Summer 6-29-2022

USING CONTINGENT PRAISE TO INCREASE VISUAL ENGAGEMENT IN AN ASYNCHRONOUS ONLINE LEARNING ENVIRONMENT: AN EYE TRACKING STUDY

Andrew Rozsa

USING CONTINGENT PRAISE TO INCREASE VISUAL ENGAGEMENT IN AN ASYNCHRONOUS ONLINE LEARNING ENVIRONMENT: AN EYE TRACKING STUDY

by

Andrew J. Rozsa III

A Dissertation
Submitted to the Graduate School,
the College of Education and Human Sciences
and the School of Psychology
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved by:

D. Joe Olmi, Ph.D., Committee Chair
Hans Stadthagen, Ph.D., Committee Member
Crystal Taylor, Ph.D., Committee Member
Michael Mong, Ph.D., Committee Member

August 2022

Published by the Graduate School

THE UNIVERSITY OF SOUTHERN MISSISSIPPI

ABSTRACT

As more students move to online learning, this results in not only new barriers but new opportunities in academia. The purpose of this study was to examine how behavior-contingent praise affects visual engagement with an online video lecture when using *WebGazer*, a publicly available eye tracking software, with a user's integrated webcam. A second aim of this study was to examine if using *WebGazer* with an integrated webcam was a valid alternative to hand scoring when collecting visual engagement data. Results of *WebGazer* measurement indicated a moderate effect size for three participants in the presence of contingent praise, and a large effect size was observed for one participant when provided contingent praise. Based on visual analysis and simple linear regression, level, shape of data paths, trend, and overall range of data were similar for three participants. One participant's *WebGazer* and hand scoring data demonstrated a notable discrepancy in range, level, and shape for the Demand and Praise conditions. These results indicate that contingent praise may result in an increase in visual engagement in online learning environments and that using *WebGazer* and an integrated webcam may be a valid tool for measuring visual engagement in online learning environments. Discrepancies in *WebGazer* and hand scoring data are discussed.

ACKNOWLEDGMENTS

I would like to thank Drs. Joe Olmi, Hans Stadthagen, Crystal Taylor, and Michael Mong for their assistance in this project. I especially want to thank Dr. Olmi for his unwavering support, guidance, and the amount of time and effort he put into providing me with the individual attention required to bring this project to fruition. Dr. Stadthagen's technical expertise, exacting outlook, and knowledge of current advancements in the field of eye tracking were invaluable to this project. I want to thank Dr. Taylor for her keen eye for detail and feedback. I also wish to thank Dr. Mong for his fresh perspective and insightful questions regarding the study's design. I am also grateful to Parker Lundy and Jordan Etherington for their assistance in data collection. I would also like to thank Parker Lundy and Meleah Ackley for their emotional support these last five years, and throughout this project. Lastly, I want to thank my parents who have been my biggest supporters my entire life. I have undoubtedly inherited their love for teaching and helping others, and they have been instrumental in continuing to grow that love in me.

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGMENTS iii

LIST OF TABLES vii

CHAPTER I - INTRODUCTION 1

Evolution of Online Learning 1

Distance Learning 2

Student Engagement in the Classroom 3

Praise in the Classroom 4

Engagement in the Online Learning Environment 4

Eye Tracking to Measure Attention 5

What Works Clearinghouse 8

Summary 9

Purpose 10

CHAPTER II - METHODS 11

Participants and Setting 11

Materials 13

Computer Hardware 13

Video Stimuli 13

WebGazer 13

Current WebGazer Application 15

Pavlovia.. 16

Zoom 17

Data Sheet 17

Procedures... 18

Baseline.. 18

Demand Condition 19

Design, Data Analysis, and Dependent Variable.. 19

Visual Engagement 21

Procedural Integrity and Interobserver Agreement.. 21

CHAPTER III – RESULTS 24

WebGazer Visual Engagement 24

Hand Scored Visual Engagement 28

WebGazer versus Hand Scored visual engagement... 32

CHAPTER IV – DISCUSSION 36

Limitations 39

Future Directions 44

Conclusion 46

APPENDIX A – IRB Approval Letter 48

APPENDIX B– Recruitment Flyer... 49

APPENDIX C – Pre-Study Questionnaire 50

APPENDIX D – Consent Form 51

APPENDIX E – Video Links 53

APPENDIX F – Data Sheet 54

APPENDIX G – Baseline Procedural integrity 55

APPENDIX H – Demand Procedural integrity 56

APPENDIX I – Praise Procedural integrity 57

REFERENCES 58

LIST OF TABLES

Table 1 Participant Demographics 12

LIST OF ILLUSTRATIONS

Figure 1. Generations of Distance Education 2

Figure 2. Example of researcher's screen with Square A and Square B showing............. 17

Figure 3. Percent visually engaged as measured by WebGazer (WG) application 25

Figure 4. Percent visual engagement as measured by hand scoring (HS) 29

Figure 5. Comparison between hand scored (HS) and WebGazer (WG) Praise data....... 33

Figure 6. Comparison between hand scored (HS) and WebGazer (WG) Demand data... 34

CHAPTER I - INTRODUCTION

Over the last three decades, the number of students enrolled in postsecondary education online courses has been growing at a phenomenal rate (Hu, Arnesen, Barbour, & Leary, 2019; Schwirzke, Vashaw, & Watson, 2018). Since 2012, the percentage of students taking at least one online class has increased by no less than .7% and as much as 1.8% per year for the period 2012-2019. By 2019, 37% of undergraduate students were taking at least one online course (U.S. Department of Education, 2021). With the movement toward online learning, education is presented with an unprecedented opportunity to assess effectiveness more efficiently and student behavior with less effort and more precision. When the public was introduced to the coronavirus disease (COVID-19) in 2019, the education system faced many struggles to ensure appropriate translation of instruction to an online environment. By Fall 2020, 74.7% of students were taking at least one online course, and 44.7% were taking only online courses. It was unknown whether many of the classroom-based strategies that had been employed for decades would be effective in this new remote environment. As the world continues to move toward online learning, it is imperative that researchers explore novel strategies to assess effectiveness. For over a century, students have sought alternative means of accessing education outside the traditional classroom setting. COVID-19 has fast-tracked the necessity to address whether this delivery format is meeting the needs of today's learners.

Evolution of Online Learning

While most publications use the terms *online learning* and *distance learning* interchangeably, there is a distinct difference between the two. A brief review of their history will serve to guide the reader to a better understanding (Singh & Thurman, 2019).

Distance Learning

According to Bates (2004), distance learning is defined as a method of education in which students have full control over when and where they study, typically without direct contact with the teacher. The defining feature of distance learning is the presence of a technology enabling the learning process. Distance learning has been evolving for decades, as can be seen in Figure 1. The first generation being characterized by using a single technology, with no direct interaction between the student and institution.

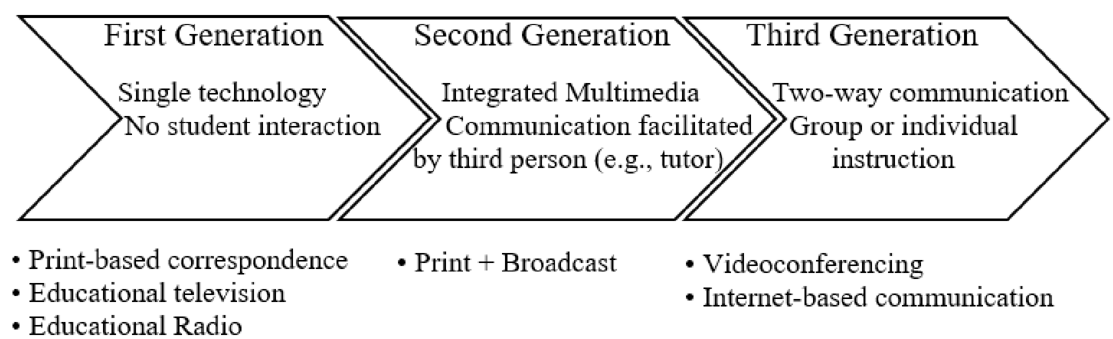

Figure 1. Generations of Distance Education

The second generation of distance learning moved from a single technology to multiple media. Lastly, the third generated incorporated the ability to teach multiple students at once and engage in live discussion. This was fueled largely by the introduction of the World Wide Web in 1989 and the subsequent introduction of the Internet for to the public in 1993 (Couldry, 2012). The third generation of distance learning has allowed easier access for otherwise isolated learners to higher education and more cost-effective means of providing education for different institutions. When both instructors and students are present during remote instruction, this is defined as synchronous learning. When instruction materials are recorded and accessed at different times, this is asynchronous learning. Many institutions have adopted more asynchronous teaching models, placing greater value on the flexibility it allows instructors to record and even re-use previous recordings (Bos, Groeneveld, van Bruggen, & Brand-Gruwel, 2015; Evans, 2008; Morris, Swinnerton, & Coop, 2019). For a more extensive review and comparison of the effectiveness of types of online learning, refer to the meta-analysis by Means and colleagues (2010), as this is outside the scope of this study. As the number of students enrolled in some form of online learning continues to increase, more research must be conducted to identify how best to increase not only effective instruction delivery, but student outcomes in these settings.

Student Engagement in the Classroom

When evaluating the effectiveness of instruction methodologies, one of the most examined variables is student engagement (Driscoll et al., 2012; Francescucci & Rohani, 2018; Watts, 2016). Although a multidimensional concept, for purposes of this study, it is defined as student involvement in academic learning, including behaviors such as attending (making eye contact) without engaging in disruptive behavior (Sinatra, Heddy, & Lombardi, 2015), asking questions in class, or engaging in class discussion (Birch & Ladd, 1997; Finn et al., 1995; Heddy, Sinatra, Seli, & Mukhopadhyay, 2014; Skinner & Belmont, 1993). This form of engagement, sometimes referred to as on-task behavior or academically engaged behavior, has decades of evidence supporting its positive correlation with academic achievement (Cobb, 1972; Greenwood, Terry, & Walker, 1994; Hecht, 1978; Lahaderne, 1968) more than any other observable behavior across all school subjects. While much of the current literature has measured behavioral engagement as a primary dependent variable to determine effectiveness of interventions, few have examined it in online environments. As recent research has indicated, measuring behavioral engagement in online environments is feasible (e.g., Rozsa, 2021).

Praise in the Classroom

Suffice it to say that this literature review is not intended to provide the reader with an extensive review of praise. For that information, the reader is encouraged to examine the review article by Floress and colleagues (2017). Briefly stated, providing praise for engagement has consistently been found to be a reliable means of increasing student engagement. Despite an extensive history investigating praise in classrooms, the overwhelming majority of studies have been with elementary school classrooms, with only a handful of studies with high school-aged students (e.g., Blaze et al., 2014; Duchaine, Jolivette, & Fredrick, 2011; Taber, 2014), and even fewer with undergraduate students (e.g., Lessard, Grossman, & Syme, 2015).

Engagement in the Online Learning Environment

As individuals continue to move to online learning environments, the increased reliance on computers will continue to increase as well. However, computer and computer-related technology is becoming cheaper and more widely available every year. Recent research has investigated the extent to which student engagement differs between synchronous online learning environments in comparison with asynchronous online learning environments (e.g., Giesbers et al., 2013), but these studies typically measure engagement in different manners. such as number and type of messages sent by students, as measuring engagement in asynchronous online learning environments may prove more problematic. Identifying a reliable method of measuring engagement that is usable across different online learning environments would allow for more accurate comparisons of interventions targeting engagement in online learning environments.

At the time, several new technologies have also become available to the public, including eye tracking. Eye tracking is a technology that measures an individual's eye gaze toward a specific stimulus, typically on a computer screen. Such technological advancements also allow for more continuous, automated observation and recording of behavior that was not previously feasible due to the need for human observers which can introduce error (Charlesworth & Spiker, 1975). By contrast, computers are bound by algorithms which dictate consistent data collection methodologies, thereby alleviating observation error.

Eye Tracking to Measure Attention

One of the most important components in assessing the effectiveness of teaching strategies is being able to measure the target behaviors, such as academic engagement or student performance (Driscoll et al., 2012). Measuring behavior in classrooms has traditionally relied upon either the observer being present in the same room as the participant or taping the individual for later coding behavior. The behaviors of interest are typically related to those expected in the classroom such as taking notes, looking toward the instructor during lectures, and staying in the assigned area. How these behaviors may appear in the classroom may be similar in an online learning environment, such as looking at the instructor or taking notes, but other barriers may be present such as other programs being open or being unable to identify if the student is taking notes or doing something else off-screen. In asynchronous classes, lectures and tasks may be completed online at a different time from when they were first recorded or uploaded, similar to what was seen in the second generation of distance learning. As more people gain access to faster internet connections and faster computers become more readily available, live lessons are more feasible for classes than ever.

Eye gaze and attention have been repeatedly shown to share a close relationship (e.g., Klein, 1980; Rafal, Calabresi, Brennan, & Sciolto, 1989; Rayner, 1998; Rayner, 2009), while research into eye movements and their relationship with academic behaviors goes as far back as 1879 (Huey, 1908). Until the advent of current technologies, most eye tracking methodologies required expensive, research-grade hardware. As a result, few groups outside research labs and universities had access to examining eye tracking applications that could be used to investigate more basic, yet important, research questions.

While some studies have successfully incorporated unique eye tracking media, such as using eye tracking glasses in face-to-face classrooms (Rosengrant et al., 2011), this type of technology is often very expensive or impractical. Most of the more recent eye tracking studies have been implemented in combination with computer screens. Some studies have examined the effect of on-screen elements such as the lecturer's video representation (e.g., Wang & Antonenko, 2017; Wang, Pi, & Hu, 2018) or monetary imbursement effects on performance and attending (Anderson, Laurent, & Yantis, 2011; Bucker & Theeuwes, 2014; Chelazzi et al., 2014; Engelmann & Pessoa, 2007; Failing & Theeuwes, 2014; Shomstein & Johnson, 2013). However, limited research has investigated the effect of praise on engagement in online learning settings.

In a recent study, Rozsa (2021) examined the effect of noncontingent praise on undergraduates' visual engagement, as measured by eye tracking software, and more formally defined as the participant looking toward the video during the lecture. Prior to this study, no other research had been conducted using live praise. All four participants experienced the baseline condition, during which no verbal interaction occurred between the researcher and participants. Each participant then experienced either the praise or neutral verbalizations condition, with half experiencing praise first, and the other half experiencing neutral verbalizations first followed by the second condition. During the neutral verbalizations condition, the researcher delivered a neutral verbalization every two mins. Neutral verbalizations were defined as verbal statements provided by the researcher that were related to video content, such as “This video is about a war,” but neither indicated approval or disapproval. During the praise condition, the researcher provided a verbal statement that indicated approval, such as “I love how you are watching the screen,” every two mins.

The researcher found that the presence of praise resulted in an overall increase in visual engagement for three out of four participants. The researcher also found that the presence of neutral verbalizations increased visual engagement for two out of four participants, though not to the same extent as praise. Because the experimental design was a counterbalanced ABC design, it was difficult to compare the neutral verbalizations condition data to baseline data for all participants. Furthermore, due to potential ceiling effects present during baseline, potential changes in visual engagement may have been masked, and the treatment may have been unnecessary for the chosen participants. The use of neutral verbalizations was also questionable, as these types of statements are unlikely to be made during a lecture. Lastly, praise was noncontingent, which, unless there are individual data to indicate otherwise, is less preferred to contingent praise

(Cooper, Heron, & Heward, 2019). As a result, more precise contingencies for praise, a more accurate experimental design, better recruitment selection, and a comparison condition that better represents what may be heard in online learning environments may provide better insight into efficacious teaching strategies for increasing visual engagement.

What Works Clearinghouse

To improve the dissemination of products, programs, policies, and education practices, the U.S. Department of Education's Institute of Education Sciences developed the What Works Clearinghouse (WWC). The primary mission of the WWC is "to be a central and trusted source of scientific evidence for what works in education" (What Works Clearinghouse, 2020). Depending on the degree of adherence to the WWC's standards, studies are rated as Meets WWC Design Standards Without Reservations, Meets WWC Design Standards With Reservations, or Does Not Meet WWC Design Standards. If a study Meets WWC Design Standards Without Reservations, the outcome variable must be measured repeatedly, the independent variable must be systematically manipulated, a second observer must also measure the outcome variable for at least 20 percent of data points per condition across all participants. In the case of alternating treatment designs (ATDs), baseline must have at least five data points, each comparison condition should have at least four data points, and each condition iteration should have no more than two data points before changing conditions. This study adhered to these guidelines.

Summary

Since the mid-twentieth century, behavioral studies and interventions have been based on classroom instruction. As more students move to online learning environments (Queens & Lewis, 2011; Taie & Goldring, 2017; Taie & Goldring, 2019; Taie & Goldring, 2020; U.S. Department of Education, 2004, 2008, 2012), more research is required to evaluate how instruction delivery may be better implemented in these environments.

The positive influence of contingent praise on student engagement has been extensively researched for decades for increasing student engagement in K-12 classrooms (Moore et al., 2018; Royer, Lane, Dunlap, & Ennis, 2019). While some research has been conducted on the effect of noncontingent praise on undergraduate student engagement in online learning environments (Rozsa, 2021), none has been conducted on the effect of contingent praise on undergraduate student engagement in online learning environments.

Classroom-based studies have traditionally used similar behavioral definitions for behavioral engagement, but there is not a consensus on defining behavioral engagement in online learning environments. As more people move to online learning as an alternative to traditional classroom-based instruction, the need for accurate online-based measures is also increasing. One of the most promising technologies in in-lab settings has been eye gaze tracking. As software is developed that may implement eye gaze tracking without the aid of additional hardware beyond what would typically be needed to access online learning, engagement data may now be accurately collected both in synchronous and asynchronous environments.

Purpose

The purpose of this study was to examine how remotely delivered contingent praise affects learners' visual engagement to video lectures, as measured by using readily available eye tracking technology. Subsequently, the research questions are:

Research Question 1: Does providing praise contingent on percent time visual engagement (as indicated by eye tracking measurement of gaze direction) affect overall percent visual engagement compared to no praise?

Research Question 2: How closely do the eye tracking program measurements of eye gaze correspond with hand-scored measurements?

CHAPTER II - METHODS

Participants and Setting

Before beginning the study, the primary investigator received approval from the University of Southern Mississippi Institutional Review Board (Appendix A). Participants were recruited from The University of Southern Mississippi during the Spring 2022 academic semester using flyers (Appendix B). Participants were paid $15 per hour of participation to be provided upon the discontinuation of their participation. To participate in the study, participants were required to be able to work at a computer screen without the aid of eyeglasses, for at least an hour. This was to increase the likelihood of the eye tracker correctly identifying the participant's pupil. Those who met the first criterion completed a questionnaire (Appendix C) that listed all the video topics. Potential participants rated each item on the questionnaire according to their perceived familiarity with the topic. Ratings were "not at all familiar," "a little familiar," "somewhat familiar," or "very familiar." To be considered for the study, a participant was required to be "somewhat familiar" or lower with 30 or more of the topics. These participants provided consent (Appendix D). Each participant was required to watch at least four videos lasting 10-15 mins. In a prior study using WebGazer with 10-15 min videos (Rozsa, 2021), the average visual engagement (weighted according to video duration) during baseline was calculated to be 79.25%. Participants continued to attend sessions until their average visual engagement was above 80% for two consecutive sessions. At that point, the participant was dismissed from the study. If visual engagement remained consistently below 80% average during baseline, individuals completed the rest of the study. Of eighteen screened participants, four met these criteria and completed the study. Relevant participant demographics for those who completed the study are presented in Table 1.

Table 1

Participant Demographics

Participants	Gender	Age	Major	Race
Annie	Female	20	Medical Laboratory Science	African American
Britta	Female	17	Public Relations in Advertising	Hispanic
Jeff	Male	19	Medical Laboratory Science	Caucasian
Shirley	Female	21	Psychology	African American

During sessions, the participant was in a 200 sq ft room in the School Psychology Program area of the School of Psychology. The researcher was placed in a nonadjacent room in the same building, about 100 ft away. All verbal interaction between participants and the researcher occurred over Zoom. The room in which the participant was located was made up of two tables, 5 chairs, an unused monitor, two shelves, and two lamps. Only the lamps were visible to the participant, who was seated at a table against the wall. Various potentially distracting stimuli were placed on the table with the participant, including a tablet with a continuously moving visual that changed colors, building blocks, magnetic letters, a puzzle ball, and a book. The two lamps were located behind the participant's laptop screen. The overhead lights remained off during trials, with motion sensors taped to prevent motion detection. The laptop was seated on 3 to 5 books, adjusted based on the participant's height so that the participant could view the screen without tilting his or her head. The laptop screen was two to three feet from the participant. Sessions lasted an average of 59.9 mins (*SD*=11.82).

Materials

Computer Hardware

A Dell Latitude 5580 and an N930AF 1080p webcam were used during the study. The Dell Latitude 5580 ran on an i7-7820HQ processor at 2.9 GHz, had 16GB of RAM, and used the Microsoft Windows 10 Pro Education operating system. The screen's diagonal length was 15 inches. The researcher used an HP Envy x360, which ran on an AMD Ryzen 7 5700U processor at 1.8 GHz, had 8GB of RAM, and used Microsoft Windows 10 Home. Both laptops had an integrated webcam and microphone.

Video Stimuli

All videos were obtained from Khan Academy (http://www.khanacademy.com), a site that provides academic instruction through videos that cover academic content for kindergarten to college level students. To be considered for use in the study, a video needed to cover some topic in the history of human civilizations between 5000 BCE and 2000 AD. The video also was required to be between 10 and 15 mins long. The lecturer had to be the same across all videos and not appear on-screen. The lecturer had to use a cursor or draw on the screen to emphasize text, timelines, or pictures. Thirty-seven videos were used, with an overall average of duration of 12.59 mins (*SD*=1.86).
Appendix E contains a list of the links and topics of videos that were used.

WebGazer

Eye gaze data were recorded using the *WebGazer* eye tracking library which is written in JavaScript (Papoutsaki, 2016). Whereas other eye tracking studies and applications typically require the installation of additional software or the use of additional equipment, *WebGazer* may be implemented with any website. *WebGazer* requires only permission to access the participant's webcam in order to identify the location of eyes, and subsequently to detect pupils and facial features. In order to calibrate the software, *WebGazer* uses cursor-gaze relationships to identify the position of eyes. When clicking on the computer screen, the user is typically focusing on the respective stimulus. For this study, multiple clickable dots appeared on the screen in random locations. Whenever the participant clicks a dot, the software identifies the location of the participant's eyes by using tracking.js (Lundgren et al., 2015), a facial feature detection library. The software first identifies the upper half of the face, then creates rectangular bounding boxes on the eyes. Next, the program sweeps the upper half of the face with a small-scale eye detection to approximate where the eyes are in order to minimize the likelihood of the program incorrectly identifying eye-like stimuli in the environment, such as a mole or a dot behind the participant. If the face is not immediately identifiable using the small-scale eye detection, full-image (large-scale) eye detection is used instead.

The pupil is the primary feature tracked when identifying where the participant is looking. First, the iris is identified, then the pupil. The program operates on three assumptions. First, the iris is darker than the rest of the eye. Second, the iris is circular. Third, the pupil is centrally located in the eye. With each click, each detected eye region is converted to 6x10 pixel images. Next, each image is grayscaled, retaining the intensity of each color. As a result, the lightest areas are whiter, while the darkest regions are black. This makes the parts of the eye easier to identify. This results in a 120D vector.

The vector and the coordinates of the clicked square are paired and fed into the linear regression algorithm is then updated.

Each subsequent click on a square increases the accuracy of the model, with the most recent data being more heavily weighted. It should be noted that in the original Papoutsaki (2016) study, the researchers used a 24-inch monitor. The researchers also required participants to engage in at least 40 mouse clicks during the task. The researchers estimated that gaze predictions had a mean error of about 175 pixels, or around 3 cm. This was the version used in the present study.

Current WebGazer Application

In the current study, the author created a script that incorporated the *WebGazer* script. Prior to each video, the participant clicked on 35 randomly placed squares to create the eye gaze prediction model to be used for that trial. The video for each trial appeared slightly skewed toward the lower right edge of the screen. The right edge and bottom edge of the video was about 2 inches and 1inch from the edge of the screen, respectively. The researcher created a programming code that placed a transparent dot that corresponds with where the participant is looking on screen. The dot's position was continuously updated on average 35.29 (*SD*=5.60) times per second. A transparent box was also created around the video. As the dot's position was updated, the program checked if the dot was within the box. If the dot was within the bounds of the box, a running tally of total visual engagement was increased by one and the total number of checks was increased by one. If the dot was not in the box, only the total number of checks were increased by one. During Praise and Demand conditions, intervals were 2 mins long. During these conditions, current interval checks and current interval visual engagement tallies were also recorded. Every 2 mins, the current interval's visual engagement tallies were divided by the current interval's number of checks. This percentage was compared against the average of the participant's visual engagement during baseline. A 1 cm x 1 cm transparent square (Square A) (See Figure 2) was placed behind the video on the participant's screen with about 2.5 mm outside the border of the video. A second transparent square (Square B) (Figure 2) was placed behind the video, slightly above Square A. The right side of Square B also protruded the video border by 2.5 mm. Both squares were always transparent except at the end of the 2-min intervals. Every 2 mins during the Demand condition, Square B appeared white for 3 s to signal the researcher to speak. During the Praise condition, every 2 mins the participant's baseline average visual engagement was compared to the current interval's average visual engagement. If the current visual engagement was above the baseline, Square A's protruding corner appeared white for 3 s, signaling the researcher to speak. The current interval's visual engagement tallies and current interval checks were reset for the next interval, while total checks and the total visual engagement continued to be recorded. At the end of each video, the total percent visual engagement was recorded.

Pavlovia

Because *WebGazer* required a website to run, a secure server to host the website was required. Pavlovia (https://pavlovia.org/docs/home/about) is a web server that was specifically designed to host online behavioral studies. Pavlovia provides a Health Insurance Portability and Accountability Act (HIPAA)-compliant web service that does not store personally identifiable information. The webpage was run on Google Chrome, a web browser that is readily available to the public.

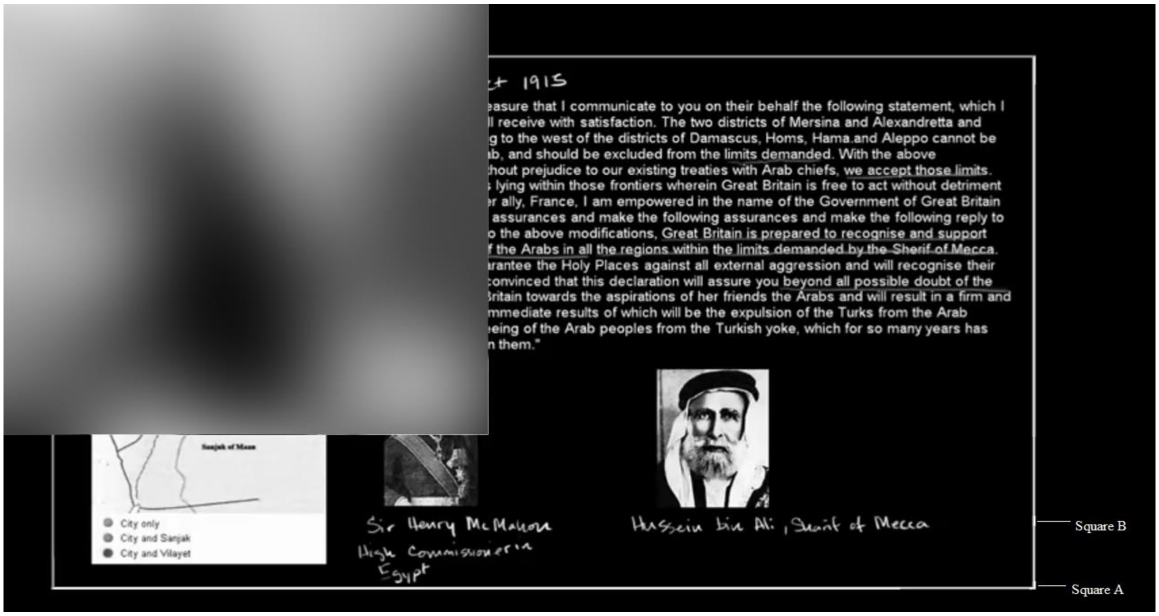

Figure 2. Example of researcher’s screen with Square A and Square B showing.

Zoom

The researcher and participant used Zoom (Zoom Video Communications Inc., 2016), a third-party video conferencing software. This software allowed the researcher to see and record the participant’s screen and face throughout the study.

Data Sheet

A data sheet with 90 blank boxes (Appendix F) was used for identifying whether the participant was visually engaged and during which 10-s interval the researcher delivered verbal feedback. as well as what type of verbal feedback was delivered to the participant.

Procedures

Sessions lasted for two to four trials each (M = 2.97, SD=.529). Trials began from when the researcher delivered the initial instructions until the video ended or 15 mins after the video had begun, elapsed, whichever occurred first. At the beginning of each trial, the researcher ensured that the equipment was working properly, distractions were present, the researcher's portrait was not visible to the participant, and that the participant's screen was visible on the researcher's screen. The study's website was on-screen at the start of each trial. The researcher instructed the participant to keep their speakers on, not to mute themselves, to take no notes, to stay at the computer, and to read aloud all directions on the screen. The participant input their assigned participant number and the current session number as indicated by the researcher. After reading through all directions, the participant calibrated the program by clicking on thirty-five squares that appeared on-screen in random locations. The participant then read the directions that appeared on-screen, then began the trial-specific video.

Baseline

Each participant initially experienced the baseline condition, in which the participant watched videos while the eye tracker assessed VE in the absence of verbal feedback from the researcher. This condition was conducted using the above procedure.

Praise Condition

This condition followed the same procedure as the baseline condition, except the researcher provided contingent praise to the participant. Upon the participant beginning the video, the researcher delivered a praise statement. Praise was defined as a verbal statement by the researcher indicating approval of the participant's engagement (e.g., "I love that you're watching the video," "Excellent job looking at the screen."). Recent literature indicates that praise delivered at a rate of as little as once per 2 mins can have a positive effect on academically-engaged behavior (e.g., Blaze, Olmi, Mercer, Dufrene, & Tingstrom, 2014). Based on these recent findings, contingent praise determined on a 2-min basis. Every 2 mins the researcher identified the outer edge of a white square (Square A) near the lower right corner of the video. If it appeared, this indicated that the participant was visually engaged at or above their average baseline level of percent visually engaged. The researcher then delivered praise within 10 s of the appearance of the white square.

Demand Condition

This condition was similar to the Praise condition, except the researcher noncontingently delivered demands instead of praise every 2 mins. Demands were defined as verbal statements instructing the participant to engage in an expected behavior, specifically looking at the video (e.g., "Watch the video," "Look at the screen."). Upon the participant beginning the video, the researcher delivered the first demand statement. Subsequently, every 2 mins a white square (Square B) appeared in the lower right corner of the video, just above the location of Square A. This signaled the researcher to deliver a demand. The researcher delivered a demand within 10 s of the square appearing throughout the video.

Design, Data Analysis, and Dependent Variable

During the study, participants were exposed to different conditions: Baseline, Demand, and Praise. All participants initially experienced the Baseline condition. Subsequent conditions were administered to each participant in a staggered multiple baseline with Randomized Block Design (RBD) (Edgington, 1967, 1980a; Onghena & Edgington, 1994, 2005). Each block was comprised of a Praise and Demand condition. Three blocks of Demand-Praise and three blocks of Praise-Demand were randomly ordered following the Baseline phase. Two participants experienced Demand-Praise first and two participants experienced Praise-Demand first. Ordering of the videos were randomized for each participant. This design *meets evidence standards* according to What Works Clearinghouse Design Standards (What Works Clearinghouse, 2020).

The first condition change occurred after a participant's five sequential data points in baseline exhibited low variability around the median. Low variability was determined using a "stability envelope" (Barton et al., 2018; Lane & Gast, 2013), which was defined by a creating a range of +/- 20% of the median. If at least 80% of the data points are within that envelope, the data are considered to have low variability. Subsequent condition changes for other participants occurred once the participant experienced at least three more baseline data points than the previous condition change, and five sequential baseline data points demonstrated low variability, also using a stability envelope.

Data were also visually analyzed for trend, level, immediacy of change, stability of data, and percent of non-overlapping data (PND; Scruggs et al., 1987). Percent of non-overlapping data was chosen for evaluating effect due to likelihood of ceiling effects encountered in past studies. Because one of the main research questions was regarding the extent to which *WebGazer* data and hand-scored data agreed, comparison of effect sizes may be masked by false negatives due to ceiling effects from other visual analysis techniques such as percent of all non-overlapping data (PAND; Parker et al., 2007). In accordance with analysis guidelines (Scruggs & Mastropieri, 1998), the independent variable was classified as being very effective if 90% or more data points fell higher than the highest data in baseline. If 70% to 90% of data fell higher than the baseline data, the effect was classified as effective. If 50% to 70% of data fell higher than the baseline data, the effect was considered questionable. For those with less than 50% of data falling above baseline data, the effect was classified as ineffective.

Visual Engagement

The primary dependent variable for this study was percent time visually engaged (VE), defined as the participant's eye gaze being directed toward the video during the lecture. Visual engagement was measured by a variation of the *WebGazer* JavaScript (Papoutsaki et al., 2016). For hand scoring, VE was defined more explicitly by the absence of specific behaviors. A participant was scored as not being visually engaged if their eyes were closed for more than 3 s continuously, if they were looking away from the screen so that the sclera of their eyes was not discernible, if their eyes were not visible due to moving their face out of the video, or if they were looking down toward the keyboard. Hand scoring was performed using a 10-s MTS procedure.

Procedural Integrity and Interobserver Agreement

To assess for procedural integrity, the primary researcher recorded whether the above procedure was followed using Appendix G, H, or I for respective conditions. Because distractions were not in the view of the camera, the primary researcher recorded if the distractions were present at the beginning of sessions (i.e., before a participant's first trial of the day) and if they were still present at the end of sessions (i.e., after the participant's final trial of the day). A second observer was trained by the primary investigator to record the dependent variable and to record whether steps were followed, using 4 videos, one from each participant and covering all 3 conditions. During training, if the second observer had not achieved at least 90% agreement with the primary researcher, the primary researcher would have explained where the errors occurred, and another video would be run. However, average IOA was 95.45% (Range: 93.9%-100%) for the dependent variable and 100% for procedural integrity IOA. Interobserver agreement was determined using scored interval IOA. The number of intervals scored for visual engagement were compared between the observer and the primary investigator. The number of agreements were divided by total number intervals, then multiplied by 100. The second observer also coded at least 30% of videos for each condition, across all participants. The observer was required to have above 80% agreement with the primary investigator in the video. Interobserver agreement was again determined using scored interval IOA.

For Annie, IOA for the dependent variable and procedural integrity were calculated for 40% of baseline trials. For Britta, IOA was calculated for 37.5% of baseline trials. For all other conditions, including across all other participants, 33% of trials were coded for IOA.

Annie's mean IOA for the dependent variable during baseline was 93.78% (range 92.10%-95.45%), Praise IOA was 100%, and Demand IOA was 100%. Procedural integrity IOA was 100% across all conditions. Britta's mean IOA for the dependent variable during baseline was 96.02% (range 93.94%-98.88%), Praise IOA was 98.88% (range 97.75%-100%), and Demand IOA was 96.15% (range 89.45%-96.15%). Procedural integrity IOA was 100% across all conditions. Shirley's mean IOA for the dependent variable during baseline was 92.78% (range 88.16%-96.83%), Praise and Demand IOA were 100%. Procedural integrity IOA was 100% across all conditions. Jeff's mean IOA for the dependent variable during baseline was 95.65% (range 89.77%-100%), Praise IOA was 99.20% (range 98.39%-100%), and Demand IOA was 99.16% (range 98.31%-100%). Procedural integrity IOA was 100% across all conditions.

CHAPTER III – RESULTS

WebGazer Visual Engagement

Annie's baseline data indicated a minimal to slightly increasing trend (Figure 3). Her baseline data indicated low to moderate variability (65.71%-83.89%) with a mean visual engagement of 76.07%. Her final five data points in baseline had a median of 77.52%, with all five falling within the 20% stability envelope (65.89% to 89.14%), indicating stable data. Annie experienced the Praise condition first upon leaving baseline. Her data did not immediately increase upon the condition's introduction. Her Praise data demonstrated moderate variability (74.02% - 98.09%) and a slight upward trend. Her average visual engagement during the Praise condition (*M*=86.20) was higher than during baseline (*M*=76.07). Of her Praise data, 50% did not overlap with baseline, indicating a questionable effect. However, all her data overlapped with the Demand condition.

Annie's Demand data immediately increased upon the condition's introduction, and demonstrated moderate variability (75.81% - 99.14%) and no apparent trend. Her average visual engagement during the Demand condition (*M*=89.58) was higher than during baseline (*M*=76.07). Of her Demand data, 83.33% did not overlap with baseline, indicating overall effectiveness, and 16.67% did not overlap with her Praise data. Overall, her visual engagement during Praise and Demand did increase above baseline, with a larger effect for the Demand condition.

Britta's baseline data indicated no apparent trend (Figure 3). Her baseline data indicated low to moderate variability (67.95%-86.57%) with a mean visual engagement of 75.59%. Britta's final five data points in baseline had a median of 78.31%, with all five falling within the 20% stability envelope (62.65% to 93.98%), indicating stable data.

Percent Visual Engagement by *WebGazer*

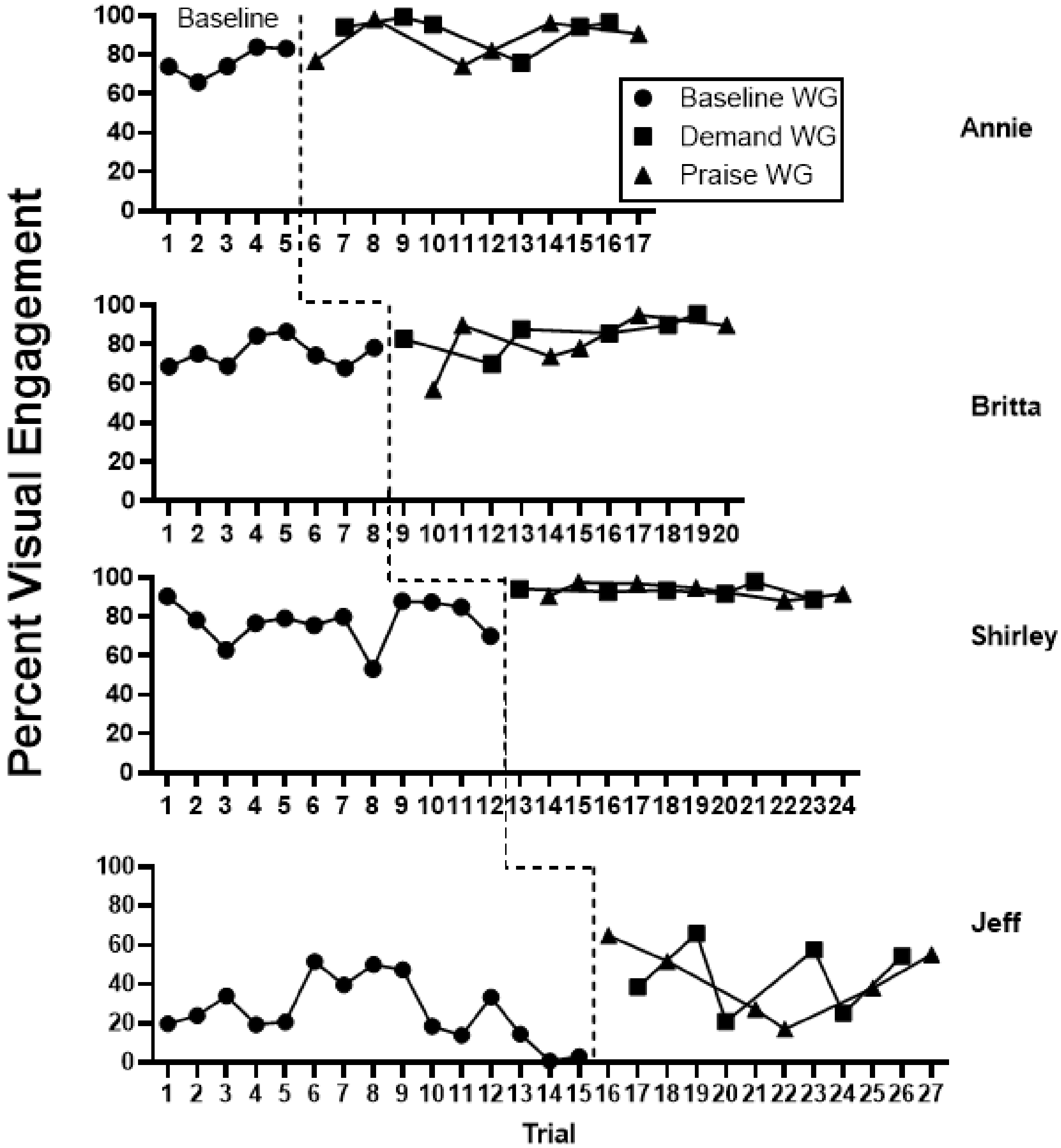

Figure 3. Percent visually engaged as measured by *WebGazer* (WG) application

Britta experienced the Demand condition first upon leaving baseline. Her data did not immediately increase upon the condition's introduction. Her Demand data demonstrated moderate to high variability (70.05% - 95.53%) and a slight upward trend. Her average visual engagement during the Demand condition (*M*=85.27) was higher than during baseline (*M*=75.59). Of her Demand data, 50% did not overlap with baseline, and 16.67% did not overlap with her Praise data.

Britta's Praise data demonstrated high variability (56.83% - 94.97%) and an upward trend. Her average visual engagement during the Praise condition (*M*=80.54) was higher than during baseline (*M*=75.59). Of her Praise data, 50% did not overlap with baseline, indicating the condition was overall ineffective. All her Praise data overlapped with the Demand condition. Overall, her visual engagement during Praise and Demand did increase above Baseline, with a similar effect for both compared to baseline, but with Demand having a slightly larger average change.

Jeff's baseline data indicated an overall decreasing trend (Figure 3). His baseline data indicated high variability (0.72%-51.63%) with a mean visual engagement of 26.08%. Jeff's final five data points in baseline had a median of 13.91%, with four of the five falling within the 20% stability envelope (11.83% to 16.70%), indicating stable data. Jeff experienced the Praise condition first upon leaving baseline. There was an immediate increase in data upon the condition's introduction. His Praise data demonstrated high variability (17.01% - 64.84%) and a slightly decreasing trend. His average visual engagement during the Praise condition (*M*=42.36) was higher than during baseline

(*M*=26.08). Of his Praise data, 50% did not overlap with baseline, indicating it was largely ineffective. However, all his data overlapped with the Demand condition.

His Demand data demonstrated high variability (21.02% - 65.87%) and no apparent trend. His average visual engagement during the Demand condition (*M*=43.81) was higher than during baseline (*M*=26.08). Of his Demand data, 50% did not overlap with baseline and 16.67% did not overlap with his Praise data. Overall, his visual engagement during Praise and Demand did increase above baseline, with a similar effect and average for both.

Shirley’s baseline data indicated no apparent trend (Figure 3). Her baseline data indicated high variability (53.21%-90.24%) with a mean visual engagement of 77.09%. Shirley’s final five data points in baseline had a median of 86.61%, with four of the five falling within the 20% stability envelope (69.3% to 103.93%), indicating stable data. Shirley experienced the Demand condition first upon leaving baseline. Her data immediately increased upon the condition’s introduction. Her Demand data exhibited low variability (88.65% - 97.80%) and no apparent trend. Her average visual engagement during the Demand condition (*M*=93.01) was higher than during baseline (*M*=77.09). Of her Demand data, 83.33% did not overlap with Baseline, indicating it was effective in increasing visual engagement, and 16.67% did not overlap with her Praise data.

Shirley’s Praise data demonstrated low variability (87.96% - 97.42%) and no apparent trend. Her average visual engagement during the Praise condition (*M*=93.15) was higher than during baseline (*M*=77.09). Of her Praise data, 83.33% did not overlap with Baseline, which indicated effectiveness in increasing visual engagement. However, all her data overlapped with the Demand condition. Overall, her visual engagement during Praise and Demand did increase above baseline, with a similar effect for both compared to baseline.

Hand Scored Visual Engagement

Annie's hand scored baseline data indicated no apparent trend (Figure 4). Her baseline data indicated low to moderate variability (69.84%-87.67%) with a mean visual engagement of 79.05%. Annie experienced the Praise condition first upon leaving baseline. Her data immediately increased upon the condition's introduction. Her Praise data demonstrated low variability (98.50% - 100.00%) and no apparent trend. Her average visual engagement during the Praise condition (*M*=99.75) was higher than during baseline (*M*=79.05). Of her Praise data, there was no overlap with baseline, indicating it was very effective. All data overlapped with Demand condition.

Annie's Demand data demonstrated low variability (98.30 – 100.00%) and no apparent trend. Her average visual engagement during the Demand condition (*M*=99.72) was higher than during baseline (*M*=79.05). Of her Demand data, there was no overlap with baseline and all data overlapped with her Praise data. Overall, the hand scored visual engagement during Praise and Demand increased above baseline, with a similar effect.

Britta's baseline data indicated slightly decreasing trend (Figure 3). Her baseline data indicated low to moderate variability (83.13%-94.37%) with a mean visual engagement of 90.81%. Britta experienced the Demand condition first upon leaving baseline. Her data immediately increased upon the condition's introduction. Demand data demonstrated low variability (92.65% - 98.53%) and no apparent trend. Average visual engagement during the Demand condition (*M*=96.60) was higher than during baseline (*M*=90.81).

Hand Scored Percent Visually Engaged

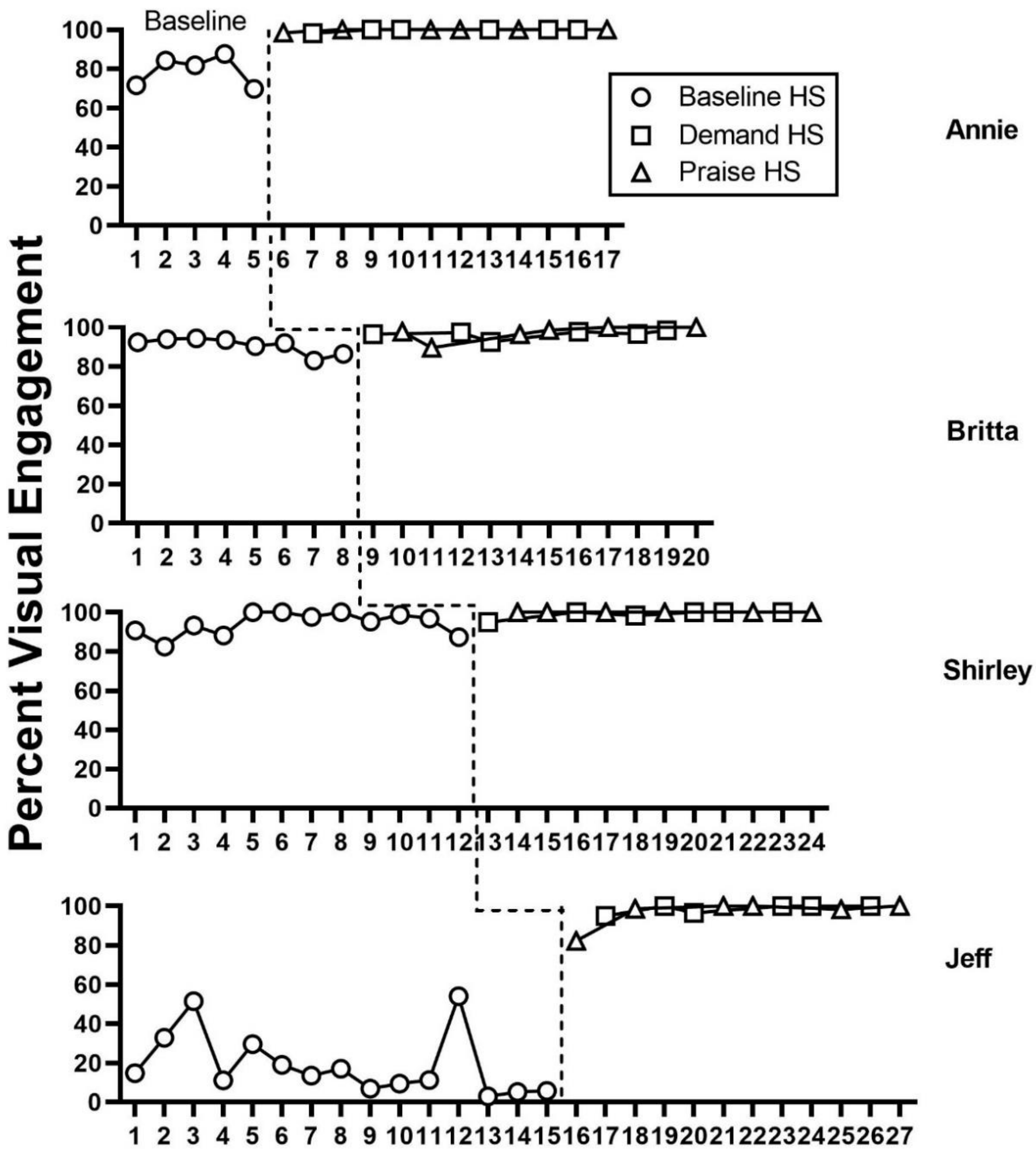

Figure 4. Percent visual engagement as measured by hand scoring (HS)

Of her Demand data, 83.33% did not overlap with baseline, indicating, but all her data overlapped with her Praise data. Britta's Praise data demonstrated low variability (89.66% - 100.00%) and no apparent trend. Her average visual engagement during the Praise condition (*M*=97.11) was higher than during baseline (*M*=90.81). Of her Praise data, 83.33% did not overlap with Baseline, indicating it was effective. Compared to her Demand condition data, 50% of her Praise data did not overlap. Overall, her visual engagement during Praise and Demand did increase above baseline, with a similar effect for both compared to baseline.

Jeff's hand scored baseline data indicated no apparent trend (Figure 3). His baseline data indicated high variability (2.94%-54.02%) with a mean visual engagement of 18.97%. Jeff experienced the Praise condition first upon leaving baseline. His data immediately increased upon the condition's introduction. His Praise data demonstrated low to moderate variability (82.26% - 100.00%) and a slight upward trend. His average visual engagement during the Praise condition (*M*=96.54) was higher than during baseline (*M*=18.97). Of his Praise data, 100% did not overlap with baseline, indicating it was very effective. However, all his data overlapped with the Demand condition.

Jeff's Demand data demonstrated low variability (95.00 – 100.00%) and no apparent trend. His average visual engagement during the Demand condition (*M*=98.60) was higher than during baseline (*M*=18.97). Of his Demand data, 100% did not overlap with baseline, and all data overlapped with his Praise data. Overall, Jeff's hand scored visual engagement during Praise and Demand did increase above baseline, with a similar effect for both conditions.

Shirley's baseline data indicated no apparent trend (Figure 3). Her baseline data indicated high variability (94.92 %-100.00%) with a mean visual engagement of 98.89%. Shirley's final five data points in baseline had a median of 78.26%, with four of the five falling within the 20% stability envelope (62.6% to 93.91%), indicating stable data. Shirley experienced the Demand condition first upon leaving baseline. Her data did not immediately increase upon the condition's introduction. Her Demand data demonstrated low variability (94.92 – 100.00%) and no apparent trend. Her average visual engagement during the Demand condition (*M*=98.89) was higher than during baseline (*M*=94.23). Of her Demand data, all overlapped with baseline, and all overlapped with her Praise data.

Shirley's Praise data demonstrated no variability (100.00% - 100.00%) and no apparent trend. Her average visual engagement during the Praise condition (*M*=100.00) was higher than during baseline (*M*=94.23). Of her Praise data, all overlapped with baseline and all her Praise data overlapped with the Demand condition. Overall, her hand scored visual engagement during Praise and Demand indicated both were ineffective.

WebGazer versus Hand Scored visual engagement

The second research question posed for this study was to what degree did visual engagement, as measured by *WebGazer*, correspond with visual engagement measured by an observer? Differences were expected due to the differences in frequency of observations (i.e., dozens per second by *WebGazer* versus one per 10-s by observer). As a result, variability, trend, and level were examined to account for potential systematic errors. Comparisons between visual engagement in the Praise condition as measured by *WebGazer* and hand scoring are graphed in Figure 4. Comparisons between visual engagement in the Demand condition as measured by *WebGazer* and hand scoring are graphed in Figure 5. Most conditions showed similar variability between the hand scored and computer scored data. The largest discrepancies in variability between *WebGazer* and hand scoring were observed in Britta's Praise condition (Range=56.83-94.97, 89.66 – 100.00), Jeff's Demand (Range=21.02-65.87, 95.00-100.00) and Praise (Range=17.01-64.84, 82.26-100.00) conditions, and Shirley's Baseline (53.21-90.24, 82.54-100.00).

Regarding level, most conditions across all participants were similar between *WebGazer* and hand scoring measurements of visual engagement. The largest discrepancies between *WebGazer* and hand scoring were observed for Jeff's Praise condition (44.94, 99.33) and Demand condition (46.58, 100.00), Britta's Baseline (74.79, 92.24), and Shirley's Baseline (78.57, 96.06).

Among the discrepancies, Jeff's Demand and Praise conditions and Shirley's Baseline were discrepant both in their range and their level. The data paths for Shirley's Baseline comparison had similar shapes, but the trends were similar. Using a simple

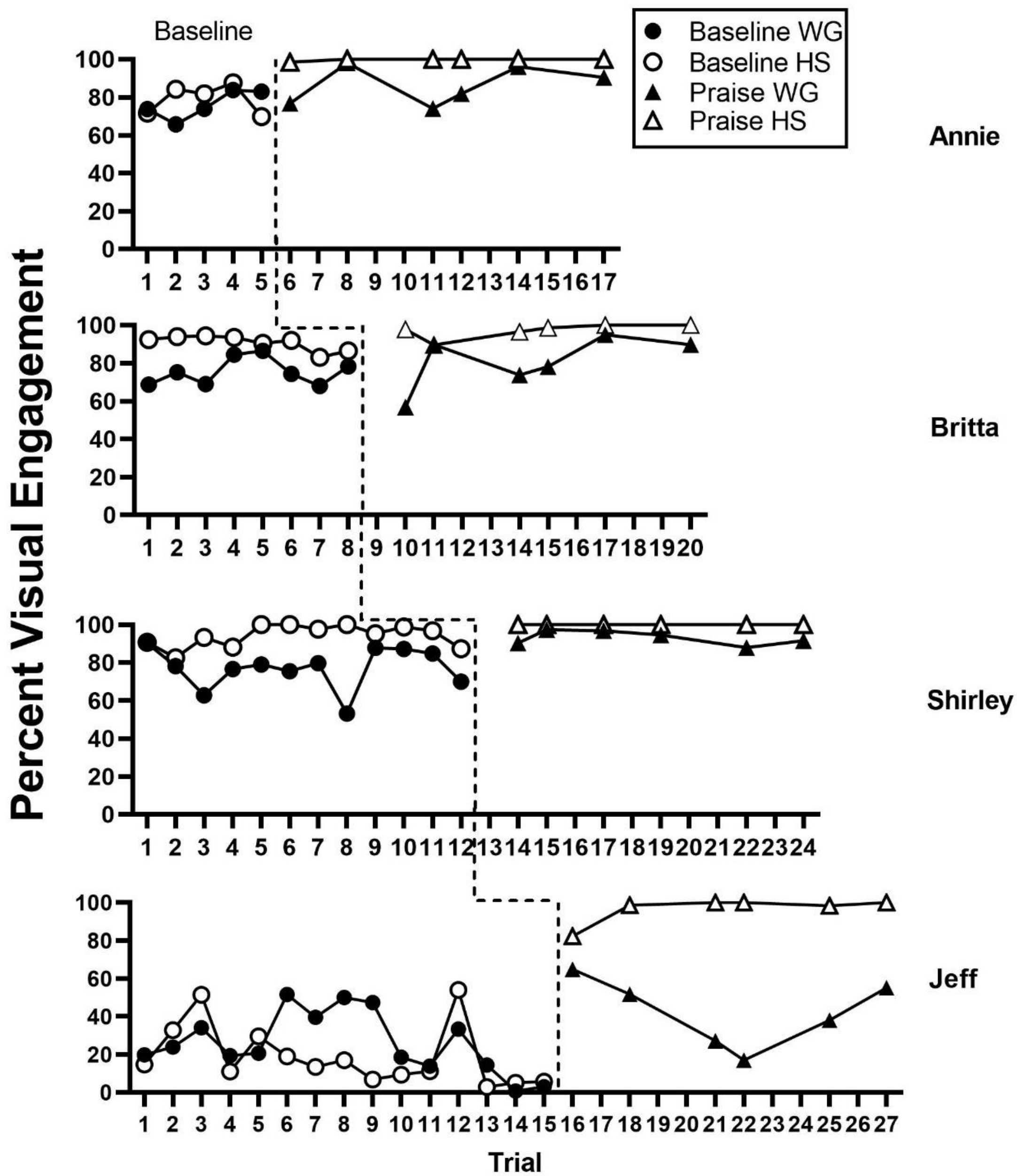

Figure 5. Comparison between hand scored (HS) and *WebGazer* (WG) Praise data.

Demand Hand Scored versus *WebGazer*

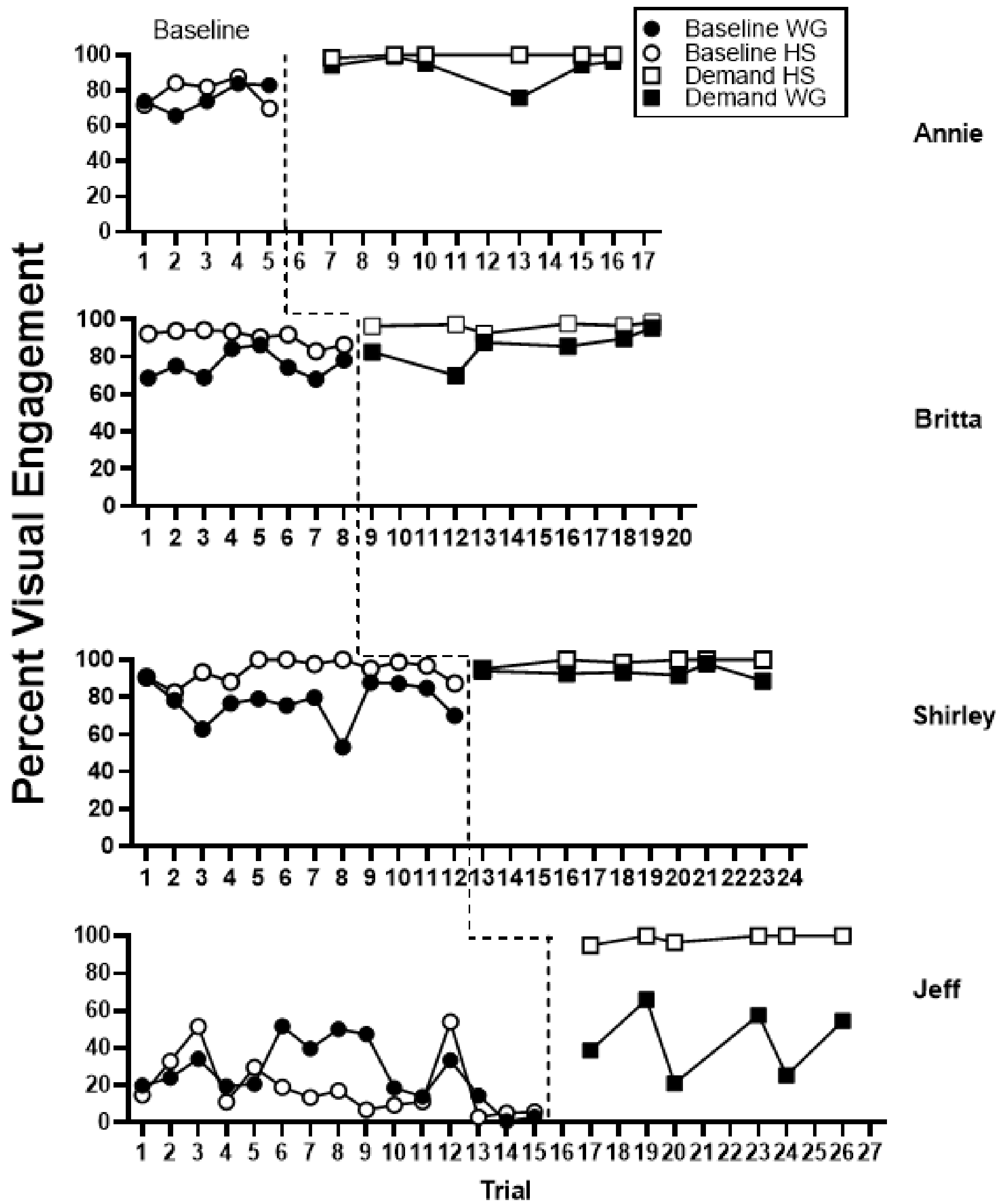

Figure 6. Comparison between hand scored (HS) and *WebGazer* (WG) Demand data.

linear regression, the slope of Shirley's baseline data was -0.03 for *WebGazer* measurement and 0.57 hand scoring. Although these slopes are in opposite directions, they are low numbers, and thus relatively close. Using a simple linear regression for Jeff's Demand condition data, the slopes for *WebGazer* and hand scoring were 0.0081 and -0.0019, respectively. These slopes indicated no apparent trend. Lastly, Jeff's Praise condition data resulted in slopes of -0.028 for *WebGazer* and 0.025 for hand scoring. Again, these were relatively close, indicating no apparent trend.

CHAPTER IV – DISCUSSION

This study was conducted to examine two questions. First, does providing contingent praise affect participants' visual engagement, as measured by eye tracking software, during video lectures? Second, how closely do the eye tracking program measurements of eye gaze correspond with hand-scored measurements?

The first research question examined the extent to which participants' percent time visually engaged, as measured by the *WebGazer* software, changed in the presence of contingent praise. All participants' visual engagement increased in the Praise condition, with three participants' data indicating questionable effectiveness and one participant's data indicating overall effectiveness. However, all data measured during Praise condition overlapped with data in the Demand condition for all participants, indicating that Praise was not discernibly better than Demand for increasing visual engagement. The level for the Praise condition also increased across all participants. The variability of *WebGazer*'s measurements means this conclusion should be interpreted with caution. However, because *WebGazer* measures hundreds more times in a 10-s period than the human observer, the software may have identified more incidents of disengagement. Without continual calibration during trials, *WebGazer*'s measurement may also be affected by head movements and repositioning, even if gradual. In answer to this first question, contingent praise had a questionable effect for three participants in the presence of contingent praise, and was overall effective for one participant. These findings are consistent with results from the Rozsa (2021) study, in which three out of four participants exhibited higher visual engagement in the presence of noncontingent praise. Like the Demand condition data in the current study, Rozsa (2021) also observed significant overlap with neutral verbalizations, finding similar effect for each.

The second research question required a comparison between the hand-scored data and data measured by *WebGazer*. When looking at the hand-scored data, both Shirley's and Britta's Baseline data averaged above 90%. Although a large effect was observed in Britta's Praise data, any difference in magnitude of effect size between her Demand and Praise data may have been masked due to potential ceiling effects. Shirley's Baseline data averaged in the mid-90s, with three trials being measured at 100% visual engagement. This resulted in no treatment effect being observable. It should be noted that the overall average increased across both conditions, and the variability greatly reduced, with her Praise data being at 100% across all Praise trials. This change in variability may indicate more consistent visual engagement, but due to the differences between hand score and *WebGazer*, this is not necessarily true. Overall, both hand-scored data and *WebGazer* measured data indicated a treatment effect for three participants, though not to the same magnitude. Despite the differences in magnitude of effect size between *WebGazer* and hand scoring, the similarity in trend, shape, level, and variability across data paths for Annie, Britta, and Shirley suggest that *WebGazer*'s measurements were likely valid representations of the respective participants' visual engagement.

The largest discrepancies (more than 40%) were observed in Britta's 10th trial, Shirley's 8th trial, and all of Jeff's Praise and Demand data, except one Praise and one Demand trial. Videos were reviewed, the percent visual engagement was calculated for each 2-min interval to identify artifacts or confounding variables responsible for disparities in measurements.

Britta’s outlying session appeared to be a result of her putting up her hair halfway through the video. Prior to putting up her hair, her average percent visual engagement was between 70% and 95%. Therefore, changing the features of her face may have reduced the tracker’s accuracy. When observing Shirley’s outlying sessions, *WebGazer* measured most of her visual engagement occurring during the first half of the video. As the video progressed, the participant started tilting her head and leaning forcefully into her hand, pushing upward on her cheek, forcing one of her eyes closed. Similar to Britta’s video, this appeared to contribute to the disparity in measurements.

Disparities in Jeff’s videos were not as apparent upon review. One of the most notable differences in appearance between Jeff and other participants was the presence of a large, dark beard. During videos in which Jeff’s *WebGazer* data were most discrepant with hand scored data (i.e., trials 20, 21, 22, and 24), Jeff almost immediately tilted his head at roughly a 30–45-degree angle upon the video beginning. In multiple trials (i.e., trials 18, 19, 20, 21) Jeff covered his mouth with his hand in a thinking pose for large portions of the video. During his baseline trials, Jeff typically immediately looked away from the video upon it beginning and started solving the nearby puzzle ball or using his phone. When asked about from what source he had learned of the study, Jeff said his roommate was dismissed from the study a couple days earlier after being screened as a potential participant. In order to continue to be a participant, Jeff may have engaged in aberrant behavior during baseline. This will be discussed in more depth in the limitations section.

Overall, the results of this study demonstrated similar results with that of a prior study which used this version of *WebGazer* (Rozsa, 2021). In both studies, the majority of participants' visual engagement increased in the presence of praise. Both studies also demonstrated observed disparities in level and variability between *WebGazer* and hand-scoring. In the previous study, ceiling effects during baseline likely masked any potential treatment effects. This study attempted to address that limitation by introducing a maximum average visual engagement of 80% during baseline as an exclusionary criterion. Unlike the previous study, this study employed an alternating treatment design to enable easier comparisons between conditions and baseline data. This study also employed more potentially distracting stimuli in the participants' immediate vicinity, whereas the previous study had minimal alternative stimuli with which to engage. Whereas the previous study used noncontingent praise with neutral verbalizations by the experimenter, this study used contingent reinforcement with demands, as both interactions were expected to be more likely to be encountered in an academic environment. Despite some disparities in data between hand scoring and *WebGazer* measured visual engagement, the advantages of having an online software collecting and calculating data live is invaluable.

Limitations

This study had multiple limitations that should be considered. Similar to other studies using similar eye tracking software (e.g., Hutt, 2020; Papoutsaki, 2018; Rozsa, 2021), eye gaze estimation is susceptible to head movements. Many older eye tracking methodologies even incorporated headrests (See Rayner, 1998 for a more comprehensive history of these methodologies). As a result, requirements of minimal head movements and clear view of the participants' eyes may limit its validity outside the experimental setting. The combination of multiple reminders about these requirements prior to each trial as well as the presence of two cameras may also have affected the attentiveness of participants. Multiple participants looked directly at the webcam multiple times, then often quickly overcorrected by looking around the room. As a result, this reactivity may have affected the results.

A few of the first participants, whom were dismissed for not meeting inclusion criteria, encountered motion-activated overhead lights mid-trial, which made their eyes difficult to see due to shadows cast by these lights. This was addressed by taping over the motion sensors and placing the two lamps behind the laptop to increase the visibility of the participants' eyes. This need for precise lighting has been noted in past studies (Papoutsaki, 2016; Rozsa, 2021). Typical eye tracking software requires more expensive hardware that uses infrared light to identify the cornea continuously. This results in less reliance on precise lighting (Inhoff & Radach, 1998). To offset this disadvantage, most studies that use *WebGazer* incorporate continually updating calibration through user mouse clicks (e.g., Papoutsaki et al., 2017). This can help adjust the prediction model to account for changes in lighting, as well as changes in the participant's position.

Because the study was conducted in a controlled setting, not all components of a typical online learning environment were present, and some, such as the second webcam, were extraneous to a typical online learning environment. For example, the lecturer would likely also be the person delivering praise or demands in a synchronous online learning environment. Because the researcher interrupted the lecture to provide demands or praise, this may have made such interactions stand out more, or the video audio may have made the researcher's verbalizations harder to understand. In a previous study, no discernible difference was observed when neutral verbalizations were provided versus praise (Rozsa, 2021). That study also had a third party providing the verbal interactions apart from the lecturer and participant. This study demonstrated similar findings. Rather than the type of interaction moderating the relationship between the independent variable and visual engagement, it may have been the interaction with a separate party itself. Once participants were aware they were being actively watched, following the first verbal interaction, this may have been enough to increase visual engagement. In a typical classroom or online learning environment, other students may also receive attention, reducing reactivity due to being observed. The addition of the square as a visual cue for the researcher may also have influenced the participants' visual engagement. In contrast with Rozsa (2021), the current study's use of contingent praise resulted in fewer instances of praise. For Jeff, this may have been partially responsible for his lower scores during Praise condition, as he was not contacting reinforcement as regularly as other participants.

Another limitation is the pool of selected participants. All but one of the participants was recruited through a psychology class. Although results were not relayed to professors in the classes, participants may have engaged in behavior that is not typical, being more attentive during trials. Furthermore, because most of the participants were in psychology classes, they may have interacted with one another, either due to being in a same class or sharing a building that houses classes of that department. For the one participant, Jeff, whom was not recruited through a class, he said he volunteered after his roommate, a dismissed participant, informed him about the study. The roommate and other participants may have identified the criteria for inclusion in the study, or the aims of the study. This may have accounted for Jeff's exceptionally low baseline data, and subsequent near 100% data during the other conditions. As a result, his data should be interpreted with caution.

Due to the small participant pool, differences between male and female participants may have also contributed to the results. Past research has indicated that when praise more consistently results in an increase in male task performance than females, for whom task performance may even decrease when praised (Carone, 1975; Deci, 1972; Deci et al., 1975; Koestner et al., 1987; Zhao & Huang, 2019; Zinser et al., 1982). Some research has also indicated a greater decrease in performance in female undergraduate participants when the researcher providing praise was female than when the researcher was a male, and a greater increase in performance for male undergraduate participants when the research was a female (Deci et al., 1975; Lessard et al., 2015). If the researcher was female, the increase in visual engagement for female participants may have been less pronounced or even decreased. Relatedly, sexual orientation may have influenced results, given the difference in performance when experimenter and participant are a different or the same gender.

Another limitation was that subjects' individual interest in topics was not considered. Although reduced familiarity with the vast majority of topics was a requirement for inclusion, this did not necessarily guarantee individual interest in the subjects. Even though the individual may not have been familiar with the rise of Hitler, they may find war-related information more interesting. Similarly, though they may be familiar with Islam or Confucius, they may find religion in general interesting.

Another limitation was the praise itself. For praise to be effective, it must have certain qualities. First, it should be delivered contingent on performance of the desired behavior. Second, the person praising should specify what behavior is being praise. Third, the praise should sound sincere (Alberto & Troutman, 2009; O'Leary & O'Leary, 1977). The contingency for praise in this study was that the participant was visually engaged above their baseline average for 2 mins. The rate at which the contingency was evaluated was chosen based on previous research that demonstrated one praise statement per 2 mins typically results an effect on academic engagement for many students, but it does not always result in an effect for every student (Blaze et al., 2014; Williamson, 2017). Though some research has explored ideal rates of praise, these studies almost exclusively examined primary and secondary education settings (e.g., Allday et al., 2012; Dufrene, Lestremau, & Zoder-Martell, 2014; Sutherland, Wehby, & Copeland, 2000). Because the 2-min intervals were evaluated as a whole, when praise was delivered it was not necessarily when the individual was visually engaged. Furthermore, if the participant was not visually engaged above their baseline average for the first couple intervals, they were not reinforced, which may have a snowball effect, meaning the person is less and less likely to be visually engage. The sincerity of praise was also not evaluated. Because the researcher had to speak quickly and above the video, as well as from out of view of the participant, the tone was often unintelligible and nonverbal signals that may suggest sincerity was not available to the participant.

The length of videos should also be considered a potential limitation. Some studies have indicated that attention during instruction may begin to decrease after about 9 mins (e.g., Davis, 1993; Guo, Kim, & Rubin, 2014; McKeachie, 1986; Wankat, 2002). Most studies commonly cite Hartley and Davies (1978) when discussing the 10-15 mins rule. The authors of that study examined the amount of note taking at different times in lectures. As previously mentioned, this type of active engagement with material is demonstrably different from more passive visual engagement. Other studies have suggested that attention may first begin to decrease after 5 mins, with a further decrease in attention after 10-18 mins into class (Johnstone & Percival, 1976). Although no consistent threshold has been established for the ideal length of a video to maintain viewer's attention, the videos in this study were typically a little longer than the above times, longer videos may be necessary to control for naturally higher sustained attention for the first 10-15 mins.

Visual engagement was also narrowly defined as the participant looking toward the video. Sustained eye contact may not always indicate attending. For example, individuals tend to fixate on fewer regions, for longer times, and on irrelevant regions during wind wandering (Reichle et al., 2010). Saccades (eye movements between fixations) also become less frequent and/or slower (Uzzaman & Joordens, 2011). What constitutes an acceptable fixation duration or saccade length also varies on an individual basis and may change based on the type of task, such as silent reading versus scene perception (Rayner, 2009).

Future Directions

Future research into other forms of contingent reinforcement may also prove beneficial. As was stated previously, in K-12 classroom settings one praise per 2-mins often results in increased academic behavior. Part of the rationale was that it was easier for teachers in classrooms to provide praise at this rate (Blaze et al., 2014). Little research exists regarding the effects of praise on young adults. Some research has indicated that behavior specific praise may have a similar effect on young adults as grade-school-aged children (Hancock, 2000; Lessard et al., 2015). Conversely, others have indicated that person- or ability-centered praise has a positive correlation with increased performance compared to those who receive behavior specific praise or no praise (e.g., Koestner et al., 1987). If praise does not have the same effect on all undergraduate students, then alternative sources of reinforcement may be needed, depending on the individual. While this study sought to use positive reinforcement, negative reinforcement may also have been effective. If finishing a lecture was important to the student, perhaps a contingency could be put in place to pause the video after detecting no visual engagement for a set amount of time. Similarly, some professional workshops or lecture series employ clickable prompts to identify if the viewer is still watching. Some pilot studies have demonstrated the utility of eye tracking in the classroom, and its use as a means of combatting mind wandering (e.g., Hutt, 2020; Hutt et al., 2019). These studies used prediction models that would intervene with a question related to the material on-screen when probable mind wandering was detected. Students reported feeling less autonomous with this approach, and these studies still used more expensive eye tracking hardware. Future applications could combine the studies' approaches so that when the individual was not engaged, a prompt may appear, which would also serve to continue calibrating the software. Similarly, with the increasing use of computers in classrooms, the use of *WebGazer* with an integrated webcam to identify if a child is engaged during assignments may assist teachers in identifying who is working, who may be struggling, and if students should be praised for working on their assignments.

Future studies may also benefit from fixed momentary differential reinforcement of other behavior (Cooper, Heron, & Heward, 2019), in which the participant is observed after a specified duration. If the individual is engaging in the desired behavior, they are provided reinforcement. If not, then nothing occurs. This may increase the likelihood of visual engagement being reinforced, rather than looking at the entire interval.

Implementing this software in a multi-student online learning environment would also be a logical next step. This would require some form of signal that was bigger than the small square used in this study, as multiple students' videos would be present at once. Finding other unintrusive signals for the teacher that don't interfere with the lecture would also be helpful.

Conclusion

These results provide additional evidence for use in applied settings. Just as praise has been shown to be an effective classroom management technique for decades, it appears praise may also be useful in online learning environments. As academia moves to more online options, behavioral strategies should also continue to increase. Technology may be used beyond simply serving as a medium for accessing online instruction and material. With hardware continually improving at an exponential rate, resources that were once completely unavailable to the public are becoming more accessible. This study has demonstrated how technology may be used to assist us in data collection methodologies, as well as calculating whether students are meeting expectations based on predetermined criteria, such as portion of time spent on-task. This may allow the teacher more freedom to attend to other curriculum-based duties.

This study provided one example of how eye tracking may be used to increase academically relevant behavior. With a few modifications to the methodology, this software may be used to facilitate data collection as well as intervention implementation both in the classroom and in online learning environments. The implications of this study suggest continued application would result in fruitful opportunities for future researchers and educators.

APPENDIX A – IRB Approval Letter

Office of Research Integrity

THE UNIVERSITY OF SOUTHERN MISSISSIPPI

118 COLLEGE DRIVE #5125 • HATTIESBURG, MS | 601.266.6576 | USM.EDU/ORI

NOTICE OF INSTITUTIONAL REVIEW BOARD ACTION

The project below has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services regulations (45 CFR Part 46), and University Policy to ensure:

- The risks to subjects are minimized and reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered involving risks to subjects must be reported immediately. Problems should be reported to ORI via the Incident template on Cayuse IRB.
- The period of approval is twelve months. An application for renewal must be submitted for projects exceeding twelve months.

PROTOCOL NUMBER: IRB-21-292
PROJECT TITLE: Using Verbal Reinforcement to Increase Visual Attending in an Asynchronous Online Learning Environment: An Eye Tracking Study
SCHOOL/PROGRAM: School of Psychology, Psychology
RESEARCHER(S): Andrew Rozsa, D Olmi

IRB COMMITTEE ACTION: Approved
CATEGORY: Expedited

7. Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

PERIOD OF APPROVAL: July 16, 2021

Donald Sacco, Jr.

Donald Sacco, Ph.D.
Institutional Review Board Chairperson

APPENDIX B– Recruitment Flyer

Undergraduate Students Needed for Eye Tracking Study

What?

The purpose of this study is to identify how different forms of researcher's verbal feedback may alter participant's attending.

Who?

- Must be a USM undergraduate student
- Must be able to work at a computer for an hour without the aid of eyeglasses
- You are not registered with the USM Office of Disability Accommodations.

Why?

You will be compensated **$15 per hour** upon completion of participation or upon the study's completion, whichever comes first. You may stop participating at any time.

When?

Sessions will begin **February 1st** and continue until the project is completed. Each session will last about **one hour**, during which you will watch 3-4 lectures about human history. You should expect to commit **up to 8 hours** spread across **2 to 5 weeks**. To finish in a timely manner, we will meet at least twice a week, depending on your availability.

How?

Confirm you meet the above criteria by scanning this QR code with your camera and filling out the required information on the qualtrics page that appears:

Alternatively, you may go directly to the site at

https://tinyurl.com/ajreyetrack

After submission, you will receive a survey by email. Please fill this out by **January 30th**. Participants who meet criteria will be contacted within one to two days of submission.

Please direct all questions to the primary researcher at **andrew.rozsa@usm.edu.**

This study has been approved by USM's Institutional Review Board (Protocol #: 21-381)

APPENDIX C – Pre-Study Questionnaire

Please indicate how familiar you consider yourself with each of the following topics on a scale from "Not at all familiar" to "Incredibly familiar."

	Not at all familiar	Little familiarity	Some familiarity	Very familiar
Golden Age of Athens, Pericles and Greek Culture				
Paris Peace Conference and Treaty of Versailles				
International Human Rights				
Hinduism: Brahman, Atman, Samsara, and Moksha				
Augustus of Rome				
Feudal System during the Middle Ages				
Ottoman, Safavid and Mughal Empires				
Indus River Valley Civilizations				
Golden age of Islam				
Ides of March and civil war				
Socrates, Plato, and Aristotle				
Punic Wars between Rome and Carthage				
Initial Rise of Hitler and the Nazis				
Theodor Herzl and the birth of political Zionism				
Alexander the Great				
Protestant Reformation: Martin Luther				
Arian Controversy and the Council of Nicaea				
Fall of the Roman Empire				
Closing Stages in World War I				
Hitler and the Nazis come to power				
Blockades, U-boats, Lusitania				
Continuity-Sikhism connections to Hinduism and Islam				
Spread of Islam				
Hittite Empire and Battle of Kadesh				
Napoleon and the Wars of the First and Second Coalitions				
Bay of Pigs Invasion				
Cyrus the Great and the Achaemenid Empire				
Ancient Egypt				
French Revolution				
Allende and Pinochet in Chile				
Confucius and the Hundred Schools of Thought				
Axis in World War II				

APPENDIX D – Consent Form

ORI | Office of Research Integrity

INSTITUTIONAL REVIEW BOARD

STANDARD (SIGNED) INFORMED CONSENT

STANDARD (SIGNED) INFORMED CONSENT PROCEDURES

This completed document must be signed by each consenting research participant.

- The Project Information and Research Description sections of this form should be completed by the Principal Investigator before submitting this form for IRB approval.
- Signed copies of the consent form should be provided to all participants.

Last Edited Feb. 3rd, 2021

Today's date:

PROJECT INFORMATION

Project Title: Visual Engagement during Online Lectures		
Principal Investigator: Andrew Rozsa	Phone: 2056168787	Email: andrew.rozsa@usm.edu
College: Education and Human Sciences	School and Program: School Psychology Doctoral Program at the Department of Psychology	

RESEARCH DESCRIPTION

1. **Purpose**:

The purpose of this study will be to identify how different forms of verbal feedback may alter attending during video lectures.

2. **Description of Study:**

Sessions will occur in February, March, April, and/or May. During the trials, you will be recorded for data verification and to assess procedural integrity. If you agree to participate in this study, you will be seated in a room on-campus with a laptop equipped with a webcam, which will be provided by the researcher. After a short calibration sequence, you will watch a randomly chosen lecture covering a historical event/period in human history. The lecture will last between 10 and 15 minutes. During the lecture, the researcher will occasionally provide you with feedback. From calibration through the end of the video, you will need to minimize head movements as much as possible. From start to finish, each trial will last about 15-20 minutes. You should expect to commit up to 8 hours spread across 2 to 5 weeks. To finish in a timely manner, we will meet at least twice a week, depending on your availability. Sessions will begin February 1st and continue until the project is completed. You may discontinue participation at any point during the study.

3. **Benefits:**

You will be compensated $15 per hour upon completion of participation. If you are not chosen to continue the remainder of the study or choose to leave the study early, you will be compensated upon the study's conclusion at $15 per hour for time attended. You may discontinue participation at any point.

4. **Risks:**

Due to the nature of the experiment requiring you to be seated and looking at a computer screen for prolonged periods of time, back pain and eye strain may occur. To decrease the likelihood of these problems, sessions will not last beyond one hour. To minimize likelihood of injury or discomfort, you will be allowed to stretch after each trial. Due to the current state of the pandemic, transmission of COVID-19 may also be possible while on-campus. To minimize this likelihood you will be required to wear a mask during all face-to-face interactions. The researcher will also engage in social distancing and masking in accordance with university policies.

5. **Confidentiality:**

During trials, your eye gaze will be tracked and recorded using numerical coordinates. These will be stored in an Excel worksheet that is stored on Pavlovia, an experiment hosting web server. This server is Health Insurance Portability and Accountability Act (HIPAA)-compliant and does not store any personally identifiable information beyond IP address for security reasons. All data will be copied from the Pavlovia server following each session and removed from the site.

Videos will also be recorded for interobserver agreement and assessment of procedural integrity. These will be recorded only on the provided laptop and stored on two encrypted thumb drives which will be stored on-campus. These thumb drives will be accessible only to volunteer observers and the primary researcher. At the conclusion of the study, all data will be stored up to a year on-campus in a locked room accessible only to faculty.

At the study's conclusion, you will be asked for demographic information regarding age, race/ethnicity, self identified gender, current major, and your undergraduate year. This information will be available only to the primary researcher for statistics reporting and will be discarded a month following the study's completion.

6. **Alternative Procedures:**

Due to the exploratory nature of this study, alternatives are not being offered.

7. **Participant's Assurance:**

This project and this consent form have been reviewed by USM's Institutional Review Board, which ensures that research projects involving human subjects follow federal regulations. Any questions or concerns about rights as a research participant should be directed to the Chair of the Institutional Review Board, The University of Southern Mississippi, 118 College Drive #5125, Hattiesburg, MS 39406-0001, 601-266-5997.

Any questions about this research project should be directed to the Principal Investigator using the contact information provided above.

CONSENT TO PARTICIPATE IN RESEARCH

Participant's Name: ______________________

I hereby consent to participate in this research project. All research procedures and their purpose were explained to me, and I had the opportunity to ask questions about both the procedures and their purpose. I received information about all expected benefits, risks, inconveniences, or discomforts, and I had the opportunity to ask questions about them. I understand my participation in the project is completely voluntary and that I may withdraw from the project at any time without penalty, prejudice, or loss of benefits. I understand the extent to which my personal information will be kept confidential. As the research proceeds, I understand that any new information that emerges and that might be relevant to my willingness to continue my participation will be provided to me.

______________________	______________________
Research Participant	**Person Explaining the Study**
______________________	______________________
Date	**Date**

APPENDIX E – Video Links

http://youtu.be/9AHqFKc3mKY	Fall of the Roman Empire
http://youtu.be/ojSkGvxFi4M	Golden Age of Athens, Pericles and Greek Culture
http://youtu.be/SGSLyp8mmMc	Spread of Islam
http://youtu.be/0t4MF9ZoppM	Ancient Egypt
http://youtu.be/Um92GZLCQ_Q	French Revolution part 2
http://youtu.be/T8O4AcTyjHc	Allende and Pinochet in Chile
http://youtu.be/Y33LnxG2L80	Arian Controversy and the Council of Nicaea
http://youtu.be/WhTpJxlJi2I	Augustus becomes first Emperor of Rome
http://youtu.be/OzyH-1p9nAg	hinduism introduction core ideas of brahman atman samsara and moksha
http://youtu.be/QCkn5bu8GgM	Napoleon and the Wars of the First and Second Coalitions
http://youtu.be/p3pYuY4buIk	Initial Rise of Hitler and the Nazis
http://youtu.be/hNpcQEGw3S4	Feudal System during the Middle Ages
http://youtu.be/iPQ6GB822x4	Ottoman, Safavid and Mughal Empires
http://youtu.be/j7N-XPi5Z0	Confucius and the Hundred Schools of Thought
http://youtu.be/mi9sMazNPxM	Hittite Empire and Battle of Kadesh
http://youtu.be/K5XKjk0-hCo	Indus River Valley Civilizations
http://youtu.be/zc_p7Mw1A7U	Golden age of Islam
http://youtu.be/pJQr77Vzwyk	Ides of March and civil war
http://youtu.be/XHVty6_XTJY	Socrates Plato Aristotle
http://youtu.be/g8sxNa-E-H0	Punic Wars between Rome and Carthage
http://youtu.be/Sa5eqaYwQ2Q	Theodor Herzl and the birth of political Zionism
http://youtu.be/xFBK9534NI8	Alexander the Great takes power
http://youtu.be/B_P48TakY3Y	Closing Stages in World War I
http://youtu.be/XmkbAduMD_E	Blockades, U-boats, Lusitania
http://youtu.be/EqEEndY0sT8	Bay of Pigs Invasion
http://youtu.be/X3bqQI7-sCg	Cyrus the Great establishes the Achaemenid Empire
http://youtu.be/a9QtIfPIQl4	Axis Momentum Accelerates in WW2
http://youtu.be/eIfQ4GfSz3U	Overview of Chinese history 1911 - 1949
http://youtu.be/9e9GWdT2pEQ	Sykes-Picot Agreement and the Balfour Declaration
http://youtu.be/MEGyRgYJKEY	Vietnam War
http://youtu.be/VO40SpSBjbc	Korean War
http://youtu.be/CH6FQhlZn6k	Cuban Missile Crisis
http://youtu.be/F_ySQvjtAxQ	Napoleon forced to abdicate
http://youtu.be/Qz5zFzvbib4	French Invasion of Russia
http://youtu.be/O3HxPDH-s7w	Haitian Revolution
http://youtu.be/ALJGz4r_VF0	Napolean and Fourth Coalition
	Napoleon and Peninsular Campaigns

APPENDIX F – Data Sheet

Date

Participant

Session

Condition Baseline Demand Praise

	1.1	1.2	1.3	1.4	1.5	1.6	2.1	2.2	2.3	2.4	2.5	2.6
Engage												
Praise												
Demand												
	3.1	3.2	3.3	3.4	3.5	3.6	4.1	4.2	4.3	4.4	4.5	4.6
Engage												
Praise												
Demand												
	5.1	5.2	5.3	5.4	5.5	5.6	6.1	6.2	6.3	6.4	6.5	6.6
Engage												
Praise												
Demand												
	7.1	7.2	7.3	7.4	7.5	7.6	8.1	8.2	8.3	8.4	8.5	8.6
Engage												
Praise												
Demand												
	9.1	9.2	9.3	9.4	9.5	9.6	10.1	10.2	10.3	10.4	10.5	10.6
Engage												
Praise												
Demand												
	11.1	11.2	11.3	11.4	11.5	11.6	12.1	12.2	12.3	12.4	12.5	12.6
Engage												
Praise												
Demand												
	13.1	13.2	13.3	13.4	13.5	13.6	14.1	14.2	14.3	14.4	14.5	14.6
Engage												
Praise												
Demand												
	15.1	15.2	15.3	15.4	15.5	15.6						
Engage												
Praise												
Demand												

APPENDIX G – Baseline Procedural integrity

Baseline	
Date: _____ Participant: ____ Trial: _____ Obs: _________	
Circle "Y" for each step each time the implementer(s) completed the step correctly. Circle "N" for each time an implementer missed or incorrectly completed a step Integrity = Yes/(Yes+No) * 100	
1. The researcher told participant to keep their speakers on, unmute themselves, to take no notes, stay at the computer, minimize head movements, and to read all directions on the screen.	Y N
2. Researcher confirmed that their video is disabled	Y N
3. The participant read all instructions out loud	Y N
4. Researcher engaged in no verbal communication during the video	Y N
5. Distractions were present, including iPad turned on, building blocks and magnetic letters, and puzzle ball.	
Total Percent Correct Implementation	%

APPENDIX H – Demand Procedural integrity

Demand	
Date: ______ Participant: _____ Trial: ______ Obs: ________	
Circle "Y" for each step each time the implementer(s) completed the step correctly. Circle "N" for each time an implementer missed or incorrectly completed a step	
1. The researcher told participant to keep their speakers on, unmute themselves, to take no notes, stay at the computer, minimize head movements, and to read all directions on the screen.	Y N
2. Researcher confirmed that their video is disabled	Y N
3. The participant read all instructions out loud	Y N
4. After the participant starts the video, the researcher delivers first demand.	Y N
5. During the video, researcher only engaged in verbal communication at the designated intervals (every 2 mins)	Y N
6. All verbalizations were demands related to the video (e.g., "Watch the video," "Look at the screen.")	Y N
7. Researcher responded with demand statement within an interval of the demand indicator (Square B) appearing	Y N
8. Distractions were present, including iPad turned on, building blocks and magnetic letters, and puzzle ball.	Y N
Total Percent Correct Implementation	%

APPENDIX I – Praise Procedural integrity

Praise	
Date: ______ Participant: _____ Trial: ______ Obs: _________	
Circle "Y" for each step each time the implementer(s) completed the step correctly. Circle "N" for each time an implementer missed or incorrectly completed a step	
1. The researcher told participant to keep their speakers on, unmute themselves, to take no notes, stay at the computer, minimize head movements, and to read all directions on the screen.	Y N
2. Researcher confirmed that their video is disabled	Y N
3. The participant read all instructions out loud	Y N
4. After the participant starts the video, the researcher delivers first praise statement.	Y N
5. Researcher only engaged in verbal communication during trials at the designated intervals (every 2 mins)	Y N
6. All verbalizations consisted of a verbal statement that signified approval (e.g., "Nice job watching the video," "Awesome attending.")	Y N
7. Researcher responded with praise statement within an interval of the praise indicator (Square A) appearing	Y N
8. Distractions were present, including iPad turned on, building blocks and magnetic letters, and puzzle ball	Y N
Total Percent Correct Implementation	%

REFERENCES

Acker, M. M., & O'Leary, S. G. (1987). Effects of reprimands and praise on appropriate behavior in the classroom. *Journal of abnormal child psychology*, *15*(4), 549–557.

Alberto & Troutman (2013). *Applied Behavior Analysis for Teachers (9th Edition).*

Allday, R. A., Hinkson-Lee, K., Hudson, T., Neilsen-Gatti, S., Kleinke, A., & Russel, C. S. (2012). Training general educators to increase behavior-specific praise: Effects on students with EBD. Behavioral Disorders, 37, 87–98.

Allen, K. E., Hart, B., Buell, J. S., Harris, F. R., & Wolf, M. M. (1964). *Effects of Social Reinforcement on Isolate Behavior of a Nursery School Child. Child Development*, *35*(2), 511.

Anderson, A. R., Christenson, S. L., Sinclair, M. F., & Lehr, C. A. (2004). Check & connect: The importance of relationships for promoting engagement with school. *Journal of School Psychology, 42*(2), 95–113.

Atteberry, A., & McEachin, A. (2016). School's out: Summer learning loss across grade levels and school contexts in the United States today. In Alexander, K., Pitcock, S., & Boulay, M. (Eds). *Summer learning and summer learning loss*, pp35-54. New York: Teachers College Press.

Austin, J. L., & Soeda, J. M. (2008*). Fixed-Time Teacher Attention to Decrease Off-Task Behaviors of Typically Developing Third Graders. Journal of Applied Behavior Analysis, 41*(2), 279–283.

Baddeley, A., Lewis, V. J., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113*, 518-540.

Banda, D. R., & Sokolosky, S. (2012). Effectiveness of Noncontingent Attention to Decrease Attention-Maintained Disruptive Behaviors in the General Education Classroom. *Child & Family Behavior Therapy, 34*(2), 130–140.

Barbour, M. (2018). A History of K-12 Distance, Online, and Blended Learning Worldwide, in K. Kennedy & R. Ferdig (Eds.), *Handbook of Research on K-12 Online and Blended Learning*, 21-32. Pittsburg: Carnegie Mellon University/ETC Press.

Barbour, M. K. (2010). Researching K-12 online learning: What do we know and what should we examine? *Distance Learning, 7*(2), 7–12.

Barton, E. E., Lloyd, B. P., Spriggs, A. D., & Gast, D. L. (2018). Visual analysis of graphic data. In J. R. Ledford & D. L. Gast (Eds.), *Single case research in behavioral sciences* (3rd ed., pp. 179–214). New York: Routledge.

Bates, A.W. (2001). Beyond button-pushing: using technology to improve learning, in R. Epper & A.W. Bates (Eds.) *Teaching Faculty How to Use Technology: best practices from leading institutions*, 141–152. Westport: American Council on Education/Oryx Press.

Berge, Z. L., & Clark, T. (2005). *Virtual schools: Planning for success.* New York, NY: Teachers College Press.

Bernard, R., Borokhovski, E., and Tamim, R. (2019). The State of Research of Distance, Online, and Blended Learning, in M. Moore & R. W. Diehl (Eds.), *Handbook of Distance Education*, 92-104. New York: Routledge.

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis*, *1*, 175–191.

Birch, S., & Ladd, G. (1997). The teacher–child relationship and children's early school adjustment. *Journal of School Psychology, 35*, 61–79.

Birnbrauer, J. ., Wolf, M. ., Kidder, J., & Tague, C. E. (1965). *Classroom behavior of retarded pupils with token reinforcement. Journal of Experimental Child Psychology, 2(2), 219–235.*

Black, M. K. (2013). The landscape of K-12 online learning: Examining what is known. In M. G. Moore (Eds.), *Handbook of distance education* (3rd ed.) (pp. 574-593). New York: Routledge.

Blatchford, P., Edmonds, S., & Martin, C. (2003). Class size, pupil attentiveness and peer relations. *British Journal of Educational Psychology*, *73*, 15-36.

Blaze, J. T., Olmi, D., Mercer, S. H., Dufrene, B. A., & Tingstrom, D. H. (2014). Loud versus quiet praise: A direct behavioral comparison in secondary classrooms. *Journal of School Psychology, 52*(4), 349–360.

Bos, N., Groeneveld, C., van Bruggen, J., & Brand-Gruwel, S. (2015). The use of recorded lectures in education and the impact on lecture attendance and exam performance. *British Journal of Educational Technology, 47*(5), 906–917.

Bradbury, N. A. (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education, 40*(4), 509–513.

Branch, M. N., & Malagodi, E. F. (1980). Where have all the behaviorists gone? *The Behavior Analyst, 3,* 31-38.

Broden, M., Bruce, C., Mitchell, M. A., Carter, V., & Hall, R. V. (1970). Effects of teacher attention on attending behavior of two boys at adjacent desks. *Journal of Applied Behavior Analysis, 3,* 205–211. doi:10.1901/jaba.1970.3-205

Brodersen, R. M., & Melluzzo, D. (2017). *Summary of research on online and blended learning programs that offer differentiated learning options* (REL 2017–228). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. Retrieved from http://ies.ed.gov/ncee/edlabs.

Buswell, G. T. (1935). How people look at pictures. Chicago: University of Chicago Press.Caldarella, P., Larsen, R. A. A., Williams, L., Downs, K. R., Wills, H. P., & Wehby, J. H. (2020). Effects of teachers' praise-to-reprimand ratios on elementary students' on-task behavior. *Educational Psychology, 40* (10), 1306–1322.

Caldarella, P., Larsen, R. A., Williams, L., Downs, K., Wills, H., & Wehby, J. (2020a): Effects of teachers' praise-to-reprimand ratios on elementary students' on-task behaviour, *Educational Psychology*., 1-17.

Caldarella, P., Larsen, R. A., Williams, L., Wills, H. P., & Wehby, J. H. (2020b). "Stop Doing That!": Effects of Teacher Reprimands on Student Disruptive Behavior and Engagement. *Journal of Positive Behavior Interventions, 109830072093510.*

Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). *Effectiveness of reading and mathematics software products: Findings from two student cohorts* (NCEE No. 2009–4041). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Carini, R.M., Kuh, G.D. and Klein, S.P. (2006) Student engagement and student learning: Testing the linkages. *Research in Higher Education, 47,* 1-24.

Carone, D. P. (1975). The effect of positive verbal feedback on females' intrinsic motivation. Unpublished master's thesis, University of Bridgeport, CT.

Carr, E. G. (1977). The motivation of self-injurious behavior: A review of some hypotheses. *Psychological Bulletin, 84,* 800-816.

Charlesworth, W. R., & Spiker, D. (1975). An ethological approach to observations in learning settings. In R. A. Weinberg & F. H. Wood (Eds.), *Observation of pupils and teachers in mainstream and special education settings: Alternative strategies.* The Council for Exceptional Children: Reston, VA, pp. 161–170.

Chelazzi, L., E to inova, J., Calletti, R., Lo Gerfo, E., Sani, I., Della Libera, C., & Santandrea, E. (2014). Altering Spatial Priority Maps via Reward-Based Learning. *Journal of Neuroscience, 34*(25), 8594–8604.

Cobb, J. A (1972). The relationship of discrete classroom behaviors to fourth-grade academic achievement. *Journal of Educational Psychology, 63,* 74-80.

Cobb, J. A. (1972). Relationship of discrete classroom behaviors to fourth-grade academic achievement. *Journal of Educational Psychology, 63*(1), 74–80.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2019). *Applied Behavior Analysis (3rd Edition)*. Hoboken, NJ: Pearson Education.

Couldry, N. (2012). *Media, Society, World: Social Theory and Digital Media Practice*. London: Polity Press. p. 2.

Crawford, C., Barker, J., & Seyam, A. (2014). *The promising role of hybrid learning in community colleges: Looking towards the future*. Littleton, CO: The Clute Institute.

Davis, B. G. *Tools for Teaching.* San Francisco, CA: Jossey-Bass, 1993.

Deci, E. L. (1972). Intrinsic motivation, extrinsic reinforcement, and inequity. *Journal of Personality and Social Psychology, 22,* 113–120.

Deci, E. L., Cascio, W. E., & Krusell, J. (1975). Cognitive evaluation theory and some comments on the Calder and Staw critique. *Journal of Personality and Social Psychology, 31,* 81–85.

Digital Learning Collaborative, (2020). *Snapshot 2020: A review of online, blended, and digital learning*. https://static1.squarespace.com/static/5a98496696d4556b01f86662/t/5e61341d879e630db4481a01/1583428708513/DLC-KP-Snapshot2020.pdf

D'Mello, S., Dieterle, E., & Duckworth, A (2017). Advanced, Analytic, Automated (AAA) Measurement of Engagement During Learning. *Educational Psychology 52*(2), 104-123.

D'Mello, S., Kopp, K., Bixler, R. E., & Bosch, N. (2016). *Attending to Attention. Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16*. doi:10.1145/2851581.2892329

Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review, 8*(2).

Downs, K. R., Caldarella, P., Larsen, R. A. A., Charlton, C. T., Wills, H. P., Kamps, D. M., & Wehby, J. H. (2019). Teacher praise and reprimands: The differential response of students at risk for emotional and behavioral disorders. *Journal of Positive Behavior Interventions*, *21,* 135–147

Dufrene, B. A., Lestremau, L., & Zoder-Martell, K. (2014). Direct behavioral consultation: Effects on teachers' praise and student disruptive behavior. *Psychology in the Schools*, *51,* 567-580.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N. Campuzano, L., et al. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort* (NCEE 2007–4005). Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Epling, W. F., & Pierce, W. D. (1983). Applied behavior analysis: New directions from the laboratory. *The Behavior Analyst, 6,* 27-37

Ervin, R., Ehrhardt, K., & Poling, A. (2001) Functional Assessment: Old Wine in New Bottles. *School Psychology Review, 30*(2), 173-179.

Evans, C. (2008). The effectiveness of m-learning in the form of podcast revision lectures in higher education. *Computers & Education*, *50*(2), 491–498.

Faber, M., Bixler, R. & D'Mello, S.K. An automated behavioral measure of mind wandering during computerized reading. *Behav Res* **50,** 134–150 (2018).

Finn, J. D. (1993). *School engagement and students at risk.* Washington, DC: National Center for Education Statistics.

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology, 82,* 221–234.

Finn, J. D., & Voelkl, K. E. (1993). School characteristics related to school engagement. *Journal of Negro Education, 62,* 249–268.

Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and inattentive withdrawn behavior and achievement among fourth graders. *Elementary School Journal, 95,* 421–454.

Floress, M. T., Jenkins, L. N., Reinke, W. M., & McKown, L. (2018). General education teachers' natural rates of praise: A preliminary investigation. *Behavioral Disorders*, *43*(4), 411–422.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109.

Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *Handbook of blended learning: Global perspectives, local designs* (pp. 3–21). San Francisco, CA: Pfeiffer.

Griffin, D. K., Mitchell, D. & Thompson, S. J. (2009). Podcasting by 63hildren63zing PowerPoint and voice: what are the pedagogical benefits? *Computers & Education, 53*(2), 532–539.

Gunter, P. L.,Venn, M. L., Patrick, J., Miller, K. A., & Kelly, L. (2003). Efficacy of using momentary time samples to determine on-task behavior of students with emotional/behavioral disorders. *Education and Treatment of Children, 26,* 400–412.

Guo, P. J., Kim, J., & Rubin, R. (2014). *How video production affects student engagement.* Proceedings of the First ACM Conference on Learning @ Scale Conference – L@S '14.

Gupta, A. & Saks, N. S. (2013). Exploring medical student decisions regarding attending live lectures and using recorded lectures. *Medical Teacher, 35*(9), 767–771.

Gysbers, V., Johnston, J., Hancock, D. & Denyer, G. (2011). Why do students still bother coming to lectures, when everything is available online? *International Journal of Innovation in Science and Mathematics Education, 19*(2), 20–36.

Hall, R. V. and Broden, M. (1967) Behavior changes in brain-injured children through social reinforcement. *Journal of Experimental Child Psychology, 5,* 463-479

Hall, R. V., Lund, D., & Jackson, D. (1968). Effects of Teacher Attention on Study Behavior . *Journal of Applied Behavior Analysis, 1*(1), 1–12.

Hansen, D. W. & Ji, Q. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions of Pattern Analysis and Machine Intelligence, 32*(3):478–500, Mar. 2010.

Harris, F., Wolf, M., & Baer, D. (1964). Effects of Adult Social Reinforcement on Child Behavior. *Young Children, 20*(1), 8-17.

Hart, B. M., Allen, K. E., Buell, J. S., Harris, F. R., & Wolf, M. M. (1964). Effects of social reinforcement on operant crying. *Journal of Experimental Child Psychology, 1*(2), 145–153.

Hasazi, J. E., & Hasazi, S. E. (1972). Effects of teacher attention on digit-reversal behavior in an elementary school child. *Journal of Applied Behavior Analysis, 5*, 157–162.

Hay, W. M., Hay, L. R., & Nelson, R. O. (1977). Direct and collateral changes in on-task and academic behavior resulting from on-task versus academic contingencies. *Behavior Therapy, 8*(3), 431–441.

Haydon T, Musti-Rao S, Kennedy A, Murphy M, Hunter W, Boone J. (2016). Using Teacher Praise with Middle and High School Students. *Beyond Behavior, 29*(2). 108-115.

Haydon, T., & Musti-Rao, S. (2011). Effective use of behavior-specific praise: A middle school case study. *Beyond Behavior, 20*(2), 31–39.

Hayward, D. A., Pereira, E. J., Otto, A. R., & Ristic, J. (2018). Smile! Social reward drives attention. *Journal of Experimental Psychology: Human Perception and Performance, 44*(2), 206–214.

Hecht, L. (1978). Measuring Student Behavior during Group Instruction. *The Journal of Educational Research, 71*(5), 283–290.

Heddy, B. C., Sinatra, G. M., Seli, H., & Mukhopadhyay, A. (2014, April). *Transformative experience as a facilitator of interest development and transfer in a college success course for at-risk students*. Paper presented at the American Educational Research Association, Philadelphia, PA.

Hollingshead, A., Kroeger, S. D., Altus, J., & Trytten, J. B. (2016). A case study of positive behavior supports-based interventions in a seventh-grade urban classroom. *Preventing School Failure, 60*(4), 1–8

Homer., R. D. (1980). The effects of an environmental “enrichment” program on the behavior of institutionalized profoundly retarded children. *Journal of Applied Behavior Analysis, 13*, 473-491.

Hong, J., Pi, Z., & Yang, J. (2018). Learning declarative and procedural knowledge via video lectures: Cognitive load and learning effectiveness. *Innovations in Education and Teaching International*, *55*(1), 74–81.

Horner, R., Swaminathan, H., Sugai, G., & Smolkowski, K. (2012). Expanding analysis and use of single-case research. *Education and Treatment of Children, 35,* 269–290.

Houghton, S., Wheldall, K., Jukes, R., & Sharpe, A. (1990). The Effects of Limited Private Reprimands And Increased Private Praise on Classroom Behaviour in Four British Secondary School Classes. *British Journal of Educational Psychology*, *60*(3), *255–265.*

Hu, M., Arnesen, K., Barbour, M.K. & Leary, H. (2019). A Newcomer's Lens: A Look at K-12 Online and Blended Learning in the Journal of Online Learning Research. Journal of Online Learning Research, 5(2), 123-144. Waynesville, NC USA: Association for the Advancement of Computing in Education (AACE). Retrieved September 10, 2021 from https://www.learntechlib.org/primary/p/195231/

Huang, J., White, R., & Buscher, G. User see, user point: Gaze and cursor alignment in web
search. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12, pages 1341–1350, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1015-4. doi:10.1145/2207676.2208591.

Hutt, S., Mills, C., White, S., Donnelly, P., & D'Mello, S. K. (2016). The eyes have it: gaze-based detection of mind wandering during learning with an intelligent tutoring system. *The 9th International Conference on Educational Data Mining* (Raleigh, NC, USA), 86–93.

Hutt, S., Krasich, K., Mills, C., Bosch, N., White, S., Brockmole, J. R., & D'Mello, S. K. (2019). Automated gaze-based mind wandering detection during computerized learning in classrooms. *User Model User-Adap Inter 29*, 821–867.

Hutt, S. "Scaling up: Moving automated gaze-based engagement detection out of the lab" (2020). Available from ProQuest Dissertations & Theses Global. at https://www.proquest.com/docview/2446610069/fulltextPDF/D36ECBD1F6DB4845PQ/

Inhoff, A.W., & Radach, R. (1998). *Definition and computation of oculomotor measures in the study of cognitive processes*. In G. Underwood (Ed.), Eye guidance in reading and scene perception (pp. 29-54). Oxford, England: Elsevier.

International Association for K-12 Online Learning. (2011). *National standards for quality on-line teaching.* Vienna, VA: Author.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982). *Toward a functional analysis of self-injury. Analysis and Intervention in Developmental Disabilities, 2*(1), 3–20.

Johnstone A.H. and Percival, F. Attention breaks in lectures. *Educ Chem, 13*: 49 –50, 1976.

Kaufman, D. (1989) Third generation course design in distance education, in Sweet, R. (Ed.) *Post-Secondary Distance Education in Canada: Policies, Practices and Priorities*. Athabasca: Athabasca University/Canadian Society for Studies in Education.

Kebritchi, M., Lipschuetz, A., & Santiague, L. (2017). Issues and Challenges for Teaching Successful Online Courses in Higher Education. *Journal of Educational Technology Systems, 46*(1), 4–29.

Khan Academy. Accessed February 14, 2021. https://www.khanacademy.org.

Khorrami, P., Vuong Le, Hart, J. C., & Huang, T. S. (2014). A system for monitoring the engagement of remote online students using eye gaze estimation. *2014 IEEE International Conference on Multimedia and Expo Workshops* (2014).

Kim, J. S., Capotosto, L., Hartry, A., & Fitzgerald, R. (2011). Can a mixed-method literacy intervention improve the reading achievement of low-performing elementary school students in an after-school program? Results from a randomized controlled trial of READ 180 Enterprise. *Educational Evaluation and Policy Analysis, 33*(2), 183–201.

Kirby, F. D., & Shields, F. (1972). Modification of arithmetic response rate and attending behavior in a seventh-grade student. *Journal of Applied Behavior Analysis, 5,* 79–84.

Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014). *Showing face in video instruction.* Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems – CHI '14.

Kodak, T., Northup, J., & Kelley, M. E. (2007). An evaluation of the types of attention that maintain problem behavior. *Journal of Applied Behavior Analysis, 40*, 167–171.

Koestner, R., Zuckerman, M., & Koestner, J. (1987). Praise, involvement, and intrinsic motivation. *Journal of Personality and Social Psychology, 53*(2), 383–390. https ://doi.org/10.1037/0022-3514.53.2.383.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Krause, K., & Coates, H. (2008). Students' engagement in first-year university. *Assessment and Evaluation in Higher Education, 33*(5), 493–505.

Lahaderne, H. (1968). Attitudinal and intellectual correlates of attention: A study of four sixth-grade classrooms. *Journal of Educational Psychology, 59*(5), 320–324.

Lane, J. D., & Gast, D. L. (2014). Visual analysis in single case experimental design studies: Brief review and guidelines. *Neuropsychological Rehabilitation: An International Journal, 24*:3-4, 445-463.

Lessard, L.M., Grossman, A., & Syme, M.L. (2015). Effects of Gender and Type of Praise on Task Performance Among Undergraduates. *Psi Chi Journal of Psychological Research, 20*, 11-17.

Li, C., & Irby, B. (2008). An Overview of online education: Attractiveness, benefits, challenges, concerns, and recommendations. *College Student Journal, Part A, 42*, 449–458.

Lindquist, S. I., & McLean, J. P. (2011). Daydreaming and its correlates in an educational environment. *Learning and Individual Differences, 21*(2), 158–167.

Lloyd, D. H. (1968, October). A concept of improvement of learning response in the taught lesson. *Visual Education,* 23–25.

Lowe, T. O., & McLaughlin, E. C. (1974). The use of verbal reinforcement by paraprofessionals in the treatment of underachieving elementary school students. *Journal of the Student Personnel Association for Teacher Education, 12*(3), 95.

Lundgren, E., Rocha, T., Rocha, Z., Carvalho, P., & Bello, M. tracking.js: A modern approach for Computer Vision on the web. http://trackingjs.com, 2014.

Luyt, I. (2013). Bridging spaces: Cross-cultural perspectives on promoting positive online learning experiences. *Journal of Educational Technology Systems, 42*, 3–20.

Lyons, J. F. (2004). Teaching U.S. history online: Problems and prospects. *The History Teacher, 37,* 447–456.

Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject researches: Percentage of data points exceeding the median. *Behavior Modification, 30,* 598–617.

Madsen, C. H., Jr., Becker, W. C., & Thomas, D. R. (1968). Rules, praise, and ignoring: Elements of elementary classroom control. *Journal of Applied Behavior Analysis, 1*, 139– 150. doi:10.1901/jaba.1968.1-139

Martens, B. K., Hiralall, A. S., & Bradley, T. A. (1997). *A note to teacher: Improving student behavior through goal setting and feedback. School Psychology Quarterly, 12*(1), 33–41.

Martin, R., Piedmont-Palladino, S., Sturlaugson, B., Penner, B., Harriss, H., Rodenbeck, J., Isenstadt, S., Livia-Brand, A., Ansari, I., Frichot, H., & Milligan, B. (2020). “Field Notes on Pandemic Teaching: 1,” *Places Journal*. Accessed 17 Jan 2021.

Mastropieri, M. A., & Scruggs, T. E. (1985-86). Early intervention for socially withdrawn 68hildren. *The Journal of Special Education, 19*, 429-441.

Matheson, A. S., & Shriver, M. D. (2005) Training Teachers to Give Effective Commands: Effects on Student Compliance and Academic Behaviors. *School Psychology Review, 34*(2), 202-219

McKeachie, W.J. *Teaching Tips: Strategies, Research and Theory for College and University Teachers.* Lexington, MA: Heath, 1986.

Means, B., Toyama, Y., Murphy, R., & Bakia, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record, 115*(3), 1–47.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies Center for Technology in Learning*.

Meyers, C. E., Attwell, A. A., & Orpet, R. E. (1968). Prediction of fifth grade achievement from kindergarten test and rating data. *Educational and Psychological Measurement, 28*(2), 457–463.

Michael, J. L. (1980). Flight from behavior analysis. *The Behavior Analyst*, *3*, 1-24.

Moore, T. C., Maggin, D. M., Thompson, K. M., Gordon, J. R., Daniels, S., & Lang, L. E. (2018). Evidence Review for Teacher Praise to Improve Students’ Classroom Behavior. *Journal of Positive Behavior Interventions, 109830071876665.*

Morgan, P., & Ritter, S. (2002). *An experimental study of the effects of Cognitive Tutor® Algebra I on student knowledge and attitude*. Pittsburgh, PA: Carnegie Learning.

Morris, N. P., Swinnerton, B., & Coop, T. (2019). Lecture recordings to support learning: A contested space between students and teachers. *Computers & Education*, *140,* 103604.

O’Handley, R. D., “A Direct Comparison of Different Schedules of Praise in Secondary Classrooms” (2016). Dissertations. 434. https://aquila.usm.edu/dissertations/434

O’Handley, R. D., Dufrene, B. A., & Whipple, H. (2018). Tactile prompting and weekly performance feedback for increasing teachers’ behavior-specific praise. *Journal of Behavioral Education, 27*(3), 324–342.

O'Leary, K. D., & O'Leary, S. G. (Eds.). (1977). *Classroom management: The successful use of behavior modification* (2nd ed.). New York: Pergamon Press.

Pacitto, G. O. (2019). *An Evaluation of the Effectiveness of Content and Quality of Praise as a Reinforcer for Skill Acquisition in Children with and without Developmental Delays* (Unpublished Doctoral Dissertation). The Chicago School of Professional Psychology.

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at Scale. *Educational Evaluation and Policy Analysis, 36*(2), 127–144.

Pane, J. F., McCaffrey, D. F., Slaughter, M. E., Steele, J. L., & Ikemoto, G. S. (2010). An experiment to evaluate the efficacy of Cognitive Tutor Geometry. *Journal of Research on Educational Effectiveness, 3*(3), 254–281.

Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., & Hays, J. (2016). *WebGazer*: Scalable Webcam Eye Tracking Using User Interactions. *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)*, 3839–3845.

Papoutsaki, A. "Democratizing Eye Tracking" (2017). *Computer Science Theses and Dissertations*. Brown Digital Repository. Brown University Library.

Papoutsaki, A., Laskey, J., & Huang, J. (2017). *SearchGazer. Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval - CHIIR '17.* doi:10.1145/3020165.3020170

Parker, R. I., Hagan-Burke, S., &Vannest, K. (2007). Percentage of All Non-Overlapping Data (PAND): An Alternative to PND. *The Journal of Special Education, 40,* 194-204.

Patterson, G. R. (1966). An application of conditioning techniques to the control of a hyperactive child, Ullman, L. P., and Krasner, L. (Eds.), *Case studies in behavior modification*, New York: Holt, Rinehart and Winston, Inc. Pp. 370-375.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods.*

Peterson, P. L., Swing, S. R., Stark, K. D., & Waas, G. A. (1984). Students' Cognitions and Time on Task During Mathematics Instruction. *American Educational Research Journal, 21*(3), *487–515.*

Polick, A. S., Carr, J. E., & Hanney, N. M. (2012). A comparison of general and descriptive praise in teaching intraverbal behavior to children with autism. *Journal of Applied Behavior Analysis*, *45*(3), *593–599.*

Poling, A., Picker, M., Grossett, D., Hall-Johnson, E., & Holbrook, M. (1981). The schism between experimental and applied behavior analysis: Is it real and who cares*? The Behavior Analyst, 4*, 93-102.

Powell, J., Martindale, B., & Kulp, S. (1975). An evaluation of time-sample measures of behavior. *Journal of Applied Behavior Analysis, 8*, 463–469.

Powell, J., Martindale, B., Kulp, S., Martindale, A., & Bauman, R. (1977). Taking a closer look: Time sampling and measurement error. *Journal of Applied Behavior Analysis, 10*, 325–332.

Queen, B., and Lewis, L. (2011). Distance Education Courses for Public Elementary and Secondary School Students: 2009–10 (NCES 2012-008). U.S. Department of Education, National Center for Education Statistics. Washington, DC: Government Printing Office.

Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. Psychological bulletin. 124, 3 (Nov. 1998), 372–422.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125–157. doi:10.1037/0033-295X.105.1.125

Riley, J. L., McKevitt, B. C., Shriver, M. D., & Allen, K. D. (2011). Increasing On-Task Behavior Using Teacher Attention Delivered on a Fixed-Time Schedule. *Journal of Behavioral Education, 20*(3), 149–162.

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). *Everyday Attention: Variation in Mind Wandering and Memory in a Lecture. Applied Cognitive Psychology, 26*(2), 234–242.

Ritter, S., Kulikowich, J., Lei, P-W., McGuire, C. L., & Morgan, P. (2007). What evidence matters? A randomized field trial of Cognitive Tutor Algebra I. In T. Hirashima, H. U. Hoppe, & S. Shwu-Cing Young (Eds.), *Supporting learning flow through integrative technologies.* Amsterdam, Netherlands: IOS Press; pp. 13–20.

Romero-Ivanova, C., Shaughnessy, M., Otto, L., Taylor, E., and Watson. E. (2020). Digital Practices & Applications in a Covid-19 Culture. *Higher Education Studies 10*(80), 80-87.

Rooney, J. E. (2003). Blending learning opportunities to enhance educational programming and meetings. *Association Management, 55*(5), 26–32.

Rosen, Y., & Beck-Hill, D. (2012). Intertwining digital content and a one-to-one laptop environment in teaching and learning: Lessons from the Time To Know program. *Journal of Research on Technology in Education, 44*(3), 225–241.

Royer, D. J., Lane, K. L., Dunlap, K. D., & Ennis, R. P. (2019). A Systematic Review of Teacher-Delivered Behavior-Specific Praise on K–12 Student Performance. *Remedial and Special Education*, *40*(2), 112–128.

Rozsa, Andrew J. III, "USING PRAISE TO INCREASE VISUAL ATTENDING IN AN ASYNCHRONOUS ONLINE LEARNING ENVIRONMENT: AN EYE TRACKING STUDY" (2021). *Master's Theses*. 856. https://aquila.usm.edu/masters_theses/856

Rubow, C. C., Noel, C. R., & Wehby, J. H. (2019). Effects of Noncontingent Attention on the Behavior of Students with Emotional/Behavioral Disorders and Staff in Alternative Settings. *Education and Treatment of Children, 42*(2), 201–223.

Ryan, S., Kaufman, J., Greenhouse, J., She, R., & Shi, J. (2015). The effectiveness of blended online learning courses at the community college level. *Community College Journal of Research and Practice, 40*(4), 285–298.

Saudargas, R. A., & Zanolli, K. (1990). Momentary time sampling as an estimate of percentage time: A field validation. *Journal of Applied Behavior Analysis, 23,* 533–537.

Schwirzke, K., Vashaw, L., & Watson, J. (2018). A history of K-12 online and blended instruction in the United States, in K. Kennedy & R. Ferdig (Eds.), *Handbook of Research on K-12 Online and Blended Learning,* 7-20. Pittsburg: Carnegie Mellon University/ETC Press.

Scruggs, T. E. & Mastropieri, M. A. (1998). Summarizing single subject research: Issues and applications. *Behavior Modification, 22*, 221-242.

Scruggs, T. E., Matropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24-33.

Seaman, J. E., Elaine Allen, I., & Seaman, J. (2018). *Grade Increase: Tracking Distance Education in the United States*. Retrieved January 13, 2021.

Sharp, C., Nelson. J., Lucas. M., Julius, J., McCrone. T. and Sims, D. (2020). *Schools' responses to Covid-19: The challenges facing schools and pupils in September 2020*. Slough: NFER.

Simpson, M. J. A., & Simpson, A. E. (1977). One-zero and scan method for sampling behavior. *Animal Behavior, 25,* 726–731.

Sinatra, G., Heddy, B., & Lombardi, D. (2015). The Challenges of Defining and Measuring Student Engagement in Science. *Educational Psychologist, 50*(1), 1-13.

Singh, V., & Thurman, A. (2019). *How* Many Ways Can We Define Online Learning? A Systematic Literature Review of Definitions of Online Learning (1988-2018). *American Journal of `Distance Education, 33*(4), 289–306.

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology, 59*(3), 623–664.

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effect of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85,* 571–581.

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience, 20,* 458−469.

Sutherland, K. S., & Wehby, J. H. (2001). The effect of self-evaluation on teaching behavior in classrooms for students with emotional and behavioral disorders. *The Journal of Special Education, 35,* 161–171.

Sutherland, K. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8*(1), 2–8.

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). *Interpolated memory tests reduce mind wandering and improve learning of online lectures*. Proceedings of the National Academy of Sciences.

Szpunar, K. K., Moulton, S. T., & Schacter, D. L. (2013). Mind wandering and education: from the classroom to online learning. *Frontiers in Psychology, 4.*

Taie, S., and Goldring, R., (2017). National Teacher and Principal Survey First Look (NCES 2017-071). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Taie, S., and Goldring, R. (2019). Characteristics of Public and Private Elementary and Secondary Schools in the United States: Results From the 2017–18 National Teacher and Principal Survey First Look (NCES 2019-140). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Taie, S., and Goldring, R. (2020). Characteristics of Public and Private Elementary and Secondary School Teachers in the United States: Results From the 2017–18 National Teacher and Principal Survey First Look (NCES 2020- 142rev). U.S. Department of Education. Washington, DC: National Center for Education Statistics

Teerlink, E., Caldarella, P., Anderson, D. H., Richardson, M. J., & Guzman, E. G. (2017). Addressing Problem Behavior at Recess Using Peer Praise Notes. *Journal of Positive Behavior Interventions*, *19*(2), 115–126.

Test, D. W., & Heward,W. L. (1983). Teaching road signs and traffic laws to learning disabled students. *Science Education, 64*, 129–139.

Tinker, M. A. (1939). Reliability and validity of eye-movement measures of reading. *Journal of Experimental Psychology, 19*, 732-746.

Tinker, M. A. (1946). The study of eye movements in reading. *Psychological Bulletin, 43*, 93-120.

Tinker, M.A. (1958). Recent studies of eye movements in reading. *Psychological Bulletin*, *55,* 215-231.

Traphagan, T., Kucsera, J. & Kishi, K. (2010). Impact of class lecture webcasting on attendance and learning. *Educational Technology Research and Development, 58*, 10–37.

UNESCO (2021). COVID-19 Impact on Education. H Retrieved from https://en.unesco.org/covid19/educationresponse on January 18, 2021.

U.S. Department of Education, National Center for Education Statistics, Parent and Family Involvement in Education Survey of the National Household Education Surveys Program (PFI-NHES:2016).

U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Spring 2016 and Spring 2017, Fall Enrollment component.

U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Spring 2020 and Spring 2021, Fall Enrollment component.

U.S. Department of Education, National Center for Education Statistics, 2003-04, 2007-08, and 2011-12 National Postsecondary Student Aid Study (NPSAS:04, NPSAS:08, and NPSAS:12).

Uzzaman, S., & Joordens, S. (2011). The eyes know what you are thinking: Eye movements as an objective measure of mind wandering. *Consciousness and Cognition*, *20*, 1882–1886. doi:10.1016/j.concog.2011.09.010

Van Wyk, M.M. (2020). Academic support under COVID-19 lockdown: what students think of online support e-tools in an OdeL course. *Interactive Technology and Smart Education*, Vol. ahead-of-print No. ahead-of-print.

Wang, H., Pi, Z., & Hu, W. (2018). The instructor's gaze guidance in video lectures improves learning. *Journal of Computer Assisted Learning*.

Wang, J., & Antonenko, P. (2017). Instructor presence in instructional video: Effects on visual attention, recall, and perceived learning. Computers in Human Behavior, 71, 79–89.

Wang, J., Antonenko, P., & Dawson, K. (2020). Does visual attention to the instructor in online video affect learning and learner perceptions? An eye-tracking analysis. *Computers & Education 146,* 103779*.*

Wankat, P. C. *The Effective Efficient Professor: Scholarship and Service*. Boston, MA: Allyn and Bacon, 2002.

Watson, J. (2008). *Blended learning: The convergence of online and face-to-face education.* Vienna, VA: North American Council for Online Learning.

Watson, J. B. (1914). *Behavior: An introduction to comparative psychology.* New York: Holt, Rinehart & Winston.

Watson, J. B. (1919). *Psychology from the standpoint of a behaviorist.* Philadelphia: Lippincott.

Watson, J. B. (1925). *Behaviorism.* New York: Norton.

What Works Clearinghouse. (2014). *Procedures and Standards Handbook, Version 3.0*. Washington, DC: Author. http://eric.ed.gov/?id=ED544775

Williamson, K. M., "Comparing the Effects of Two Rates of Specific Praise on Student Behavior" (2017). Doctoral Dissertations. 1356. https://opencommons.uconn.edu/dissertations/135

Wilson, K.A., & Korn, J.H. (2007). Attention during Lectures: Beyond Ten Minutes. *Teaching of Psychology, 34,* 85 - 89.

Wisher, R. A., & Olson, T. M. (2003). *The effectiveness of web-based training* (Research Report No. 1802). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

World Health Organization (2020). *How it spreads.* https://www.who.int/docs/default-source/coronaviruse/risk-comms-updates/update-20-epi-win-covid-19.pdf?sfvrsn=5e0b2d74_2. Retrieved January 15, 2021.

Wu, H.-K. and Huang, Y.-L. (2007), Ninth-grade student engagement in teacher-centered and student-centered technology-enhanced learning environments. *Science. Education, 91*. 727-749.

Zhang, Z., Li, Z., Liu, H., Cao, T., & Liu, S. (2019). Data-drived Online Learning Engagement Detection via Facial Expression and Mouse Behavior Recognition Technology. *Journal of Educational Computing Research,* 073563311982557, 1-24.

Zhao, Q., & Huang, X. (2019). Individual differences in response to attributional praise in an online learning environment. Educational Technology Research and Development. Zimmerman, E H. and Zimmerman, J. (1962). The alteration of behavior in a special classroom situation. *Journal of the Experimental Analysis of Behavior, 5,* 59-60.

Zinser, O., Young, J. G., & King, P. E. (1982). The influence of verbal reward on intrinsic motivation in children. *Journal of General Psychology, 106,* 85–91.

Zoom Video Communications Inc . (2016). Security guide. Zoom Video Communications Inc. Retrieved from http://www.zoom.us