000
Standard Error	22.962
Standardized Test Statistic	-1.328
Asymptotic Sig. (2-sided test)	.184

SVDcondition = "Small Self Video Display"

Statistics[a]

		percent_gazeatpartner	est_percentgazeatpartner
N	Valid	16	18
	Missing	2	0
Median		67.1247	62.5000
Percentiles	25	50.4266	43.7500
	50	67.1247	62.5000
	75	77.4133	95.0000

a. SVDcondition = "Small Self Video Display"



Total N	16
Test Statistic	71.000
Standard Error	19.339
Standardized Test Statistic	.155
Asymptotic Sig. (2-sided test)	.877

Table B13

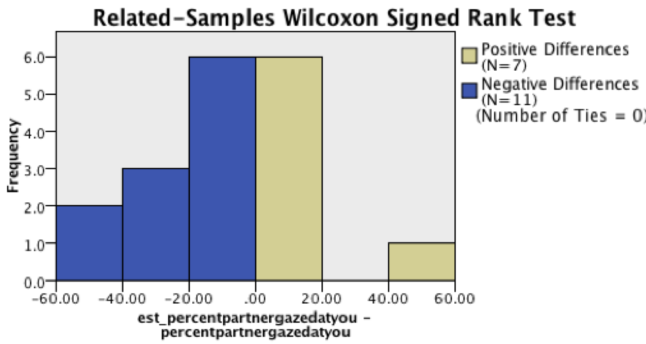
Participant's Gaze Awareness of Partner Looking at the Participant Tests

SVDcondition = "Large Self Video Display"

Statistics[a]

		percentpartnergazedatyou	est_percentpartnergazedatyou
N	Valid	18	22
	Missing	4	0
Median		76.1369	75.0000
Percentiles	25	63.8850	50.0000
	50	76.1369	75.0000
	75	78.7475	90.0000

a. SVDcondition = "Large Self Video Display"



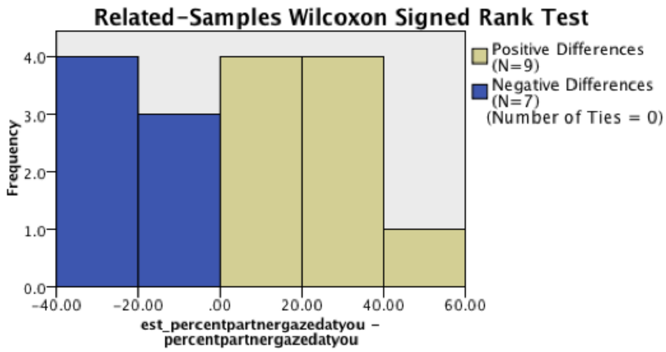
Total N	18
Test Statistic	67.000
Standard Error	22.962
Standardized Test Statistic	-.806
Asymptotic Sig. (2-sided test)	.420

SVDcondition = "Small Self Video Display"

Statistics[a]

		percentpartnergazedatyou	est_percentpartnergazedatyou
N	Valid	16	18
	Missing	2	0
Median		67.1250	70.0000
Percentiles	25	50.4300	37.5000
	50	67.1250	70.0000
	75	77.4100	95.0000

a. SVDcondition = "Small Self Video Display"



Total N	16
Test Statistic	68.000
Standard Error	19.339
Standardized Test Statistic	.000
Asymptotic Sig. (2-sided test)	1.000

Table B14

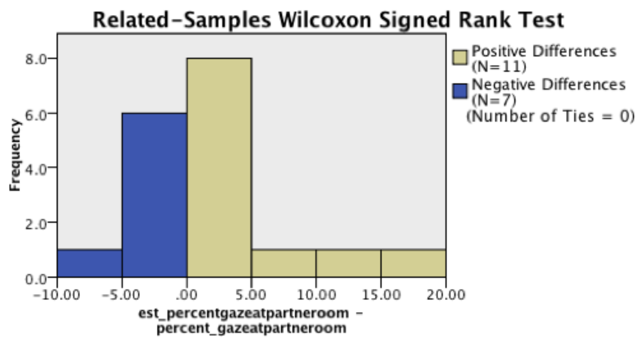
Participant's Gaze Awareness of Looking at their Partner's Room

SVDcondition = "Large Self Video Display"

Statistics[a]

		percent_gazeatpartneroom	est_percentgazeatpartneroom
N	Valid	18	22
	Missing	4	0
Median		4.7757	5.0000
Percentiles	25	2.0360	.0000
	50	4.7757	5.0000
	75	6.6668	10.0000

a. SVDcondition = "Large Self Video Display"



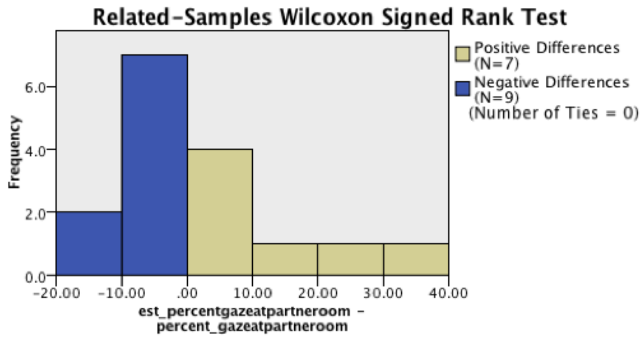
Total N	18
Test Statistic	127.000
Standard Error	22.962
Standardized Test Statistic	1.807
Asymptotic Sig. (2-sided test)	.071

SVDcondition = "Small Self Video Display"

Statistics[a]

		percent_gazeatpartneroom	est_percentgazeatpartneroom
N	Valid	16	18
	Missing	2	0
Median		5.0680	7.5000
Percentiles	25	3.6412	.0000
	50	5.0680	7.5000
	75	12.7101	15.0000

a. SVDcondition = "Small Self Video Display"



Total N	16
Test Statistic	67.000
Standard Error	19.339
Standardized Test Statistic	-.052
Asymptotic Sig. (2-sided test)	.959

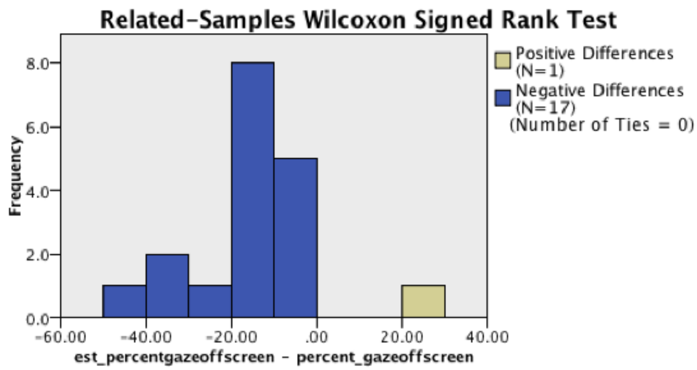
Table B15
Participant's Gaze Awareness of Looking Offscreen

SVDcondition = "Large Self Video Display"

Statistics[a]

		percent_gazeoffscreen	est_percentgazeoffscreen
N	Valid	18	22
	Missing	4	0
Median		14.8803	.0000
Percentiles	25	12.0157	.0000
	50	14.8803	.0000
	75	22.5094	5.0000

a. SVDcondition = "Large Self Video Display"



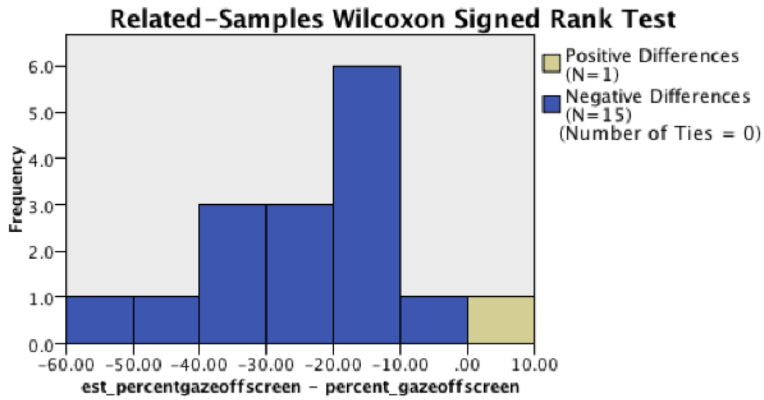
Total N	18
Test Statistic	15.000
Standard Error	22.962
Standardized Test Statistic	-3.070
Asymptotic Sig. (2-sided test)	.002

SVDcondition = "Small Self Video Display"

Statistics[a]

		percent_gazeoffscreen	est_percentgazeoffscreen
N	Valid	16	18
	Missing	2	0
Median		21.6175	3.5000
Percentiles	25	15.7421	.0000
	50	21.6175	3.5000
	75	40.6756	5.0000

a. SVDcondition = "Small Self Video Display"



Total N	16
Test Statistic	2.000
Standard Error	19.339
Standardized Test Statistic	-3.413
Asymptotic Sig. (2-sided test)	.001

Table B16

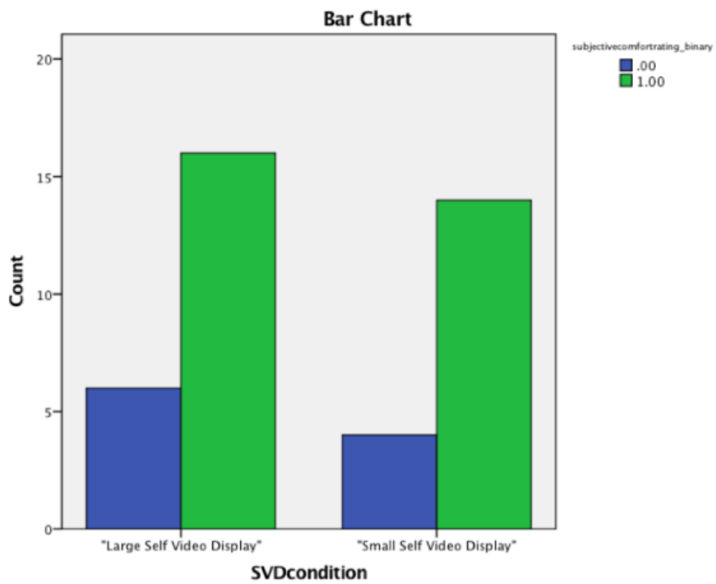
Self-video Display Evaluation Tests

SVD Perceived Comfort

SVDcondition * subjectivecomfortrating_binary Crosstabulation

			subjectivecomfortrating_binary		
			.00	1.00	Total
SVDcondition	"Large Self Video Display"	Count	6a	16a	22
		Expected Count	5.5	16.5	22.0
		% within SVDcondition	27.3%	72.7%	100.0%
		% within subjectivecomfortrating_binary	60.0%	53.3%	55.0%
		% of Total	15.0%	40.0%	55.0%
		Standardized Residual	.2	-.1	
	"Small Self Video Display"	Count	4a	14a	18
		Expected Count	4.5	13.5	18.0
		% within SVDcondition	22.2%	77.8%	100.0%
		% within subjectivecomfortrating_binary	40.0%	46.7%	45.0%
		% of Total	10.0%	35.0%	45.0%
		Standardized Residual	-.2	.1	
Total		Count	10	30	40
		Expected Count	10.0	30.0	40.0
		% within SVDcondition	25.0%	75.0%	100.0%
		% within subjectivecomfortrating_binary	100.0%	100.0%	100.0%
		% of Total	25.0%	75.0%	100.0%

Each subscript letter denotes a subset of subjectivecomfortrating_binary categories whose column proportions do not differ significantly from each other at the .05 level.



Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.135[a]	1	.714	1.000	.503	
Continuity Correction[b]	.000	1	1.000			
Likelihood Ratio	.135	1	.713	.734	.503	
Fisher's Exact Test				1.000	.503	
Linear-by-Linear Association	.131[c]	1	.717	1.000	.503	.269
N of Valid Cases	40					

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 4.50.

b. Computed only for a 2x2 table

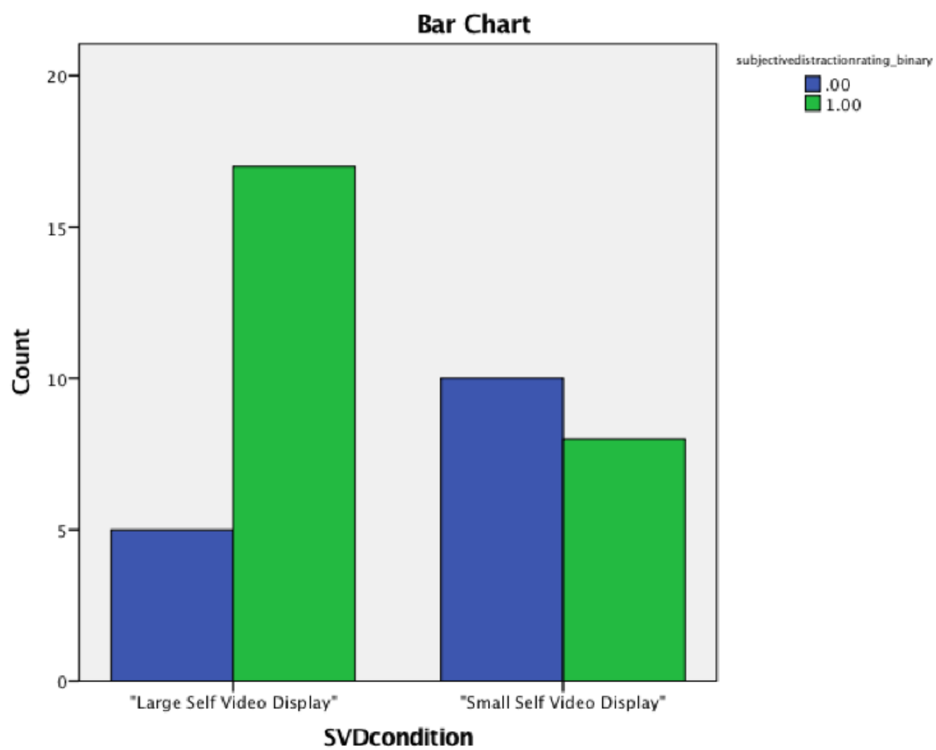
c. The standardized statistic is .362.

SVD Perceived Distraction

SVDcondition * subjectivedistractionrating_binary Crosstabulation

			subjectivedistractionrating_binary		
			.00	1.00	Total
SVDcondition	"Large Self Video Display"	Count	5a	17b	22
		Expected Count	8.3	13.8	22.0
		% within SVDcondition	22.7%	77.3%	100.0%
		% within subjectivedistractionrating_binary	33.3%	68.0%	55.0%
		% of Total	12.5%	42.5%	55.0%
		Standardized Residual	-1.1	.9	
	"Small Self Video Display"	Count	10a	8b	18
		Expected Count	6.8	11.3	18.0
		% within SVDcondition	55.6%	44.4%	100.0%
		% within subjectivedistractionrating_binary	66.7%	32.0%	45.0%
		% of Total	25.0%	20.0%	45.0%
		Standardized Residual	1.3	-1.0	
Total		Count	15	25	40
		Expected Count	15.0	25.0	40.0
		% within SVDcondition	37.5%	62.5%	100.0%
		% within subjectivedistractionrating_binary	100.0%	100.0%	100.0%
		% of Total	37.5%	62.5%	100.0%

Each subscript letter denotes a subset of subjectivedistractionrating_binary categories whose column proportions do not differ significantly from each other at the .05 level.



Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	4.552[a]	1	.033	.050	.035	
Continuity Correction[b]	3.259	1	.071			
Likelihood Ratio	4.612	1	.032	.050	.035	
Fisher's Exact Test				.050	.035	
Linear-by-Linear Association	4.438[c]	1	.035	.050	.035	.029
N of Valid Cases	40					

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.75.

b. Computed only for a 2x2 table

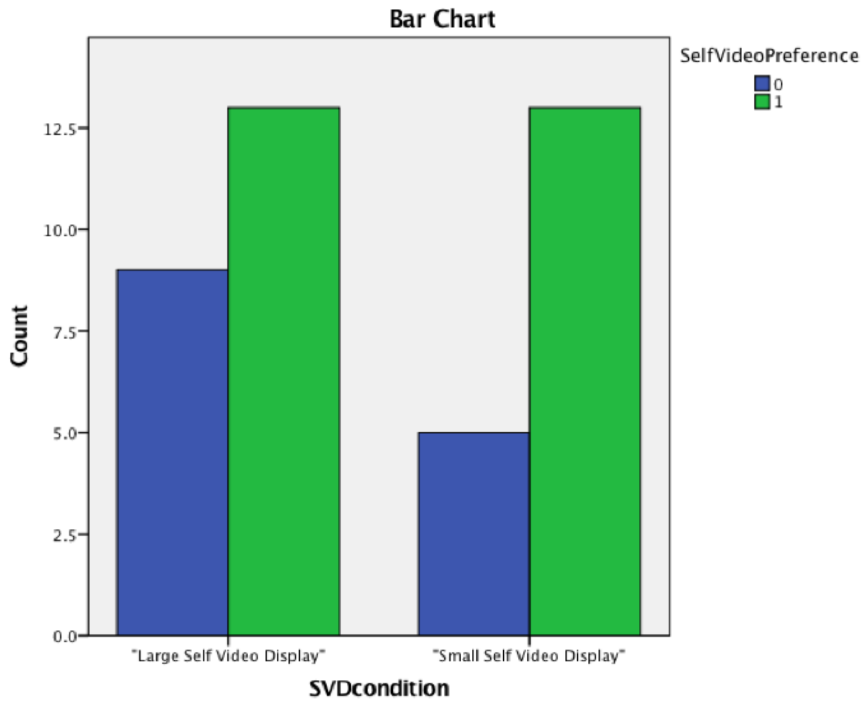
c. The standardized statistic is -2.107.

SVD Preference

SVDcondition * SelfVideoPreference Crosstabulation

			SelfVideoPreference		
			0	1	Total
SVDcondition	"Large Self Video Display"	Count	9a	13a	22
		Expected Count	7.7	14.3	22.0
		% within SVDcondition	40.9%	59.1%	100.0%
		% within SelfVideoPreference	64.3%	50.0%	55.0%
		% of Total	22.5%	32.5%	55.0%
		Standardized Residual	.5	-.3	
	"Small Self Video Display"	Count	5a	13a	18
		Expected Count	6.3	11.7	18.0
		% within SVDcondition	27.8%	72.2%	100.0%
		% within SelfVideoPreference	35.7%	50.0%	45.0%
		% of Total	12.5%	32.5%	45.0%
		Standardized Residual	-.5	.4	
Total		Count	14	26	40
		Expected Count	14.0	26.0	40.0
		% within SVDcondition	35.0%	65.0%	100.0%
		% within SelfVideoPreference	100.0%	100.0%	100.0%
		% of Total	35.0%	65.0%	100.0%

Each subscript letter denotes a subset of SelfVideoPreference categories whose column proportions do not differ significantly from each other at the .05 level.



Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.750[a]	1	.386	.510	.298	
Continuity Correction[b]	.284	1	.594			
Likelihood Ratio	.758	1	.384	.510	.298	
Fisher's Exact Test				.510	.298	
Linear-by-Linear Association	.732[c]	1	.392	.510	.298	.184
N of Valid Cases	40					

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.30.

b. Computed only for a 2x2 table

c. The standardized statistic is .855.