JOURNAL OF SECONDARY PSYCHOLOGICAL STUDIES

Volume 4 Issue 2

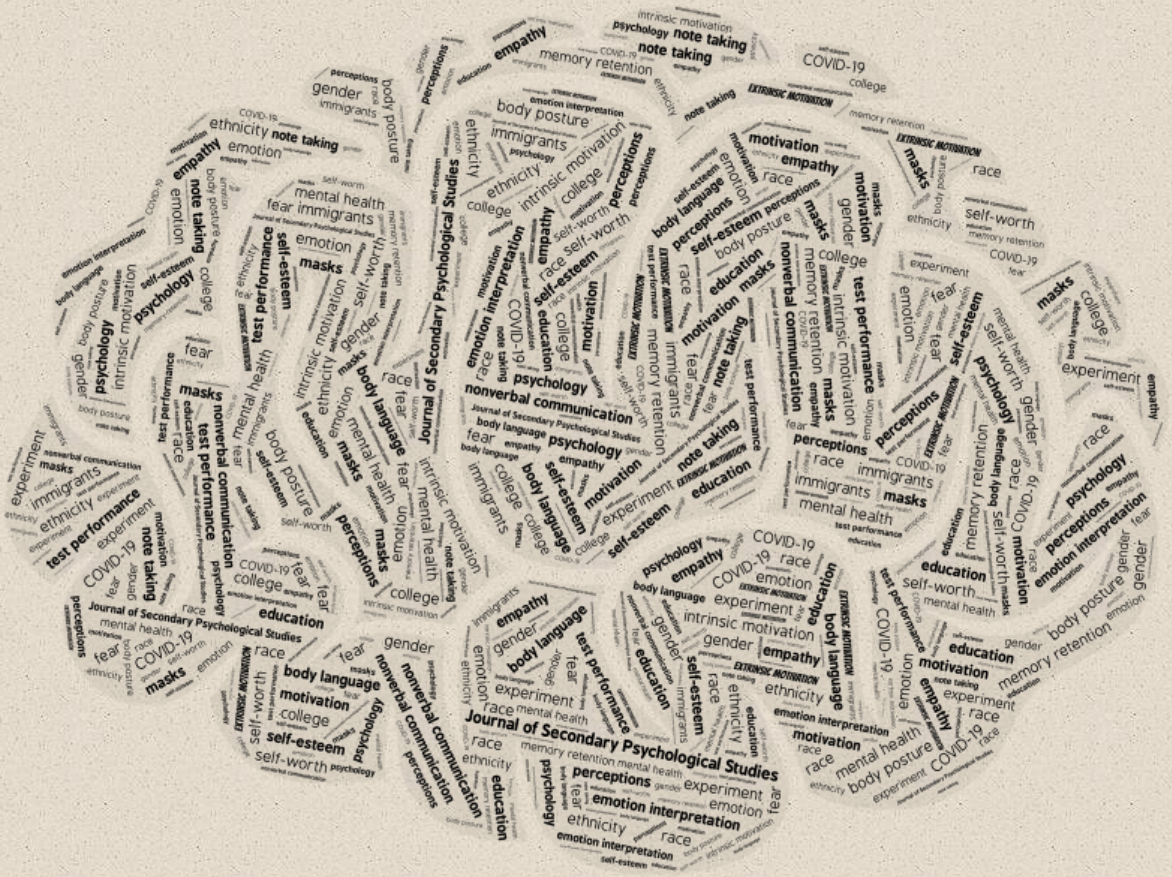


Table of Contents.

Staff.

Letter from the Editor.

Dear Readers,

The editorial board of the *Journal of Secondary Psychological Studies* is pleased to present the second issue of the fourth volume of our publication. I am incredibly grateful to have been able to work in conjunction with such talented authors and hardworking staff to produce another quality edition.

With over two years passing since the emergence of COVID-19, a key focus of this publication has been to encourage high schoolers to keep persevering and making scientific inquiries in this post-pandemic world. Thus, we were delighted to receive such a high quantity of submissions from across the country and the globe. In response, the journal took steps to provide increased support at a more sophisticated level, by expanding our editorial board and implementing new strategies of providing feedback to authors.

This edition features the work of researchers who are making the most of the new landscape with which they have been presented, with investigations into the impact of masks on emotional perceptions and how to better the ever-changing academic environment in regards to motivation and memory retention. We hope this publication inspires students to keep learning, growing, and discovering, pairing the difficulties of the past with curiosity and determination to work towards creating a better future.

Arya Sinha
Editor-in-Chief

Emotion Recognition in Asian Faces: How Much of a Difference Do Masks Make?

Rena Chen, Great Neck South High School, Great Neck, NY

E-mail: renazchen@gmail.com

Abstract

COVID-19 has introduced face masks into everyday life. The social implications of such a phenomenon are important to understand as they affect our day-to-day interactions. Difficulties in emotion recognition, in particular, have been linked to the usage of face masks, but related studies are few in number, present conflicting results, and focus almost exclusively on Caucasian faces. This study assesses the effect of face masks on emotion recognition of Asian faces by high school students. Students (N = 115) at a Long Island High School were recruited to take a survey on SurveyMonkey. Participants were asked to identify the emotion (out of 8) displayed by a given target face and indicate their confidence in their assessment using a bipolar scale rating from 1 = "very unconfident" to 7 = "very confident". A total of 64 Chinese target faces including eight emotions, two males, two females, and two conditions (unmasked vs masked) were presented to participants in a randomized order. Emotions included anger, content, disgust, fear, happiness, sadness, and surprise. This study found that face masks were associated with a significant reduction in accuracy and confidence in assessment of emotion recognition. Recognition of all emotions except neutral was significantly impaired by masks. Recognition of fear in particular was impaired when masked, which was not found in Caucasian faces in previous studies. Emotion misinterpretations also presented some concerning patterns of confusion including the tendency to confuse disgust for anger and fear for surprise. disgust, neutral, content, and happiness had the greatest reductions in confidence as a result of masks. Confusions in the emotion recognition of negative emotions are particularly concerning because the implications of misinterpreting negative emotions are greater than those of positive ones. The results of this study suggest that face masks hinder the ability to accurately assess emotions, posing a threat to everyday interactions and communication. Impairments in recognition as a result of masking were not associated with participant ethnicity. Findings of this study suggest that the use of face masks compromises emotional connections involving Asian target faces to a similar or even greater extent than Caucasian faces. Some limitations included the use of only four individual target faces, a lack of participant diversity, and difficulty in photoshopping face masks onto faces uniformly. Future research can further explore the relationships between acculturation and race/ethnicity of participant, incorporate multiple ethnicities and ages of target faces, and incorporate participants with greater diversity in terms of both age and ethnicity.

Literature Review

Since the outbreak of COVID-19 as a pandemic, wearing a mask has become an integral part of everyday life. While the public health benefits of mask-wearing pertaining to the transmission of COVID-19 has been extensively studied (Chu et al., 2020), the social implications of wearing masks have received less attention.

In the medical setting, the use of face masks has raised concerns about doctor-patient communication. A study conducted by Kratzke et al. (2021) found that patients had lowered positive perceptions of doctors wearing a face mask. On average, patients in the study believed a surgeon wearing a face mask was less empathetic and trustworthy compared to a surgeon wearing a clear face shield. They were also less comfortable with

the idea of having the surgeon they met wearing a face mask operate on them.

Emotion recognition is an important part of human interactions. Faces are the primary method of recognizing emotions and informing reciprocal expressions (Bruce & Young, 1986; Dimberg et al., 2000). In the education setting, face masks can negatively affect the relationship between teachers and students, a relationship built on emotional connections. Difficulties in emotion recognition can interfere with outward emotional responses to a peer's face, making it difficult to adjust one's behavior to match behavioral norms. Masks also impair verbal and non-verbal communication, opening the door to miscommunication (Spitzer, 2020).

Face Masks and Emotion Recognition

Recent studies have already suggested that face masks confuse emotion recognition. A study conducted by Carbon (2020) shows a significant decrease in accuracy of the emotions of anger, disgust, happiness, and sadness. Accuracy dropped by 14.2, 50.2, 24.6, and 13.4 percentage points, respectively. Confidence in perceived emotions for the emotions of anger, disgust, happiness, neutral, and sadness under a face mask were also significantly impaired. Each of the 36 participants in this study was presented with 72 pictures and was asked to assess the emotion depicted from a list of six choices (angry, disgusted, fearful, happy, neutral, and sad). They were also asked to indicate their personal confidence for each assessment using a Likert scale from 1 (very unconfident) to 7 (very confident). This study was conducted using Caucasian faces from the MPI FACES database (Ebner et al., 2010). No mention of participant ethnicity was made. Results were mirrored in a similar study that also used Caucasian faces, finding that target faces wearing face masks were associated with a significant decrease in accuracy ($p < .001$) of emotion recognition (in the same six emotions) compared to unmasked faces. Participants of this study lived in Germany and 90% indicated that German was their sole ethnicity (Grundmann et al., 2021).

It is generally agreed that the eye and mouth region are the most important regions of the face for emotion recognition (Blais et al., 2012; Spitzer, 2020). Previous emotion recognition studies have utilized tiled target images to determine which regions of the face are most informative. Tiled portions of a face image would randomly reveal themselves and respondents were instructed to stop the sequence once they recognized the emotion. Respondents that successfully identified emotions generally relied on tiles containing the eye and mouth region (Spitzer, 2020; Wegryzn et al., 2017). Therefore, when the mouth, one of these informative regions of the face, is obscured with a face mask, there is a large potential for impaired emotion recognition.

There is general agreement on how specific regions of the face may most accurately predict certain emotions, with a few contentions. Overall, detection of the emotions anger, fear, and sadness relies heavily on the eye region (Bombari et al., 2013; Schurgin et al., 2014; Wegrzyn et al., 2017). While assessment of the emotions disgust and happiness (described as "joy" in Schurgin et al., 2014) showed prolonged fixation in the mouth region for studies conducted by Wegrzyn et al. (2017) and Schurgin et al. (2014), Bombari et al. (2013) presented a slight contention, finding that the mouth region (in addition to the eye region) was also important for recognition of fear.

Existing studies pertaining to the effect of face masks on emotion recognition during the pandemic have generally limited their target faces to Caucasian faces (Carbon, 2020; Grundmann et al., 2021). Participants surveyed have also been of the same race/ethnicity as target faces (Caucasian). Thus, results of existing studies may not be applicable to faces of race/ethnicities other than Caucasian due to cultural differences in emotion recognition and expression (Jack et al., 2009; Beaupré & Hess, 2005).

In several studies, the ethnicity of participants has been found to influence emotion recognition. Jack et al. (2009) found that East Asian observers assessed emotions with bias towards the eye region while Western Caucasian observers distributed

their attention more evenly across the face. A similar study conducted among Sub-Saharan African, Chinese, and French Canadian individuals found that French Canadians were more accurate in decoding shame and sadness. The expression of emotion may also be influenced by the ethnicity of the target face. Fear, when expressed by Sub-Saharan Africans, was recognized with the greatest accuracy by all groups, possibly due to expressive morphological features of the face (Beaupré & Hess, 2005).

There are various arguments as to why individuals may have generally greater accuracy when decoding emotions expressed by their own ethnic group. One argument points to subtle differences in expression across different cultural groups, making it more difficult for out-group members to recognize emotions (Elfenbein & Ambady, 2002). Another suggests that general differences in cultural decoding (Matsumoto, 2002), attributed to culturally learned display rules and cultural norms (Ekman & Friesen, 1969), may be the culprit for variations in cross-cultural interpretations of emotions. However, previous studies have mixed results regarding the role of race/ethnic concordance in emotion recognition (Matsumoto, 1992; Prado et al., 2014). In Prado et al. (2014), although Australian Caucasian participants recognized Caucasian expressions significantly better than participants from Mainland China, Mainland Chinese respondents did not recognize Chinese expressions more accurately than Australian Caucasians. In Matsumoto (1992), ethnic concordance between American and Japanese participants/judges and poser faces was not found to significantly influence accuracy of emotion recognition.

Previous studies have generally found Asian observers to less accurately decode the negative emotions of anger, sadness, and fear compared to Caucasian counterparts (Biehl et al., 1997; Matsumoto, 1992). Explanations for this phenomenon support the existence of general differences in cultural decoding as a result of cultural norms. Asian culture emphasizes a collectivist nature, encouraging moderation of emotions and a lack of expression of negative emotions, which in turn leads to lower recognition of negative emotions compared to more individualistic western cultures (Beaupré & Hess, 2005; Prado et al., 2014). Collective cultures may be less tolerant of negative emotions than individualistic cultures, encouraging displays of emotion that limit group disharmony (Matsumoto, 1990). Contributing to this point, a study conducted by Prado et al. (2014) found that the emotions of fear, anger, and disgust in Chinese faces were least accurately assessed across Australian Caucasians, people of Chinese heritage living in Australia, and mainland Chinese respondents within the emotions tested (happiness, sadness, fear, anger, surprise and disgust). Exploring Asian faces in the context of masking can help contribute to a greater understanding of the cross-cultural effects of face masks on emotion recognition.

Many studies exploring emotion recognition have included participants with a wide range of ages but have often omitted the high school age of participants, generally including participants of elementary school age and adults (Carbon, 2020; Roberson et al., 2012). Younger participants tended to focus on the eye region of the face while adults, better versed in configural processing (the ability to analyze multiple facial features at the same time), put more emphasis on the mouth (Roberson et al., 2012; Schwarzer, 2000). Although the general accuracy of perceived emotions of younger participants was below that of adults (3–4-year-olds having about 60% accuracy, 5–6-year-olds with 70%, and 7–8-year-olds with around 85% compared to adults having >95% accuracy), the emotion recognition of participants under the age of nine was not impaired by face masks. This was unlike older children (9–10-year-olds) and adults who dropped from 90–100% correct to 60–70% correct emotion assessment (Roberson et al., 2012) when a face was masked. This result is likely due to adults being well versed in configural processing. Adults are used to analyzing multiple features of the face as a whole, so obscuring certain regions has led to larger decreases in accuracy compared to children who focus on one region of the face. Configural

processing is predicted to reach maturity around 15 years of age (Mondloch et al., 2002), which, combined with the use of face masks, has the potential to confuse emotion recognition in teenagers to a greater degree than younger children.

Another demographic factor in emotion recognition is gender. Females have historically performed better than males in emotion recognition tasks (Joseph & Newman, 2010). In a recent study, adult males compared to females experienced a significantly greater decline (odds ratio = 0.79) in accuracy of emotion recognition (Grundmann et al., 2021), indicating that emotion recognition was less impacted by face masks for female respondents.

The aim of this study was to determine the effect of face masks on emotion recognition of Asian faces by high school students. Studies similar in nature to this study have utilized only Caucasian faces and have also excluded participants in high school. This study seeks to broaden the understanding of the effects of face masks on emotion recognition of different racial/ethnic groups and age groups. I hypothesized that: *Hypothesis 1:* Accuracy and confidence of emotion recognition in masked faces would be lower than those without a mask. *Hypothesis 2:* Since high school students are at the age when configural processing matures, I hypothesized a significant difference between unmasked and masked conditions, more similar to confusion trends in adults (Roberson et al., 2012). *Hypothesis 3:* Because shared cultural heritage may aid Asian participants in assessing emotions of Asian faces, mask-induced impairment in emotion recognition would be lower for Asian participants than for non-Asian participants. *Hypothesis 4:* Recognition of disgust, happiness, and fear would be impaired to a greater degree because of their reliance on the mouth region for expression. *Hypothesis 5.* Female participants would perform better than males in emotion recognition under both unmasked and masked conditions.

Method

Participants

High school students were recruited from AP Psychology and Science Research classes at Great Neck South High School. Extra credit was offered as an incentive for participation in the study.

Materials

Participants were asked to complete an online survey. In the survey, participants were presented with a picture of a person's face, then asked to identify the emotion expressed (from a set of eight options) as well as to indicate how confident they were in their assessment. Confidence in assessment was presented as a 7-point bipolar scale with values of very unconfident, unconfident, slightly unconfident, neutral, slightly confident, confident, and very confident. Permission was obtained from Dr. Pei Sun to use target faces from the Tsinghua Facial Expression Database (FED) in the survey. Demographic questions regarding participant race/ethnicity, gender, and age were also asked at the conclusion of the survey.

The target faces in the Tsinghua FED have an overall 79.1% correct emotion identification rate, validated by 34 young (ages 19–35) and 31 older (aged 58–72) native Chinese face raters (Yang et al., 2020). Although other facial expression databases including Asian faces exist, the Tsinghua FED has the greatest number of target faces (a total of 110 individuals) and includes both older and younger faces. The pilot study compared emotion recognition of older versus younger faces. However, there was no significant difference in accuracy found. This study utilized the same group of target faces, this time only using younger faces, out of convenience. Additionally, the FED includes specifically Chinese faces while other databases include Japanese or other Asian groups. Chinese faces in particular were of interest for this study because a previous study found that negative emotions (such as anger, disgust, fear) were recognized less often

than positive emotions: making a case for exploring emotion recognition of specifically Chinese faces (Prado et al., 2014).

This study used pictures of four unique Asian faces: two young males and two young females (ages 19–35). For each individual face, eight emotions were shown, including anger, disgust, fear, sadness, surprise, neutral, content, and happiness. To develop a masked version of the target faces, a stock image of a blue surgical mask was photoshopped onto faces using the editing software GIMP. Face masks were individually placed on images and adjusted to fully cover the mouth and nose, the region obscured by a face mask. Refer to Appendix for examples of target face images.

In total, 64 face stimuli were used in this study. Each of the four individual faces (two females, two males) were shown expressing all eight emotions and were then shown again in the masked condition (4 individuals x 8 emotions x 2 conditions = 64). Questions were randomized and presented in a survey hosted by SurveyMonkey. All responses were anonymous.

Procedure

Participants were invited to a Google Classroom where they were given access to a PDF consent form. The consent form details survey instructions, procedures, and provides example questions. Although joining a Google Classroom requires an email address, surveys on SurveyMonkey do not require email addresses to be filled out. Therefore, all responses were anonymous and unable to be connected back to participants. Participants had the option of discontinuing the survey whenever they chose to and were not required to indicate their demographics. After submission of the parental assent form, students were emailed instructions and a link to the survey. The survey took on average 12 minutes to complete, and no time limit was imposed for any response.

Data Analysis

Data was exported to Excel using the "export file" button on SurveyMonkey and analyzed using linear regressions in STATA. Sample means and 95% confidence intervals were derived for accuracy in emotional recognition by the study condition (unmasked vs. masked), both overall and for each of the eight expressed emotions of the target faces. Distribution of reported confidence was examined by the study condition. To gain deeper insight into emotion misinterpretations, two confusion matrices for the unmasked and masked conditions were generated.

Analysis was conducted to further examine whether masks impaired emotional recognition differently by respondent gender, gender of the target face, and gender concordance between the respondent and the target face. Impairment in emotion recognition was compared between Asian and non-Asian respondents. Respondents who reported their race/ethnicity as "mixed" and had Asian heritage were counted as "Asian" in this analysis. Respondents who chose not to indicate their ethnicity or gender were included in analysis of accuracy for unmasked vs. masked conditions but excluded from comparisons of accuracy by demographic group.

Each respondent contributed 64 observations to the analysis. Because answers contributed by the same respondent were not independent, robust standard errors and 95% confidence intervals were derived by taking into account clustering at the respondent level by using the cluster() option in STATA 16.0 commands.

Results

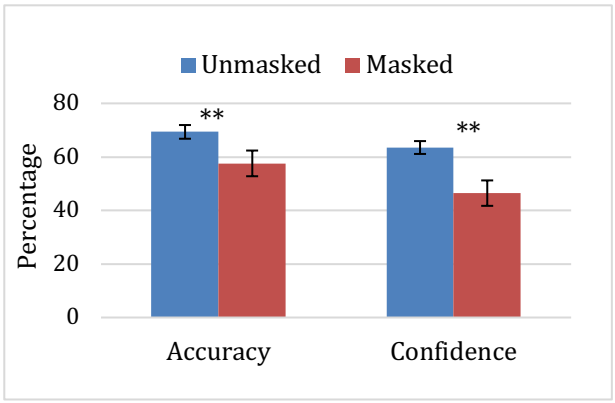
Participant Demographics

This study collected a total of 115 responses. Participants had ages ranging from 15–18 with an average age of 16.3. Roughly 68.70% of participants were Asian or Asian American, 18.26% White, 2.61% Hispanic or Latino, 0.87% Black or African American, 4.35% Mixed Race, and 5.22% preferred not to indicate their ethnicity. There were 67 females, 44 males, and four

participants preferred not to respond to the gender question.

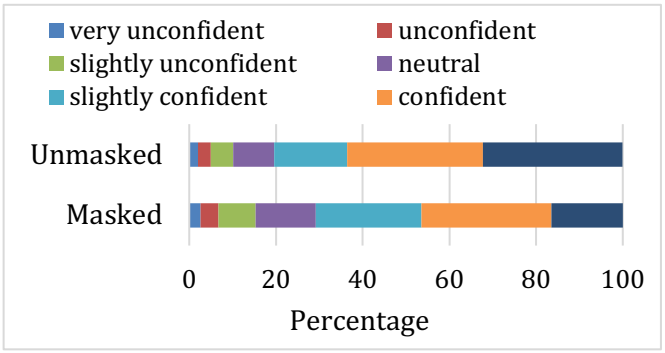
Mask-Induced Impairment in Recognition and Confidence

Overall, face masks were observed to significantly impair the accuracy of emotion recognition by 11.77 percentage points (from approximately 69.35% to 57.58%, $p < .001$; Figure 1). Overall confidence in assessment was also impaired by face masks. In the unmasked condition, 63.51% of participants were either confident or very confident in their assessment of a given emotion. Only 46.47% of participants were confident or very confident in their assessment of emotions in the masked condition (a 17.04 percentage-point decrease). The distribution of confidence in unmasked vs masked conditions is also of interest (Figure 2). In the masked condition, the frequency of respondents feeling "very confident" in their assessment of emotions dropped from 32.28% to 16.47% (a 15.81 percentage-point reduction) compared to the unmasked condition. Those reporting "confident" did not change much (30.00% vs. 31.22%) while the lower confidence ratings (neutral, slightly unconfident, unconfident, and very unconfident) all increased.



Note. Confidence level percentages represent the proportion of confidence ratings indicating "confident" and "very confident" in emotion recognition assessment. Error bars represent 95% confidence interval. $N = 115$ participants x 64 questions. ***$p < .001$.

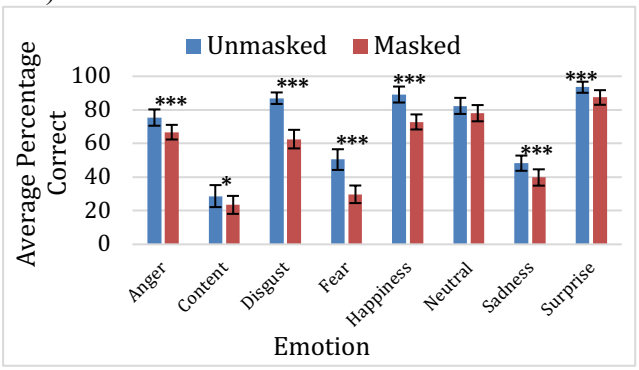
Figure 1. Accuracy of Emotion Recognition and Confidence Level Across Unmasked and Masked Conditions



Note. $N = 115$ respondents x 64 questions/respondent.

Figure 2. Distribution of Confidence Ratings Unmasked vs Masked

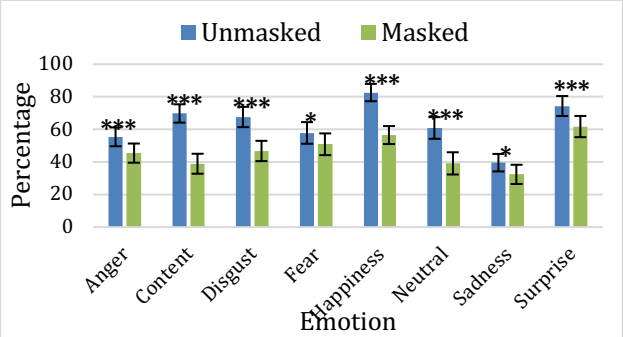
The degree of face-mask-induced impairment on recognition differed by emotion. All eight emotions (with the exception of Neutral) saw a significant decrease in the accuracy of emotion recognition. Neutral still had a borderline significant impairment of accuracy ($p = 0.059$). The magnitude of impairment differed across emotions (Figure 3). Compared to the unmasked condition, accuracy of emotion recognition in the masked condition for disgust went from 86.96% to 62.61% (24.35 percentage points, $p < .001$). Fear decreased from 50.43% to 29.78% (20.65 percentage points, $p < .001$). That of happiness from 89.13% to 72.83% (16.3 percentage points, $p < .001$). Anger decreased from 75.43% to 66.74% (an 8.69 percentage point difference, $p < .001$). Sadness from 48.26% to 39.78% (8.48 percentage points, $p = 0.001$). Surprise from 93.48% to 87.39% (6.09 percentage points, $p = .001$). Content dropped from 28.70% to 23.48% (a 5.22 percentage point difference, $p < .05$). Neutral from 82.39% to 78.04% (4.35 percentage points, $p = .059$).



Note. Error bars represent 95% confidence interval. $N = 115$ respondents x 64 questions/respondent (four unique target faces individually expressed all eight emotions).*$p < .05$. ***$p < .001$.

Figure 3. ***Accuracy of Emotion Recognition Across 8 Emotions, Unmasked vs. Masked***

Confidence in assessment decreased in all eight emotions when comparing masked to unmasked conditions (Figure 4). The greatest percentage point decreases of confidence were found in the emotions of Disgust (20.87 percentage points, $p < .001$), Neutral (21.74 percentage points, $p < .001$), Content (30.87 percentage points, $p < .001$), and Happiness (26.09 percentage points, $p < .001$).



Note. Confidence level percentages represent the proportion of confidence ratings indicating "confident" and "very confident" in emotion recognition assessment. Error bars represent 95% confidence interval. $N = 115$ participants x 64 questions. $*p < .05$. $***p < .001$.

Figure 4. ***Confidence in Assessment Across Eight Emotions, Unmasked vs Masked***

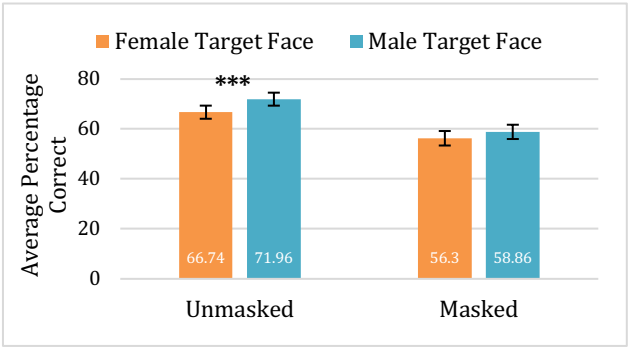
Confusion in Assessment of Emotions

In the confusion matrix presented (Table 1), the dark green diagonal stretching from the upper left to bottom right hand corner indicates a greater level of agreement between emotions expressed by target faces and those perceived by respondents. The darker the green, the greater the agreement. The diagonal is noticeably more uniform and darker in the unmasked condition compared to the masked condition. In the masked condition, emotions slightly confused with each other in the unmasked condition were generally similarly confused, just to a greater degree. In the unmasked condition, the emotions of Content, Fear, and Sadness had low accuracy, the expressed and perceived emotions agreeing less than 60% of the time. Content was accurately recognized 28.7% of the time, confused to the greatest degree with happiness (63.91%). Fear was correctly recognized 50.43% of the time, being confused with Disgust (24.35%) followed by Surprise (18.04%). Sadness was correctly recognized 48.26% of the time, commonly confused with Disgust (22.61%), followed by Anger (11.09%). Some other observations include Anger's tendency to be confused with Disgust 10.43% of the time, Happiness' slight confusion with Content (7.39%), and Neutral's with Content (6.09%).

Accuracy in recognition for all eight emotions declined in the masked condition (Figure 3). Several emotions were confused for each other more than others (Appendix A). For anger, confusion with disgust became less prominent and confusion spread out more evenly amongst the other emotions. In particular, anger's confusion for neutral increased from 4.78% to 10%. Content's tendency to be confused with Happiness (63.91% in the unmasked condition) decreased by 28.04 percentage points in the masked condition. Misinterpretations of content when masked spread more evenly between happiness and neutral (35.87% and 31.52%, respectively). Disgust, which showed no leanings towards confusion with other emotions when unmasked, was confused with anger 25.22% of the time when masked. The emotion of fear, which showed previous tendencies to be confused with disgust and surprise, leaned heavily towards surprise (rising from 18.04% to 53.48%) in the masked condition. The tendency for happiness to be confused for content was worsened under a mask (7.39% confusion unmasked vs 15.65% masked) with content. Neutral's leaning towards confusion with content (6.09% confusion when unmasked) was increased slightly in the masked condition (7.61%). Neutral's confusion with sadness was also worsened with the addition of a face mask (5.65% unmasked vs. 10.0% masked). Sadness's tendency to be confused with anger and disgust when unmasked decreased slightly in the masked condition and confusion with neutral increased from 8.91% to 16.52%. Surprise continued to show minimal confusion with other emotions. Overall, face masks increased the ambiguity of each emotion.

Associations Between Mask-Induced Impairment and Demographic Variables

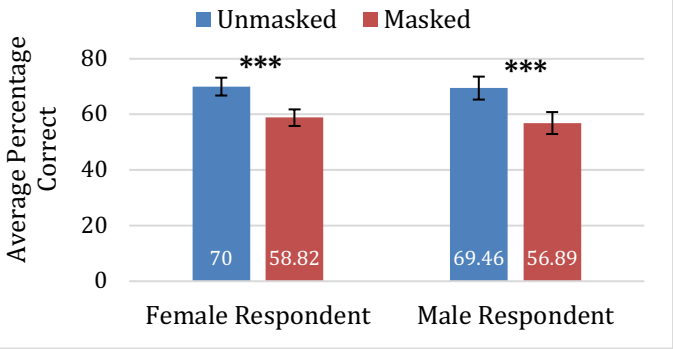
Masks significantly impaired emotion recognition for target faces of both genders ($p < .001$) but lessened the difference between emotion recognition of the two target face genders (Figure 5). In the unmasked condition, male target faces were associated with higher accuracy in emotion recognition when compared to female target faces in both unmasked (5.22-percentage-point higher, $p < .001$) and masked (2.56-percentage-point higher, not significant difference) conditions.



Note. Error bars represent 95% confidence interval. $N = 115$ respondents x 64 questions/participant. ***$p < .001$.

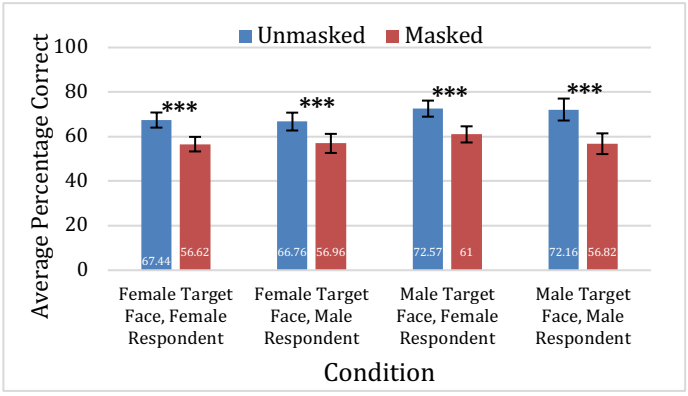
Figure 5. ***Accuracy of Emotion Recognition by Target Face Gender***

In both respondent genders, the masked condition significantly impaired accuracy of emotion recognition ($p < .001$) (Figure 6). Impairment of recognition did not differ by respondent gender. Males had a 1.38 percentage point greater impairment than females, but this difference was not statistically significant ($p = .426$). Respondent and target face gender concordance was not associated with mask-induced impairment (Figure 7).



Note. Error bars represent 95% confidence interval. $N = 111$ participants x 64 questions/participant (four participants preferred not to indicate their gender). ***$p < .001$.

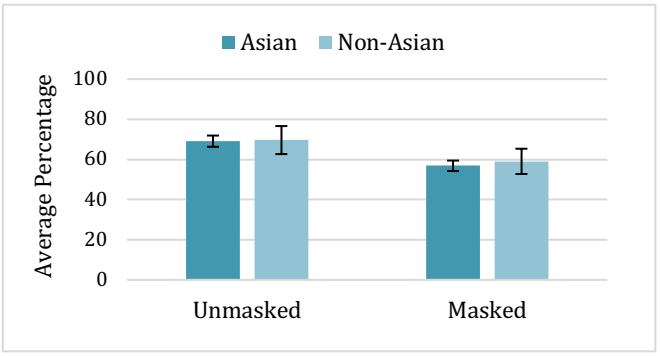
Figure 6. ***Accuracy of Emotion Recognition Across Respondent Gender***



Note. Error bars represent 95% confidence interval. $N = 111$ participants times 64 questions (four participants preferred not to indicate their gender). ***$p < .001$.

Figure 7. ***Accuracy of Emotion Recognition by Target Face and Respondent Gender Concordance***

Race/ethnicity of respondents was not associated with mask-induced impairment in emotion recognition (Figure 8). The mean accuracy for Asian respondents was 1.67-percentage-points lower than that of non-Asian respondents but this difference was not significant ($p = 0.323$).



Note. Average percent accuracy of emotion recognition by Asian vs. Non-Asian participant. Error bars represent 95% CI. $N = 109$ respondents x 64 questions/respondent. Results graphed are not significant.

Figure 8. ***Average Accuracy of Emotion Recognition Asian vs. Non-Asian Participants***

Discussion

This experiment sought to explore face masks' effects on emotion recognition in Asian faces by high school students. Previous studies of such effects have predominantly focused on Caucasian faces (Carbon, 2020; Grundmann et al., 2021). Asian faces, often described as less expressive with their emotions (Yamamoto & Li, 2012), are worth studying in the context of face masks given their baseline ambiguity and

heightened risks of misinterpretation when masked. Exploring the effect of face masks on emotion recognition in Asian faces can also increase the overall understanding of emotion recognition across different race/ethnicities. Respondents of the high school age have not been represented in previous studies (Roberson et al., 2012) and are worth studying in part because of the immense social changes present at school. As the COVID-19 pandemic continues and schools reopen with mask mandates, understanding how masks may impair social interactions will be important first steps towards strategies to mitigate such impairment.

Overall, face masks were associated with a 16.97% relative reduction in the accuracy of emotion recognition in Asian target faces. Participants correctly identified unmasked target faces 69.35% of the time (Figure 1), 9.73-percentage-points lower than the Tsinghua FED's accuracy in validation tests. This slight change in accuracy may be attributed to the FED being validated by adult respondents (Yang et al., 2020). Since configural processing matures at age 15 (Mondloch et al., 2002), it is probably not surprising that adults assessed emotions with greater accuracy than high schoolers. It is also possible that the validators of the FED (Chinese people who live in China) had a greater cultural advantage in identifying emotions compared to high school students in the United States, leading to higher accuracy of emotion recognition in the validation versus the results of this study.

The substantial and significant decline in expressed confidence by respondents indicates that face masks impaired respondents' sureness in their assessments. Interestingly, the "very confident" rating dropped drastically (15.81-percentage-points) while "confident" stayed fairly consistent, displaying a mere 1.22 percentage point increase in the masked condition (Figure 2). It is possible that respondents switched from "very confident" to "confident" in their assessments and from "confident" to the other lower confidence levels, inadvertently keeping the percentage of "confident" constant. Respondents could have also just been generally confident in their recognition of emotions. Mask-induced emotion recognition impairment was significant for seven out of eight emotions tested. The emotions of anger (11.52% relative decrease in accuracy in the masked condition compared to unmasked), content (18.19% decrease), disgust (28.0% decrease), Fear (40.95% decrease), happiness (18.29% decrease), sadness (17.57% decrease), and surprise (6.51% decrease) all saw significant impairment in recognition (Figure 3). These findings were also consistent with those of Carbon's 2020 study (conducted without the emotions of content and surprise) which found that masks significantly impaired emotion recognition in Caucasian faces for the emotions of anger, disgust, happiness, and sadness ($p < .001$). Also consistent with Carbon (2020), recognition of the neutral emotion in this study was not significantly impaired. Although fear was not significantly impaired by face masks in Caucasian faces (Carbon, 2020), it was for Asian faces in this study.

The relative decreases in accuracy correspond with the regions of the face thought to best predict their recognition. Anger and sadness, thought to be expressed predominantly in the eye region (Bombari et al., 2013; Schurgin et al., 2014; Wegrzyn et al., 2017), had relatively low decreases in accuracy (11.52% and 17.57% decreases respectively) compared to other emotions when masked. Surprise also had a relatively low decrease in accuracy (6.51%) which suggests its reliance on the eye region for expression. On the other hand, disgust and happiness, thought to be expressed predominantly in the mouth region (Bombari et al., 2013; Schurgin et al., 2014; Wegrzyn et al., 2017), had relatively higher decreases in accuracy compared to other emotions when masked (28% and 18.29% decreases respectively). This study found that fear was associated with a startling decrease in accuracy (40.95% relative reduction) in the masked condition. A previous study found that recognition of fear in Caucasian faces relied more heavily on the eye region (Wegryzn et al., 2017), which is left unobscured by a face mask. It is possible that Asians tend to utilize more of the obscured mouth region to express fear compared to Caucasians,

leading to greater impairment. In a previous study, East Asian observers were presented with Japanese faces and Western Caucasian observers with Caucasian faces. The study found a significant deficit in East Asian observers in recognizing fear, but not among Western Caucasians (Jack et al., 2009). Results of this study suggest that Asian faces rely heavily on the mouth region to express fear, thus, leading to significant impairment in assessment in the masked condition.

In this study, certain emotions were confused with others in both the unmasked and masked conditions (Table 1). Negative emotions such as anger, disgust, and sadness were confused with each other, and positive or neutral emotions such as content, neutral, and happiness were confused with each other. Compared to other emotions, neutral, surprise, anger, happiness, and sadness displayed relatively smaller degrees of disagreement between the expressed and perceived emotion.

Sadness had the lowest rate of accuracy unmasked (48.26%) second only to content and was confused with multiple emotions in both the unmasked and masked conditions. Confusion of sadness for other emotions became more disbursed under the masked condition. One noticeable result is that misinterpretation of sadness shifted primarily from disgust when unmasked but to neutral when masked. This change suggests that masks have the effect of dampening the intensity or severity of emotions, a particularly concerning possibility given that emotions are essential to communication and facilitate human interactions (Bruce & Young, 1986).

Of note, under the unmasked condition, content (28.7% accurate) was confused for happiness at an alarming 63.91% of the time. In the masked condition, content was confused with happiness (35.87%) and neutral (31.52%), suggesting the ambiguity of content to study respondents. In the FED, content is described to be a "smile without teeth" or a subtle version of happiness. Participants in this study were not given definitions of emotions shown. Instead, they answered questions based on their personal interpretations. The personal interpretations of emotions may have also caused some emotions to be more accurately assessed than others.

The emotions of disgust and fear had prominent decreases in recognition in the masked condition (24.35 and 20.65 percentage points, respectively). Unmasked, disgust had a 86.96% accuracy rate with a small 5% confusion with anger. This confusion was amplified in the presence of a face mask: assessment of disgust fell to 62.61% accurate and was confused for anger 25.22% of the time. Masked disgust (43.7%) in Carbon (2020) was also heavily confused with anger (37.8% of the time). This confusion is of concern because when masked, a person who is aversive of a situation may be perceived as an irritated or even aggressive person. Fear in the unmasked condition was correctly assessed 50.43% of the time, confused with disgust (24.35%) and surprise (18.04%) most prominently. Masked fear, however, had an accuracy rate as low as 29.78%, and was perceived as surprise 53.48% of the time. This is another particularly concerning finding because of the sheer magnitude of confusion, but also the implications: a person who is afraid of something in a certain situation can be misinterpreted as a person who is feeling shocked at an unexpected circumstance. In addition, this finding seems to apply specifically to Asian faces. The recognition of fear in Caucasian faces showed little confusion with other emotions (92.5% correct assessment when unmasked, 93.5% correct masked) (Carbon, 2020). It is of interest to note that a confusion matrix generated in the validation of the FED displayed similar unmasked confusions in recognition. Disgust was confused with anger 12.09% of the time and fear confused with surprise 18.89% of the time. The slight inherent confusions between database images may play a role in accuracy of emotion recognition in this study.

The gender of target faces had no significant effect on emotion recognition. In a previous study conducted by McDuff et al (2017) comparing the expressiveness of female versus male faces in participants from France, Germany, UK, US, and China, results suggested that female faces were generally more expressive. The same

study found that male facial actions tended to be centered in brow furrows while women used more smiles and inner brow raises. Although brows are a dominant part of the upper region of the face, a region unobscured by face masks, emotion recognition of male target faces in this study only outperformed female target faces by 2.56 percentage points in the masked condition (Figure 5). These results highlight the importance of the mouth region for both genders in emotion expression and recognition.

This study surveyed participants of different racial/ethnic backgrounds. This is in contrast with existing studies of mask-induced impairment in emotion recognition whose participants were the same race/ethnicity as target faces (Carbon, 2020; Grundmann et al., 2021). This study found no difference in mask-induced impairment between Asian and non-Asian participants (Figure 8). The smaller percentage of non-Asian participants in the sample (~30%) may have made it difficult to statistically detect small differences. Participants in this survey attended a high school in the United States and were either born and raised in the United States or highly acculturated, therefore sharing the same race/ethnicity with a target face may have not affected the ability to accurately assess emotions in Asian faces. A previous study comparing emotion recognition of Australian Caucasians (unmasked faces) and people of Chinese heritage living in Australia had comparable findings. Overall, recognition scores between the two groups were similar (82.6 vs. 80.8 % respectively), supporting the idea that acculturation may dull differences in recognition associated with decoding rules pertaining to certain racial/ethnic roots (Prado et al., 2014). It is also possible that non-Asian participants in this study were accustomed to interpreting emotions expressed by Asians as the study population was from a school where Asians are the majority.

Comparing the accuracy of assessment for the eight emotions of the FED (validation) with the unmasked accuracy in this study, the emotions of anger, happiness, and neutral all had differences of less than 10 percentage points (Yang et al., 2020). Jack et al. (2009) suggest that compared to Western Caucasian observers, East Asian observers demonstrate a deficit in recognizing the emotions of fear and disgust. In this study, disgust was recognized 86.76% of the time compared to the 71.06% in the validation, reflecting this trend. However, results of this study may be impacted by the particular target faces used (only four out of the FED's total 110) as unique facial features may facilitate recognition of certain emotions better than others. Recognition of fear was higher in the validation than in this study (62.29% vs 50.43%), and surprise was lower in the validation than in this study (80.29% vs. 93.48%). Content's accuracy in this study is drastically different from its validation (90.71%), coming in at a concerning 28.7%. sadness was validated at 76.41% but was accurately assessed 48.29% of the time in this study. The drastic differences in accuracy for content and sadness are possibly also the result of a difference in age and cultural differences between validators of the FED (adults) and participants in this study (high schoolers).

Although face masks impair emotional recognition and negatively impact social interactions, this is not a reason to disregard their use during a pandemic such as the one we are experiencing. Body language, verbal communication, and social context are all tools that can contribute to recognition of emotional states (Abramson et al., 2021; Golan et al., 2006). Awareness of how masks cause ambiguity in emotions can encourage people to assess emotions more carefully and with greater sensitivity. Findings of this study suggest that such attention and sensitivity should be exercised universally regardless of demographic attributes or concordance between the two sides of a conversation or interaction.

Limitations and Future Directions

This study had a few limitations. Participants in this study lacked diversity as they were 60.36% female and 68.7% Asian, all from the same community. The sample population was not completely random—all were either in an AP Psychology class or a Science Research class. This

study also only used Chinese target faces. It is unknown how results may be applicable to other Asian faces, such as Indian, Japanese, or Korean faces.

In addition, to contain the length and burden of the survey, older target faces were not included in this study. Although a pilot study conducted previously did not find significant differences in impairment of recognition by the age of the target faces, this study was thus unable to formally compare emotion recognition and mask-induced impairment by target face age. In the validation, the Tsinghua FED had an overall 79.08% accuracy of emotion recognition, so the database itself may not accurately represent all emotions.

Another measure taken to reduce the length of the survey was the use of only four individual target faces out of the 110 available in the Tsinghua FED. With the limited number of target faces, an individual's unique facial expressions may have had undue influence on emotion recognition by respondents. However, subjecting all respondents to the same set of faces (rather than presenting each with a random set of faces) boosts the internal validity of the study. A lack of uniformity in the photoshopping of face masks onto target faces may also have impacted study results since it is impossible to manually edit images and achieve unvaried results.

Findings of this study raised an important question: are cultural influences more important in interpreting emotions than someone's race/ethnicity? This study found no significant differences in emotion recognition between Asian and non-Asian respondents, suggesting that shared culture may play a greater role than race concordance. To more specifically tackle the question regarding the role of culture (vs. race/ethnicity) on emotion recognition, future studies may consider surveying participants with both racial/ethnic and cultural diversity, such as native Chinese, American-born Chinese, and American-born Caucasians, and comparing their accuracy of emotion recognition. Differences in emotion confusion patterns across different cultures and ethnicities would also be an interesting topic to pursue due to the importance of cross-cultural understanding. A study exploring this could include Asian participants assessing emotions of Caucasian target faces. It would also be of interest to compare impairments in recognition across age groups in the future to further explore the effect age has on emotion recognition.

Conclusion

This study assessed the effect of face masks on emotion recognition in Asian faces. It was found that the use of face masks significantly impaired emotion recognition of Asian faces for seven out of eight of the tested emotions (all except neutral) as well as lowered participant's confidence in their assessments. Notable mask-induced impairment was found in the emotions of disgust and fear. The severe impairment in recognizing fear was unique to Asian target faces and not observed in Caucasian faces from previous studies (Carbon, 2020). In the masked condition, disgust was often confused for anger, and fear was often confused for surprise. Demographic factors including target face gender, participant gender, and participant race/ethnicity were not significantly associated with accuracy of emotion recognition. Awareness of impaired ability to recognize emotions when masked can help people to be more conscious during face-to-face interactions and to leverage other modes of communication, potentially minimizing mask-induced impairment to social interactions.

Author Note

I would not have been able to complete this project without the guidance of several mentors throughout this past year. Thank you to Dr. Truglio and Dr. Hersh for always encouraging me and never giving up on me, especially when I felt discouraged. Thank you for accompanying me through survey design conundrums, and for all the time you have spent giving me feedback and advice. I would also like to thank my mother, Dr. Yuhua Bao (Weill Cornell Medical School) for guiding me through data analysis. A special thank

you as well to Dr. Pei Sun (Tsinghua University) for providing me access to the Tsinghua Facial Expression Database. This study was conducted using human subjects at Great Neck South High School under the supervision of teachers at Great Neck South High School: Dr. James Truglio, Dr. Carol Hersh, and Ms. Spinelli. All survey procedures were approved by the high school IRB and all data collected was anonymous.

References

Abramson, L., Petranker, R., Marom, I., & Aviezer, H. (2021). Social interaction context shapes emotion recognition through body language, not facial expressions. *Emotion, 21*(3), 557–568. https://doi.org/10.1037/emo0000718

Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology, 36*(3), 355–370. https://doi.org/10.1177%2F0022022104273656

Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., & Ton, V. (1997). Matsumoto and Ekman's Japanese and Caucasian facial expressions of emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior, 21*(1), 3–21. https://doi.org/10.1023/A:1024902500935

Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia, 50*(12), 2830–2838. https://doi.org/10.1016/j.neuropsychologia.2012.08.010

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*(3), 305–327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *Quarterly Journal of Experimental Psychology, 66*(12), 2426–2442. https://doi.org/10.1080%2F17470218.2013.789065

Carbon, C. C. (2020). Wearing face masks strongly confuses counterparts in reading emotions. *Frontiers in Psychology, 11*, 1–8. https://doi.org/10.3389/fpsyg.2020.566886

Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., & Schünemann, H. J. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and Covid-19: A systematic review and meta-analysis. *The Lancet, 395*(10242), 1973–1987. https://doi.org/10.1016/S0140-6736(20)31142-9

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*(1), 86–89. https://doi.org/10.1111%2F1467-9280.00221

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavioral Research Methods, 42*, 351–362. https://doi.org/10.3758/BRM.42.1.351

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*(1), 49–98. https://doi.org/10.1515/9783110880021.57

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*(2), 203–235. https://doi.org/10.1037/0033-2909.128.2.203

Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 36*(2), 169–183. https://doi.org/10.1007/s10803-005-0057-y

Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. *PLOS ONE, 16*(4). https://doi.org/10.1371/journal.pone.0249792

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology, 19*(18), 1543–1548. https://doi.org/10.1016/j.cub.2009.07.051

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology General, 141*(1), 19–25. https://doi.apa.org/doi/10.1037/a0023463

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95*(1), 54–78. https://psycnet.apa.org/doi/10.1037/a0017286

Kratzke I. M., Rosenbaum, M. E., & Cox, C. (2021). Effect of clear vs standard covered masks on communication with patients during surgical clinic encounters. *JAMA Sugery, 156*(4), 372–378. https://doi.org/10.1001/jamasurg.2021.0836

Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and Emotion, 14*, 195–214. https://doi.org/10.1007/BF00995569

Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-Cultural Psychology, 23*(1), 72–84. https://doi.org/10.1177%2F0022022192231005

Matsumoto, D. (2002). Methodological requirements to test a possible in-group advantage in judging emotions across cultures: Comment on Elfenbein and Ambady (2002) and evidence. *Psychological Bulletin, 128*(2), 236–242. https://doi.org/10.1037/0033-2909.128.2.236

McDuff, D., Kodra, E., Kaliouby, R. E., & LaFrance, M. (2017). A large-scale analysis of sex differences in facial expressions. *PLOS ONE*, *12*(4). https://doi.org/10.1371/journal.pone.0173942

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*(5), 553–566. https://doi.org/10.1068/p3339

Prado, C., Mellor, D., Byrne, L. K., Wilson, C., Xu, X., & Liu, H. (2014). Facial emotion recognition: A cross-cultural comparison of Chinese, Chinese living in Australia, and Anglo-Australians. *Motivation and Emotion, 38*(3), 420–428. https://psycnet.apa.org/doi/10.1007/s11031-013-9383-0

Roberson, D., Kikutani, M., Döge, P., Whitaker, L., & Majid, A. (2012). Shades of emotion: What the addition of sunglasses or masks to faces reveals about the development of facial expression processing. *Cognition, 125*(2), 195–206. https://doi.org/10.1016/j.cognition.2012.06.018.

Schurgin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J. Y., & Franconeri, S. L. (2014). Eye movements during emotion

recognition in faces. *Journal of Vision, 14*(13). https://doi.org/10.1167/14.13.14

Schwarzer, G. (2000). Development of face processing: The effect of face inversion. *Child development, 71*(2), 391–401. https://doi.org/10.1111/1467-8624.00152

Spitzer, M. (2020). Masked education? The benefits and burdens of wearing face masks in schools during the current Corona pandemic. *Trends in Neuroscience and Education, 20*. https://doi.org/10.1016/j.tine.2020.100138

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLOS ONE, 12*(5). https://doi.org/10.1371/journal.pone.0177239

Yamamoto, Y., & Li, J. (2011). Quiet in the eye of the beholder: Teacher perceptions of Asian immigrant children. *The Impact of Immigration on Children's Development,* 1–16. https://doi.org/10.1159/000331021

Yang, T., Yang, Z., Xu, G., Gao, D., Zhang, Z., Wang, H., Liu, S., Han, L., Zhu, Z., Tian, Y., Huang, Y., Zhao, L., Zhong, K., Shi, B., Li, J., Fu, S., Liang, P., Banissy, M. J., & Sun, P. (2020). Tsinghua facial expression database – A database of facial expressions in Chinese young and older women and men: Development and validation. *PLOS ONE, 15*(4). https://doi.org/10.1371/journal.pone.0231304

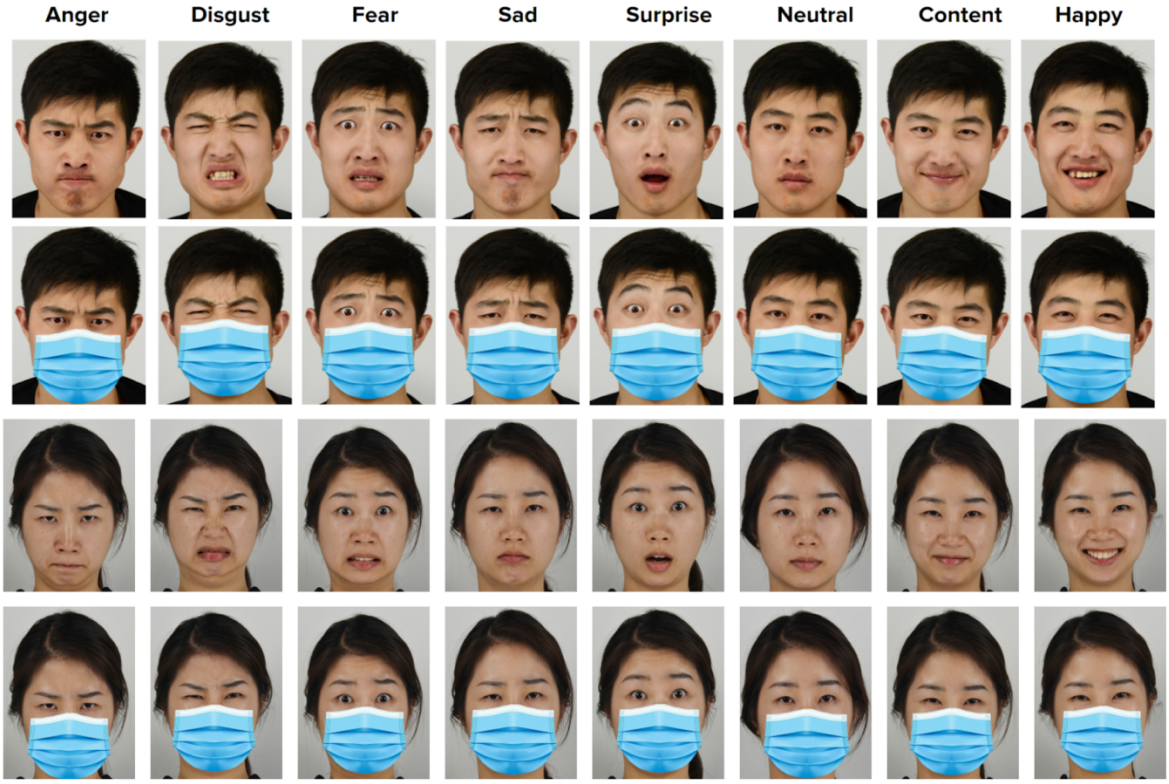
Appendix A: *Confusion Matrix of Expressed Emotion and Perceived Emotion Across Eight Emotions, Unmasked vs Masked*

Expressed Emotion

Perceived Emotion	Anger	Content	Disgust	Fear	Happiness	Neutral	Sadness	Surprise	
Anger	75.43	0.43	5	2.17	0	2.17	11.09	0.22	**Unmasked**
Content	2.17	28.7	0.43	1.52	7.39	6.09	4.78	0.65	
Disgust	10.43	1.96	86.96	24.35	0.87	1.3	22.61	0.65	
Fear	1.96	0.43	3.48	50.43	0.65	1.09	2.39	2.39	
Happiness	0.87	63.91	0.65	0.65	89.13	1.3	0.43	1.09	
Neutral	4.78	3.26	0	1.74	1.52	82.39	8.91	1.3	
Sadness	3.48	0.87	3.26	1.09	0.43	5.65	48.26	0.22	
Surprise	0.87	0.43	0.22	18.04	0	0	1.52	93.48	
Anger	66.74	1.96	25.22	6.74	1.09	0.43	9.78	0.43	**Masked**
Content	4.78	23.48	0.87	1.52	15.65	7.61	4.57	1.74	
Disgust	7.83	3.48	62.61	4.78	1.3	0.87	17.17	1.3	
Fear	1.74	1.3	1.96	29.78	0.65	1.09	8.48	3.48	
Happiness	5	35.87	1.96	0.87	72.83	1.96	1.09	2.61	
Neutral	10	31.52	1.74	1.09	7.39	78.04	16.52	2.17	
Sadness	1.96	1.96	4.57	1.74	0.87	10	39.78	0.87	
Surprise	1.96	0.43	1.09	53.48	0.22	0	2.61	87.39	

Note. Numbers are column percentages out of 100. Darker shading indicates higher percentages (White = 0%, Dark Green = 100%). Cells on the diagonal stretching from the upper left-hand corner to the lower right-hand corner for each condition represent agreement between perceived and expressed emotions.

Appendix B: *Experimental Stimuli*



The Effect of Note Taking on Memory Retention

Jake Konigsberg, Roslyn High School, Roslyn, NY

E-mail: jkonigsberg23@roslynschools.org

Abstract

In the age of technology, students have become reliant on laptops to take notes. They do not have to worry about carrying around a binder full of paper and the notes are automatically saved on their device. While convenient, previous studies have shown contradictory results over whether the benefits of note taking by hand are also seen when taking typed notes. The present study investigated the effect of using a computer or paper to take notes and teaching note taking on memory retention. To conduct this study, 154 ninth grade students were randomly assigned to one of five conditions: no notes, handwriting notes and not taught how to take notes, handwriting notes and taught how to take notes, typing notes and not taught how to take notes, and typing notes and taught how to take notes. Participants who were taught to take notes were shown a 2.5-minute clip going over the process of outline notes. Then, all participants were shown a 15-minute video with those in note taking conditions instructed to take notes on the content presented. Afterwards, everyone was given a 10-minute test with 10 points worth of factual questions and four points worth of conceptual questions. It was found that those who did not take notes did not perform significantly differently than those who took typed or handwritten notes for factual questions, yet those who took handwritten notes performed significantly better than those who took typed notes. For conceptual questions, it was found that those who took handwritten notes performed significantly better than those who took typed notes and those who did not take notes. Additionally, it was revealed that teaching how to take notes did not have a significant impact on test performance for either factual or conceptual questions. Overall, this experiment suggests that schools should make a concerted effort to reduce the use of electronic devices in the classroom, resorting back to pen and paper.

Literature Review

Problem

Over the past two decades, the total number of applicants applying to college increased by over 150%, causing colleges to become more competitive than ever (Selingo, 2022). With colleges becoming extremely competitive, high school students are increasing the rigor of their course loads as a way to compete with their peers. For example, the proportion of high school students taking at least one Advanced Placement course, typically seen as the highest-level course available to high schoolers, has increased from 28.6% in 2011 to 34.9% in 2021 (College Board, n.d.). Since student course loads are harder, they have to worry about absorbing more information through the practice of notetaking during lectures (Akintunde, 2013). However, in recent years, the landscape for notetaking has greatly changed. For one, the rise of technology means there is a greater availability of internet resources, such as Google Classroom, which allows information, which previously could only be accessed during a lecture, to be posted online (KIPP NJ, 2016). Therefore, notetaking is no longer the sole method of revisiting information, which was previously one of the major benefits of notetaking (Jacobs, 2008). As a result, the practice of notetaking has been increasingly questioned and its importance has declined in school curricula (Singer & Samson, 2019). Additionally, the rise of technology means that the use of laptops and tablets in schools have greatly increased, resulting in the number of students who type their notes to also increase (KIPP NJ, 2016). However, there is disagreement over the difference in effectiveness of typing versus handwriting notes (Mueller & Oppenheimer, 2014; Urry et al., 2021). Therefore,

the purpose of this study was to determine how the different ways of recording notes and being taught to take notes affects memory retention.

Notetaking on Retention of Information

While there is growing skepticism towards the practice of notetaking (Singer & Samson, 2019), it has been found to increase the retention of information regardless of the type of notes one takes. Some suggest it is because notetaking allows students to process the information by summarizing and synthesizing it (Jacobs, 2008). Others suggest that notetaking is beneficial because it allows students an opportunity to record information so that it is accessible when needed to study for a test (Jacobs, 2008). For the most part, students who take notes do better on both essay and multiple-choice tests. For instance, Akintunde (2013) found that on a multiple-choice test based on a video about stress, those who took handwritten notes did significantly better on the test with an average score of 6.5 out of 10 than those who did not take any notes who had an average score of 5.25 out of 10. On the essay test, those who took handwritten notes also did significantly better as they had an average score of 11.98 out of 15, while those who did not take notes had an average score of 10.95 out of 15 (Akintunde, 2013). The improvement in test performance caused by note-taking was also seen in another study where college students took a test on a lecture. Those that took notes had a mean score of 15.3 while those who did not take notes had a mean score of 12.4 (Fisher & Harris, 1973).

Type of Notes

While there is agreement that taking notes is significantly more beneficial than not taking notes, the rise of technology has changed the practice of note taking with the introduction of typing notes (KIPP NJ, 2016). Studies investigated the effect of typing and handwriting notes on the ability to answer factual questions and found handwriting notes does not significantly improve test performance compared to typing notes (Mueller & Oppenheimer, 2014). Replications of this study by Mitchell and Zheng (2017) and Urry et al. (2021) also found that there was no significant difference between handwriting and typing notes when it came to factual questions.

On the other hand, studies that have compared the effect of typing to handwriting notes on students' test performance on open-ended, conceptual questions have found contradictory results. Oppenheimer and Mueller (2014) found that for conceptual questions, those who typed their notes did significantly worse than those who handwrote their notes. The study also found that students who typed notes on a video had an average of 14.6% verbatim overlap while those who hand wrote their notes only had an average of 8.8% verbatim overlap (Mueller & Oppenheimer, 2014). These findings suggest that handwriting notes results in more processing and less passive typing of what the lecturer says, which may have caused students who took handwritten notes to remember more conceptual material (Mueller & Oppenheimer, 2014). Another study found through a meta-analysis of 14 experimental design studies that compared to typing notes, handwriting notes resulted in 9% more college students getting an A or B in their class (Allen et al., 2020). Allen et al. (2020) suggested this difference could be attributed to the fact that handwriting notes is not as distracting as typing notes because applications and notifications are readily available on a computer unlike with pen and paper. However, other studies found that there is no significant difference between typing and writing notes when answering conceptual questions. For example, multiple replications of Oppenheimer and Mueller's experiment (2014) found no significant difference exists between taking handwritten notes and typed notes (Mitchel & Zheng, 2017; Urry et al., 2021). Although there is contradictory research, there is more evidence to support the idea that handwriting notes will have a positive impact on retention compared to typing notes especially since the study conducted by Allen et al. (2020) is a meta-analysis of multiple different experimental studies which all suggest the benefits of handwriting outweigh that of typing.

Taught vs. Not Taught

While it is suggested that note taking is a beneficial practice, most students take bad notes, missing the most important ideas which reduces the effectiveness of such notes (Robin et al., 1977). However, it has been shown that students who are taught how to take notes tend to take better notes containing more of the important ideas, resulting in greater retention of information (Robin et al., 1977). For example, one study taught students how to record and recognize the most important information in a lecture by hosting five practice lectures where they first gave the students an outline of what notes to take and then gradually reduced the outlines until the students were on their own. They measured note-taking quality in critical idea units (ideas that are essential to the lecture), finding that after teaching underachieving college students how to take notes, their notes contained 28.3% more critical idea units than a group that was not taught how to take notes (Robin et al., 1977). Another study found that when one is instructed how to take notes, the usefulness of such notes increases, leading student performance on tests to improve (Chang & Ku, 2014). In the teaching note-taking group for this study, participants were given instructional material which compared good notes, which contained the essential ideas, to bad notes, which lacked those ideas. Those who were instructed on how to take quality notes scored significantly higher on a reading comprehension test, with a mean score of 12.3, than those who were not taught how to take quality notes, with a mean score of 10.8 (Chang & Ku, 2014). Contrary to the aforementioned studies, the current study utilized a different approach, using a short video to teach notetaking rather than physical handouts with follow-up lectures. A video was used to ensure that the "teaching" each participant in the teaching conditions got was identical. Additionally, the rise in technology in classrooms has also led to many lessons being taught via video, with videos becoming regularly used in 79% of classrooms (Schaffhauser, 2019). Therefore, a video was used in this study to mimic the likely method in which note taking will be taught in the future. The current study also was conducted in a high school atmosphere, unlike the two mentioned studies which utilized college and elementary school students.

Hypotheses

In the current study, groups were taught to take notes and follow the outline method. They were instructed to follow this method because it is one of the most popular note-taking strategies in the college and high school atmospheres (GoodNotes, 2018). Additionally, multiple other studies, such as that conducted by Akitunde (2013), utilized outline notes. In this method, information is bulleted with the least general information beginning at the left and more specific information indented to the right (Akitunde, 2013; GoodNotes, 2018). The following hypotheses were tested: *Hypothesis 1:* Those who handwrite their notes will not perform significantly different on factual questions than those who type their notes, but they will both perform significantly better than those who do not take notes. *Hypothesis 2*: Those who are taught how to take notes will do better on factual and conceptual questions than those who are not taught how to take notes

Method

Participants

One hundred fifty-four students were recruited from a suburban high school's 9th grade English classes. Such classes were used as the 9th grade English classes are heterogeneously grouped, providing a mix of students of all academic abilities.

Participants in the study were told that the study would "examine how well they can retain information from a video." They were also notified that they would be taking a test after the video that would measure how well they could remember and apply what was said in the video. To incentivize participation in the study, participants were told that at the end of the study they could enter a raffle to win a $25 Amazon gift card. They entered the raffle by filling out a ticket with their school email address, and the winner was notified.

Experimental Manipulation

This study had two independent variables. The first was whether the student took no notes, handwritten notes, or typed notes. The second was whether or not the student was taught to take notes. Each of the seven classes were randomly assigned to one of five conditions: no notes, handwriting notes and not taught how to take notes, handwriting notes and taught how to take notes, typing notes and not taught how to take notes, and typing notes and taught how to take notes. Three of the seven classes were randomly assigned to a condition that required them to be taught how to take notes. Consequently, they were shown a video that explained how to take outline notes. The video, which was approximately 2.5 minutes long, was titled "Formal Outline" and went through the goal of taking and formatting outline notes and provided a comprehensive and high-quality example of such notes (Ferreira, 2014).

Procedure

To test the effectiveness of the note taking strategies on memory retention, participants watched a video on the Indus River Valley. Prior to watching the video, those assigned to handwriting notes conditions were given a pen and paper to take notes with, those assigned to typing notes conditions were told to take out their school supplied laptop to take notes with, and those assigned to the no notes condition were instructed to completely clear their desk and provide their undivided attention to the video. All participants then watched the video titled "Computing a Rosetta Stone for the Indus Script." This video was approximately 15-minutes in length and was chosen because it is about ancient history, which is loosely related to the 9th grade ancient literature curriculum (Rao, n.d.). This video was also used as it is the same video used in the study conducted by Mueller and Oppenheimer (2014). The video discussed the fact that the language of the Indus River Valley civilization has not been decoded and the different methods in which it may be decoded in the future.

After participants watched the video and took notes, they completed a short answer test based off of the video. They had no time to review their notes and did not use their notes or any other outside information during the test. The test was taken from the study conducted by Mueller and Oppenheimer (2014) and was out of 14 points with 10 points based on factual information and the other four points based on conceptual information (see Appendix). Factual questions were questions that were directly answered by the video while conceptual questions were questions that took the information said in the video and required it to be applied to the questions.

Data Analysis

To investigate the impact of note taking on memory retention, univariate ANOVAs with subsequent post hocs and independent sample t-tests were used. The univariate ANOVA was used to compare the no notes, handwriting, and typing average test scores for both conceptual and factual questions. Independent sample t-tests were used to compare the taught conditions to the not taught conditions. A significance level of $\alpha = .05$ was used, and error bars on the graphs show ± 1 standard error.

Results

Type of Notes

A one-way ANOVA showed that the effect of note type on test score for factual questions is significant, $F(2, 155) = 3.34$, $p = .04$. Subsequent post hocs revealed that those who took handwritten notes performed significantly better on factual questions than those who took typed notes. Students who did not take notes scored an average of 25.3%, and students who took typed notes scored an average of 23.0%, while students who took handwritten notes scored an average of 33.1% (Figure 1). These results do not support hypothesis 1a in that those who took handwritten notes performed significantly better than those who took typed notes.

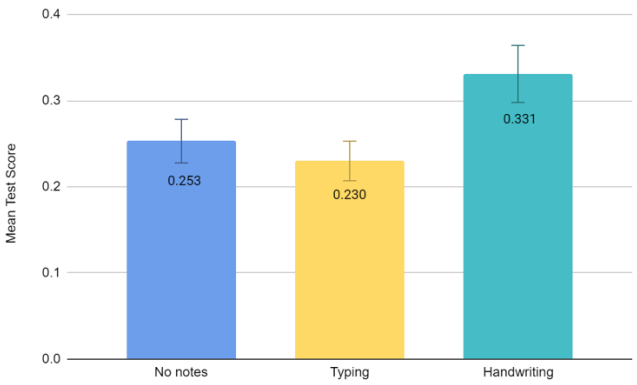


Figure 1. ***The Effect of Type of Notes on Factual Test***

Another one-way ANOVA showed that the effect of note type on test score for conceptual questions is significant, $F(2,155) = 4.34$, $p = .02$. Subsequent post hoc tests showed that those who took handwritten notes performed significantly better than those who took typed notes and those who did not take notes for conceptual questions. Students who did not take notes scored an average of 13.2%, and students who typed notes scored an average of 12.9%, while students who hand wrote notes scored an average of 24.6% (Figure 2). These results substantiated my hypothesis that those who took handwritten notes would perform significantly better than those who took typed notes and no notes for conceptual questions. Yet, my hypothesis was not supported in that those who took typed notes did not perform significantly differently from those who did not take notes.

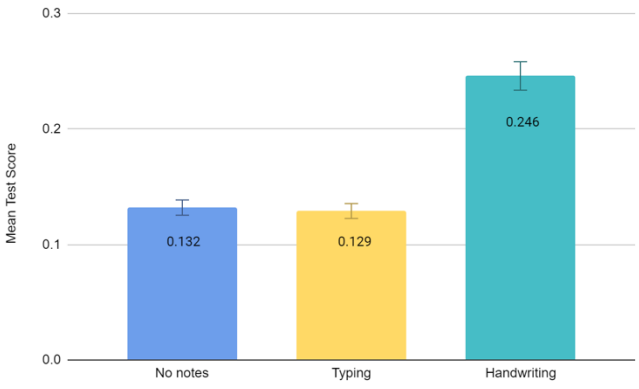


Figure 2. ***The Effect of Type of Notes on Conceptual Test Score***

Taught vs Not Taught

Students who were taught to take notes did no better on the exam questions than students who were not taught to take notes. An independent samples t-test found that being taught to take notes did not significantly improve test performance compared to not being taught to takes notes for factual test questions, $t(153) = 1.48$, $p = .14$, with students who were taught to take notes scoring an average of 23.8% and those who were not taught to take notes scoring an average of 29.5% (Figure 3).

Another independent sample t-test revealed that being taught to take notes did not significantly alter test performance compared to not being taught to take notes for conceptual test questions, $t(153) = -0.76$, $p = .45$. Students who were taught to take notes scored an average of 20.8%, while students who were not taught to take notes scored an average of 25.1% (Figure 4). These results refuted my hypothesis that those who were taught to take notes would perform significantly better than those who were not taught to take notes for both factual and conceptual questions.

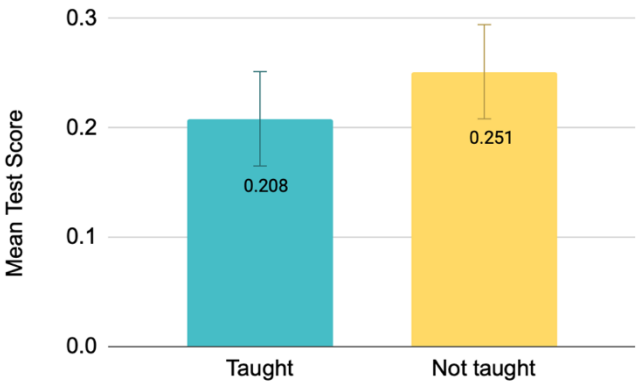


Figure 3. ***The Effect of Teaching Note Taking on Factual Test Score***

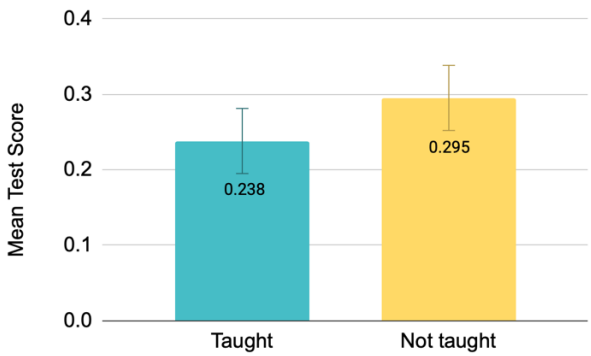


Figure 4. ***The Effect of Teaching Note Taking on Conceptual Test Score***

Discussion

Type of Notes

The data gathered from the study refuted hypothesis 1a in the respect that for factual questions, those who took handwritten notes performed significantly better than those who took typed notes, but neither performed significantly differently from those who did not take notes. The finding was unexpected as previous literature found that those who took handwritten notes and those who took typed notes did not perform significantly differently on a memory retention test for factual questions (Mitchell & Zheng, 2017; Mueller & Oppenheimer, 2014; Urry et al., 2021). Additionally, it was unexpected as previous literature suggested that note taking led to the synthesis and processing of information (Jacobs, 2008), causing those who take notes, both handwritten and typed, to perform better on memory retention tests than those who do not take notes (Akintunde, 2013; Fisher & Harris, 1973).

A possible explanation for the unexpected finding is that during the video of the Indus River Valley Civilization, participants who were typing notes often seemed very distracted, and their eyes rarely left their computer screens. Therefore, participants who took typed notes could have performed worse for factual questions than those who took handwritten notes because they may have been distracted by gaming, messaging, and social media applications on their device, resulting in them being off task and not paying attention to the video. To explore this possibility, further research should be conducted comparing test performance of students typing on devices with internet access to those typing on devices without internet access.

Additionally, taking notes did not significantly alter test performance compared to not taking notes for factual questions which may have been because participants who took notes were fixated on getting the perfect notes and were caught up in the details of the notes causing them to not pay attention to the video on the board. The no notes group, however, was able to provide their undivided attention to the video allowing them to retain the factual information presented in the video (DeWitt, 2007). Therefore, the greater attentiveness of the no-notes group possibly counteracted the processing benefit faced by the note taking groups for such questions.

As expected, those who took handwritten notes performed better than those who took typed notes and those who did not take notes for conceptual questions on the test, thus substantiating hypothesis 1b. Unlike what was expected, students who did not take notes performed no worse than students who typed their notes, refuting hypothesis 1b. It is unsurprising that those who took handwritten notes performed better than those who took typed notes as past literature, such as the study conducted by Jacobs (2008), suggested that those who handwrite their notes cannot keep up with the lecturer, so they are forced to synthesize the information rather than typing it verbatim. The synthesizing of information results in thoughtful processing of it as it forces one to consider how to best condense the information to record it in a timely manner. On the other hand, typing notes verbatim is a thoughtless process which requires simply recording whatever comes out of the mouth of the lecturer. Additionally, those who took typed notes appeared very distracted by the device they were taking notes with which could have resulted in them being off task and not paying attention to the video. Due to the extra synthesis necessary and the fewer distractions available, those who took handwritten notes performed significantly better than those who took typed notes for conceptual questions. It was also foreseeable that the handwriting condition would perform significantly better than the no notes condition as handwriting notes results in the processing and retention of information (Jacobs, 2008).

The finding that typing notes did not improve students' performance as compared to not taking notes suggests that the distractions of using an electronic device may minimize the processing benefit of note taking. While taking notes would be expected to be beneficial as note taking helps retain and process information (Jacobs, 2008),

using devices in the classroom can be very distracting (Awwad et al., 2013). The fact that the devices were distracting likely reduced any processing and retention benefits of note taking as students did not provide their full and undivided attention to the video and task at hand. On the other hand, while the no notes group did not have the processing benefit of note taking, they had the benefit of closely watching the video. As a result, the no notes group performed similarly to the typing group.

Taught vs. Not Taught

The data from the study refuted hypothesis 2 as no significant difference was found between the taught and not taught conditions in terms of test performance for both factual and conceptual questions. The finding was unexpected as previous literature suggested that being taught to take notes would improve one's notes resulting in the processing benefit of note taking being more profound (Chang & Ku, 2014; Robin et al., 1977).

The likely reason for the insignificant difference was that a three-minute video with no practice may be insufficient to teach the process of note taking. Past literature employed interventions to teach the proper taking of notes which took place over many weeks, whereas this study employed a short video to ensure that the same thing was taught to each class in the teaching condition and to reduce the amount of class time that this study interfered with. It is possible that since the participants had no time to practice taking notes using the strategies they were taught, the teaching manipulation was not very effective.

typed notes, but neither performed significantly differently from those

Strengths and Limitations

This study is notable because it addresses the dispute over whether handwriting notes are beneficial for answering conceptual questions compared to typing notes. Studies have employed this same method and found contradictory results with some studies finding that handwriting notes does significantly improve test performance for conceptual questions compared to typing notes (Allen et al., 2020; Mueller & Oppenheimer, 2014), while other studies suggest that there is no significant difference between handwriting and typing notes for such questions (Mitchell & Zheng, 2017; Urry et al., 2021). This study supported literature suggesting that handwriting notes do significantly improve test performance for conceptual questions compared to typing notes. Also, this study disputed the results of previous studies which suggested no significant difference would be found between handwriting and typing notes for factual questions (Allen et al., 2020; Mitchell & Zheng, 2017; Mueller & Oppenheimer, 2014; Urry et al., 2021). Therefore, this study can be used as a basis for future research aiming at whether handwriting notes positively affects one's performance for factual questions.

One limitation of the study was the inability to analyze the notes of each participant for idea units and verbatim overlap. Idea units are considered main ideas in the video, and each participant's notes would be analyzed to determine how many of these units their notes cover. Additionally, it would be interesting to analyze verbatim overlap by looking for any phrases in one's notes which were taken word-for-word from the lecture. Checking for verbatim overlap would be important as by copying the words of the video, it reduces the processing benefit of note taking as such processing is the result of synthesizing material to include on the notes. Analyzing the notes of each participant may have explained why the taught and not taught conditions had, on average, similar test scores as it is possible that the teaching manipulation was not effective which would be evident by an insignificant difference in the average number of idea units and verbatim overlap in the notes in each condition. Additionally, if the notes were analyzed and the typing condition had, on average, significantly more verbatim overlap than that of the handwriting condition it would provide evidence to suggest that typing notes results in more copying word-for-word and, therefore, leads to less processing.

Future Study

In a future study, one could test different note taking techniques. Specifically, one could test Cornell notes, outline notes, and diagram-based notes as those are the most common methods of note taking (GoodNotes, 2018). Identifying the best note taking technique can help provide teachers and students with guidance and advice when deciding how they will teach notes and take notes, respectively.

Another possible future study would be to examine how different teaching styles to teach notes affects how well students take notes. One could test the difference between a lecture, a video, and a handout on how well students take notes as those are the three teaching methods with the most background literature (Chang & Ku, 2014; Robin, et al., 1977). The results from this study would be particularly useful for teachers in advising them on how to most effectively teach their students how to take quality notes.

Lastly, in a future study, one could test typing on devices with and without internet access and see how they differ in terms of average test performance. This comparison would help provide insight as to why typing notes was found to be overall less effective than handwriting notes. If the group without access to the internet, on average, performs significantly better on the memory retention test that the group with access to the internet, then it could be safely concluded that since distracting applications such as games, messaging, and social media are easily accessible on devices where students often type notes, it results in typing notes being less effective than handwriting notes.

Conclusion

The findings suggest that teachers should limit the use of electronic devices in the classroom as they do not help students retain information and probably serve as a distraction. In terms of teaching note taking, it is likely that further research is needed to test different methods of teaching and compare the resulting notes to determine the most effective method. The results also suggest that in some cases involving more factual based information, providing one's undivided attention to the video can be a useful alternative to note taking. However, in terms of conceptual based questions, the results reaffirm the importance of note taking.

References

Akintunde, O.O. (2013). Effects of Cornell, verbatim and outline note-taking strategies on students' retrieval of lecture information in Nigeria. *Journal of Education and Practice, 4*(25), 67–73. https://core.ac.uk/download/pdf/234635029.pdf

Allen, M., Lefebvre, L., Lefebvre, L., & Bourhis, J. (2020). Is the pencil mightier than the keyboard? A meta-analysis comparing the method of notetaking outcomes. *Southern Communication Journal*, *85*(3), 143–154. https://doi.org/10.1080/1041794X.2020.1764613

Awwad, F., Ayesh, A., & Awwad, S. (2013). Are laptops distracting educational tools in classrooms. *Procedia - Social and Behavioral Sciences*, *103*, 154–160. https://doi.org/10.1016/j.sbspro.2013.10.320

Chang, W.-C., & Ku, Y.-M. (2014). The effects of note-taking skills instruction on elementary students' reading. *The Journal of Educational Research*, *108*(4), 278–291. https://doi.org/10.1080/00220671.2014.886175

College Board. (n.d.). *AP program results: Class of 2021*. https://reports.collegeboard.org/ap-program-results/2021

DeWitt, S. (2007). The effects of note taking and mental rehearsal on memory. *Journal of Undergraduate Psychological Research*, *2*, 46–49. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.521.9802&rep=rep1&type=pdf

Ferreira, D. (2014, May 1). *Formal outline* [Video]. YouTube.

https://www.youtube.com/watch?v=CZ57Jh1xdRw

Fisher, J. L., & Harris, M. B. (1973). Effect of note taking and review on recall. *Journal of Educational Psychology*, *65*(3), 321–325. https://doi.org/10.1037/h0035640

GoodNotes. (2018, May 9). *The best note-taking methods for college students*. https://medium.goodnotes.com/the-best-note-taking-methods-for-college-students-451f412e264e

Jacobs, K. (n.d.). *A comparison of two note taking methods in a secondary English classroom*. http://insanitek.net/wp-content/uploads/2017/07/grasp-2008-56.pdf

KIPP NJ. (2016, October 24). Should we still teach kids to take notes by hand? https://blog.kippnj.org/resources/teaching-note-taking/

Mitchell, A., & Zheng, L. (2017). Examining longhand vs. laptop debate: Evidence from a replication. *AMCIS*. https://www.semanticscholar.org/paper/Examining-Longhand-vs.-Laptop-Debate%3A-Evidence-from-Mitchell-Zheng/1897d8c7d013a4b4f715508135e2b8dbce0efdd0

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard. *Psychological Science*, *25*(6), 1159–1168. https://doi.org/10.1177%2F0956797614524581

Rao, R. (n.d.). *A Rosetta Stone for a lost language* [Video]. Ted. https://www.ted.com/talks/rajesh_rao_a_rosetta_stone_for_a_lost_language

Robin, A., Foxx, R. M., Martello, J., & Archable, C. (1977). Teaching note: Taking skills to underachieving college students. *The Journal of Educational Research*, *71*(2), 81–85. https://www.jstor.org/stable/27537081

Singer, F., & Samson, P. (2019, August 7). *The end of note taking?* Inside Higher ED. https://www.insidehighered.com/digital-learning/views/2019/08/07/end-note-taking

Schaffenhauser, D. (2019, July 10). *Video use in education evolving beyond the classroom.* Campus Technology. https://campustechnology.com/articles/2019/07/10/video-use-in-education-evolving-beyond-the-classroom.aspx

Selingo, J. (2022, March 23). *The college-admissions process is completely broken*. The Atlantic. https://www.theatlantic.com/ideas/archive/2022/03/change-college-acceptance-application-process/627581/

Urry, H. L., Crittle, C. S., Floerke, V. A., Leonard, M. Z., Perry, C. S., Akdilek, N., Albert, E. R., Block, A. J., Bollinger, C. A., Bowers, E. M., Brody, R. S., Burk, K. C., Burnstein, A., Chan, A. K., Chan, P. C., Chang, L. J., Chen, E., Chiarawongse, C. P., Chin, G., . . . Duong, V. N. (2021). Don't ditch the laptop just yet: A direct replication of Mueller and Oppenheimer's (2014) study 1 plus mini meta-analyses across similar studies. *Psychological Science*, *32*(3), 326–339. https://doi.org/10.1177%2F0956797620965541

Appendix: *Test*

Factual:

1. Approximately how many years ago did the Indus civilization exist? (1)
2. Many of the remaining examples of the Indus script are small seals that were used for what purpose? (1)
3. The talk mentions three hypotheses about what the Indus script could represent. What are they? (3)
4. Which hypothesis does the speaker support? (1)
5. What are the three problems the speaker mentions in trying to decode the Indus script? (3)
6. What word/sound do researchers think this symbol might represent? (1)

Conceptual:

1. What evidence exists that the Indus script encodes a language? (1)
2. What is significant about the examples of the Indus script found in Mesopotamia? (1)
3. How did they develop and test the computer model they created to help decode the Indus script? (1)
4. Why do researchers think some of the texts represent names of constellations and other heavenly bodies? (1)

A+ Self-Esteem

The Relationship Between Academic Competence as a Contingency of Self Worth, Attitudes Toward College, and Motivation of High School Students

Crystal C. Li, Richard Montgomery High School, Rockville, MD

crystal20058@gmail.com

Abstract

Past research has established that individuals' contingencies of self-worth—domains in which outcomes influence one's sense of self-worth—increase motivation to maximize success in those domains (Crocker & Wolfe, 2001). The present study investigates the relationship between having an academic contingency of self-worth, motivation to succeed in school (intrinsic and extrinsic), college-oriented motivation, and perceived importance of college. A survey was administered to high school students (N = 164), who responded to scale items measuring these constructs. In line with hypotheses, Pearson's correlation analyses revealed that all variables were positively correlated with medium to high effect sizes: students who reported having self-worth contingent on academic competence tended to be more motivated in school (r = .35), as were students who attributed greater importance to college (r = .48). As such, the students who attributed greater importance to college also tended to be those who had higher academic contingencies of self-worth (r = .44). Notably, greater perceived importance of college and higher academic contingencies of self-worth were positively associated with both intrinsic and extrinsic motivation, and there was no significant difference between how strongly these variables associated with each type of motivation. The findings reveal that motivation to work hard in school and to pursue college may stem in part from students' beliefs that they should do well in academics in order to feel worthy.

Literature Review

Problem

The K-12 education system in the US is often criticized for its emphasis on scores and its failure to teach useful and provocative content (The Learning Network, 2019)—flaws which render it incapable of instilling in young students a passion for learning. Indeed, educators are regularly hard-pressed to find methods of incentivizing students to actively engage in lessons, beyond the pressure of exams.

To understand what motivates students in academic settings, it is useful to examine factors that predict high intrinsic academic motivation and high extrinsic motivation, respectively. To that end, the present study investigates the relationship between high school students' level of academic motivation in school (both intrinsic and extrinsic), the extent to which self-worth is contingent on academics, and the extent to which college is perceived to be both important and an object of ambition.

Motivation and Instrumentality Beliefs

For decades, scholars have made a distinction between intrinsic and extrinsic motivation (Deci, 1971; White, 1959). Extrinsic motivation is a drive to engage in a task because of an external incentive associated with it. For example, operant conditioning demonstrates that organisms are more

likely to repeat a voluntary behavior if they are rewarded and less likely to repeat it if they are punished (Skinner, 1953). Crucially, behaviors that are extrinsically regulated are prone to extinction: if a desired behavior is no longer rewarded, it will readily cease.

On the other hand, intrinsic motivation is a desire to adopt a behavior because of its inherent value. An individual who is intrinsically driven to engage in an activity does not require external incentives to do so because the activity is rewarding in its own right (Deci, 1971). As such, intrinsic motivation is often a more sustainable and desirable form of motivation than extrinsic motivation. An extensive body of research has concluded that the introduction of extrinsic rewards diminishes individuals' motivation to engage in activities they already intrinsically enjoy (Deci et al., 1999). If rewards intended to induce extrinsic motivation undermine intrinsic motivation and prevent self-regulation, then it may be logical to conclude that there is an inverse relationship between the two forms of motivation such that the more extrinsically compelled an individual is to engage in a task, the less intrinsic motivation they are to engage in it.

Despite this potential relationship, intrinsic and extrinsic motivation may not always be mutually exclusive. As Miller et al. (1999) have shown, instrumentality beliefs among college students are positively correlated with both kinds of motivation. Broadly, an instrumentality belief is the view that an activity is valuable because it is useful to the individual. In this case, students who believed that tasks in school would prove useful to them in the future reported higher extrinsic *and* intrinsic motivation. However, when this study was replicated using a sample of high-achieving high schoolers, the relationship between instrumentality beliefs and intrinsic motivation was not found (Kover & Worrell, 2010). It is therefore unclear whether students' instrumentality beliefs about school (its utility in helping them obtain desired future outcomes such as admissions to prestigious colleges and respectable careers) represent external and not intrinsic motivators. In continuation of this line of research, one goal of the present study is to compare how concern about college and academic performance relate to intrinsic motivation with how they relate to extrinsic motivation.

Contingencies of Self-Worth

Individuals derive their self-esteem and sense of self-worth from various sources. Crocker et al. (2003) identified seven such "domains of contingency": appearance, approval from others, academic competence, competition, family support, God's love, and virtue. They define contingencies of self-worth as self-imposed conditions that must be fulfilled in order for an individual to feel valuable and worthy. For example, if a person has a need to look attractive in order to feel confident, then "appearance" is said to be one of the individual's contingencies of self-worth.

The researchers also made a distinction between internal contingencies and external contingencies. Much like the two forms of motivation, internal contingencies of self-worth are rooted in one's core identity and decisions. For example, virtue is considered an internal contingency because here, individuals' sense of self-worth is grounded in their own ability to uphold a moral code an internal factor. On the other hand, external contingencies—such as appearance, approval from others, and competition—are external contingencies because they either depend on other people or are superficial components of identity.

Given that an axiomatic characteristic of humans is their desire to protect their self-esteem, scholars have theorized that contingencies of self-worth influence individuals' emotions, cognition, and behavior (Crocker & Wolfe, 2001). If a person's sense of self-worth hinges on a particular area, then they are compelled to increase their success in that area in order to maximize their perceived self-worth. Rieger et al. (2021) demonstrated this theory in the context of eating habits, finding that individuals whose self-worth was contingent on body image were more likely to report symptoms of eating disorders. Tellingly, the symptoms were manifestations of a desire to maintain body image and, by extension, maintain self-worth.

Attitudes Toward College: Its Importance and Motivating Power

In many school districts, especially in the US, high school students are conditioned to become

college graduates: course offerings and school counseling efforts are all geared toward ushering students into college (An, 2020). Thus, it's no surprise that from the viewpoint of many high school students today, college is one of the most critical milestones in their lives. The evidence of this is nearly ubiquitous: many high school students have expressed the feeling that college is the best and only option for them after graduation (Hansen, 2018).

It's clear that college and college admissions have become an exceedingly salient academic motivator for high school students. In this study, I have referred to this mindset as "college-oriented motivation." This describes the extent to which students are motivated by the prospect of college. While I also look at intrinsic and extrinsic moti==vation in the context of school academics, in this study I treat college-oriented motivation as a separate, specific form of motivation that arises from thinking about the need to get into college.

The other facet of college attitudes that I examine in the study is "perceived college importance," which is the extent to which students care about college and the level of importance they attribute to it.

Hypotheses

This study seeks to extend the theory advanced by Crocker and Wolfe (2001) regarding the impact of contingencies of self-worth on motivation applies in the realm of academics. In particular, the study determines whether, as I anticipated, students who have their sense of self-worth highly contingent on their academic success tend to be more motivated in school. Further, I investigate whether an academic contingency of self-worth (CSW) is more highly correlated with intrinsic motivation or with extrinsic motivation. I predicted a higher correlation with extrinsic motivation because working hard in school for the sake of preserving self-esteem represents focus on an *outcome* (i.e., reward, the hallmark of extrinsic motivation), not the inherent value of course content itself. Academic competence might also be considered an external contingency because it relies on external academic results and thus may principally produce extrinsic motivation.

In summary, from speculation and past work on motivation and contingencies of self-worth, the six hypotheses follow. The central focus of the study is on how an academic contingency of self-worth, college attitudes, and motivation all interrelate, and whether college and academic proficiency serve as intrinsic motivators or extrinsic motivators. Based on the background literature, the following hypotheses were tested: *Hypothesis 1:* Holding beliefs that self-worth is contingent on academic competence is positively correlated with academic motivation. *Hypothesis 2:* Academic contingency of self-worth is more strongly correlated with extrinsic motivation than with intrinsic motivation. *Hypothesis 3:* Holding beliefs that self-worth is contingent on academic competence is positively correlated with the extent to which high school students value college. *Hypothesis 4:* The importance students attach to college is negatively correlated with intrinsic motivation and positively correlated with extrinsic motivation. *Hypothesis 5:* College-oriented motivation is positively correlated with in-school motivation. Hypothesis *6:* College-oriented motivation is more strongly correlated with extrinsic motivation than with intrinsic motivation.

Method

Participants

Students at a large public high school in an upper-middle class suburb in Maryland were invited to complete a digital survey. The survey link was shared with psychology and language classes through the online instruction platform Canvas Classroom. Participation was voluntary and all survey responses were anonymous. A total of 164 participants were recruited, of which 33.5% were 12th graders, 40.9% were 11th graders, 19.5% were 10th graders, and 6.1% were 9th graders.

Intrinsic and Extrinsic Motivation

Participants were shown several items from the Adapted Academic Intrinsic Motivation Scale for High School Students (Vo et al., 2021). Using a 9-point Likert scale, they indicated the extent to which they agreed with three statements, such as "I study because I want to learn new concepts and skills." (1 = "strongly disagree", 9 = "strongly agree"). To

measure extrinsic motivation, I created the two items "The main reason I study and complete assignments in school is to receive good grades" and "I prioritize performing well on tests over fully understanding course content." The intrinsic motivation subscale yielded high internal consistency (α = .80, *M* = 6.2, *SD* = 1.5), and the extrinsic motivation subscale had moderate internal consistency (α = 0.63, *M* = 7.3, *SD* = 1.5). An intrinsic motivation composite and an extrinsic motivation composite were created by taking the mean scores of the items in each of the two subscales, and an overall motivation composite was created by computing the mean of the two subscale composites.

Contingencies of Self-Worth

The Academic Competence subscale from the Contingencies of Self-Worth questionnaire (Crocker et al., 2003) was used in this study. The subscale is designed to gauge the extent to which participants feel academics is a domain of contingency of self-worth and it contained three Likert scale items such as "My academic performance influences my self-esteem" (1 = "strongly disagree," 9 = "strongly agree"; α = 0.75, *M* = 7.1, *SD* = 1.6).

Attitudes Toward College

To assess the ambition and importance participants attach to college, I created two scales: the first measures the extent to which participants feel college is important (perceived college importance), and the second measures the extent to which college is a salient motivator in students' lives (college-oriented motivation). The former contained three statements such as "I feel that college is one of the most important milestones in my life" and "College matters a lot to me" (α = 0.89, *M* = 7.5, *SD* = 1.5). The latter contained four statements such as "Thinking about college is what motivates me most" as well as one statement assessing instrumentality beliefs, adapted from items used by Miller et al. (1999): "I work hard in school because doing so will help me achieve my college and career goals." This item was added to the college-oriented motivation composite because it gauges whether students' source of academic motivation arises from the knowledge that school is valuable for achieving college-related goals. This modified scale yielded high internal consistency (α = 0.87, *M* = 6.5, *SD* = 1.7).

Procedures

Prior to completing the survey, students were informed of the purpose of the study and reminded that participation was both voluntary and anonymous. The first section of the survey consisted of the motivation scales and instrumentality belief items. The next section assessed participants' contingencies of self-worth, and the final section asked for their opinions on the importance of college and their college-oriented motivation. After completing the survey, participants were thanked, received two chocolate bars, and had their participant code entered into a raffle to win a $25 Amazon gift card as compensation.

Results

This study set out to examine the correlation between three overarching variables: scores on academic competence as a contingency of self-worth, intrinsic and extrinsic motivation, and beliefs regarding the college. I outline the results of each of the relationships in turn. With the exception of the linear regression model at the end of this section, all analyses were completed using Pearson's correlation.

Academic Competence as a Contingency of Self-Worth and Motivation

There was a significant relationship between the degree to which self-worth is contingent on academics and general motivation in school (Figure 1:$r = .35$, $p < .001$). Correlation analyses run with intrinsic and extrinsic motivation separated as distinct composites revealed that having an academic contingency of self-worth is more strongly correlated with extrinsic motivation ($r = .33$, $p < .001$) than with intrinsic motivation ($r = .21$, $p = .006$). However, a *t*-test comparing these two correlation coefficients found that the difference between the correlations is not significant, $t(161) = 1.22$, $p = .22$, failing to support Hypothesis 1a.

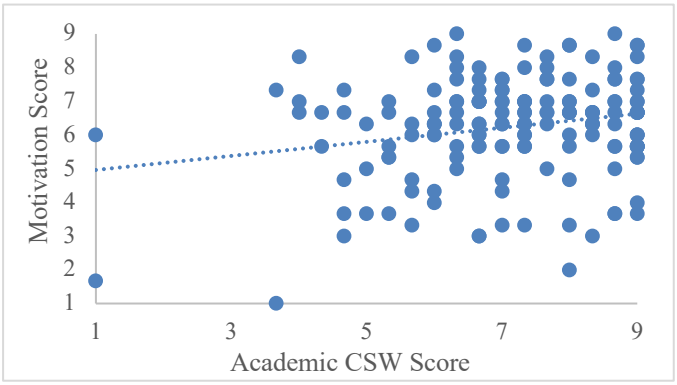


Figure 1. ***Motivation in School vs. Academic Contingency of Self-Worth***

Academic Competence as a Contingency of Self-Worth and Attitudes Toward College

In line with Hypothesis 2, the degree to which college is perceived as important was significantly associated with the degree to which sense of self-worth is contingent on academic competence (r = .44, p < .001), with a medium to large effect size. Thus, the more students derive their self-esteem from academic performance, the more they tend to value college

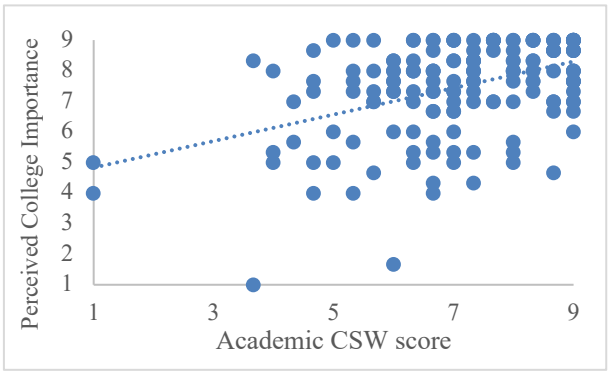


Figure 2. ***Perceived Importance of College vs. Academic CSW***

Furthermore, participants' level of college-oriented motivation—the extent to which college is felt to be a primary source of motivation in their lives—was positively correlated with scores on academic contingency of self-worth (r = .37, p < .001) such that individuals who feel that their self-esteem is dependent on their academics tended to feel that the prospect of college is a powerful motivator

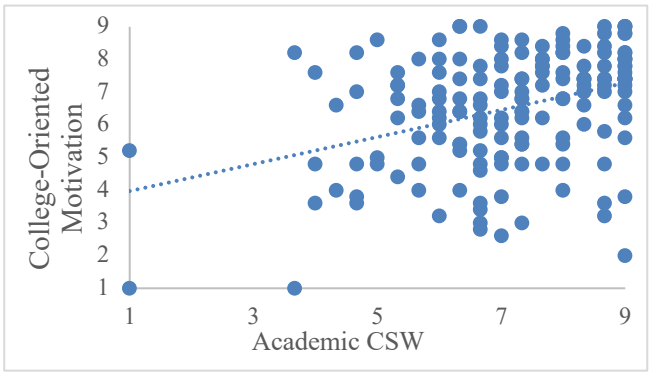


Figure 3. ***College Oriented Motivation vs. Academic CSW***

Attitudes Toward College and Motivation

Hypotheses 3, 4, and 4a dealt with the two scales measuring students' attitudes toward college (perceived college importance and college-oriented motivation). As a preliminary finding, scores on these two scales are significantly and strongly correlated (r = .73, p < .001). As one might expect, the more participants believe that college is important, the more they agreed that college is a predominant source of motivation for them

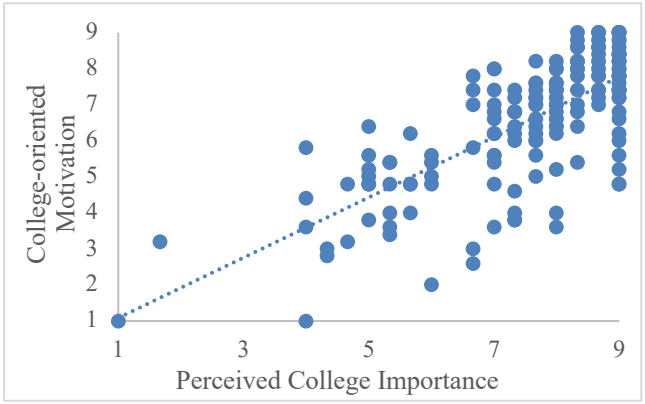


Figure 4. ***College-Oriented Motivation vs. Perceived Importance of College***

There was a significant correlation between perceived college importance and academic motivation in school with a large effect size (r = .48, p < .001). However, the level of importance students attribute to college related positively to scores on *both* extrinsic motivation and intrinsic motivation, thus refuting Hypothesis 3, which assumed that perceived college importance was an extrinsic motivator and therefore predicted that it would relate to reduced intrinsic motivation.

In Hypothesis 4, I predicted that scores on the college-oriented motivation scale would correlate positively with scores on academic motivation in school (the general motivation composite), and this was the case (Figure 5; $r = .45$, $p < .001$).

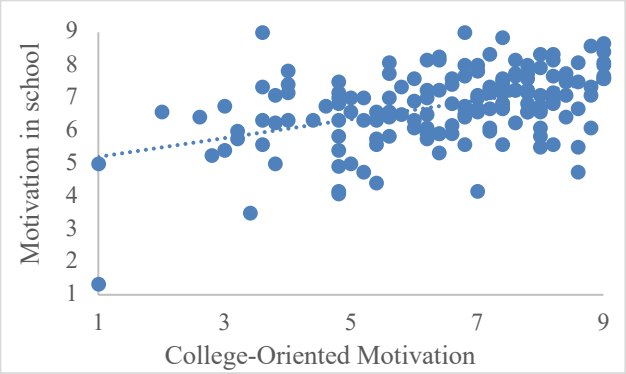


Figure 5. ***School Motivation vs. College-Oriented Motivation***

Finally, contrary to Hypothesis 4a, college-oriented motivation was not more strongly correlated with extrinsic motivation ($r = .29$, $p < .001$) than with intrinsic motivation ($r = .38$, $p < .001$). In fact, there was no significant difference between the two r values, $t(161) = -0.9$, $p = .37$.

Overall Results

After the correlation analyses, I ran a forced-entry multiple regression analysis with academic contingency of self-worth, perceived college importance, and college-oriented motivation as predictors of general motivation in school. The forced-entry method was used (e.g., rather than hierarchical regression) because each of the constructs studied are plausibly equally good predictors of motivation in school; there is no prior literature or theory indicating that any one of the factors studied is a better predictor of motivation in school than the others. The model is a significantly good fit, $F(3) = 19.95$, $p < .001$, and the three variables explain about 27% of the total variation in motivation scores. This was to be expected since the three predictors studied here are far from the only factors that may influence motivation in school. All three predictors make a significant contribution to the model, according to the significant slope estimates.

	b	*SE b*	β
Intercept	3.568	0.433	
Academic CSW	0.114	0.055	0.157*
Perceived Importance of College	0.190	0.076	0.257*
College-Oriented Motivation	0.135	0.064	0.209*

Note. *b* = slope estimate; *SE b* = standard error of B; β = standardized slope estimate *$p < .05$

Figure 6. ***Parameter Estimates for the Linear Regression Model for Motivation Using Academic Contingency of Self-Worth, College Oriented Motivation, and Perceived College Importance as Predictors***

Discussion

Out of the variables examined in this study—academic contingency of self-worth, motivation in school, perceived importance of college, and college-oriented motivation—all were positively correlated with one another, and results yielded medium to large effect sizes. However, no negative correlation between perceived college importance and intrinsic motivation was found, nor a significant difference between the variables' correlation with intrinsic motivation and with extrinsic motivation.

Students who are the most academically motivated are driven by factors rooted in their sense of identity and self-worth; the most highly motivated feel that their academic performance (at least in part) determines how worthy they are. Such a finding confirms Crocker's (2002) theory on the effect of domains of contingency on individuals' motivation to maximize success in those domains.

Furthermore, since participants who scored higher in academic contingency of self-worth tended to attribute more importance to college, students' contingencies of self-worth correlate with and may influence their values and perceptions regarding what is important. Those who perceived college to be more important also tended to be more academically motivated, suggesting that the anticipation of college has potential to incentivize hard work in school.

These findings lend insight into high school students' sources of intrinsic and extrinsic motivation. An academic competence contingency of self-worth and college-oriented motivation each

better predict extrinsic motivation than intrinsic motivation, but to a non-significant extent. This implies that academic competence and college may not be exclusively extrinsic motivators, and I offer potential explanations below:

Academic competence might intrinsically motivate students by presenting them with the opportunity for achieving mastery—not merely desired results and performance. According to the distinction Grant and Dweck (2003) made between learning goals (those aimed at acquiring a skill and becoming genuinely proficient at a task) and performance goals (those focusing on obtaining successful results and minimizing failure), the former is linked with intrinsic motivation while the latter is not. Also, the opportunity to achieve competence in a task is one of the three main facets of self-determination (Deci et al., 2017). Therefore, if students view academic competence as one of their contingencies of self-worth because mastering a skill makes them feel good about themselves, then their domain of contingency would be conducive to intrinsic motivation. However, if students' self-esteem hinges on academic *performance* and not academic *mastery*, then they are more likely to display extrinsic motivation.

Simultaneously, the idea of college might intrinsically motivate students by reminding them that the skills and concepts they learn in school are inherently valuable and necessary (*instrumental*) in their future endeavors.

Thus, although grades, test scores, and college admissions are external factors, they may have the power to both intrinsically *and* extrinsically incentivize students.

Limitations

This study establishes only an association between the measured variables. Due to its correlational design, no conclusions about causality can be drawn; it cannot be inferred that possessing an academic contingency of self-worth increases motivation and subjective importance of college. If causality exists, the correlational nature of the study also prevents conclusions regarding the direction of causality. For example, students' contingencies of self-worth may plausibly be either a cause or effect of high academic motivation. In addition, each of the scales used in the study contained relatively few items and two of the subscales, the academic competence contingency of self-worth subscale and the extrinsic motivation subscale, had only moderate internal consistency, which may have led to less reliable results. Replications of the study would need to aim for higher internal consistency of scales by including more items to measure each construct. Furthermore, the study's sample imposes several limitations. Firstly, participants came from a single high school in a liberal and high-income area, and as such, conclusions drawn from this sample may not apply to high school students in other locations. For example, attitudes regarding college are highly dependent on personal and family conditions, culture, and the environment in which students live. The vast majority of students in this study are set on attending college, but this is unrepresentative of the views of all American high schoolers. Additionally, all participants were recruited from psychology and language classes, (with most being from psychology classes), making the sample further unrepresentative.

Future Studies

Since all variables studied here were found to be positively correlated, more work is needed to disentangle which variables cause changes in others. In order to determine causality, future experimental studies may prime participants to think about college and determine the extent to which doing so increases motivation and competitiveness among high schoolers. Further research is also needed to determine whether and how individuals' contingencies of self-worth can be manipulated such that their dependence on academics as a determinant of self-worth can be increased or decreased. This would allow for research on whether and how increased or decreased emphasis on an academic domain of contingency can affect the perceived importance of college.

Conclusion

The findings from this research have evidenced a positive relationship between having a

contingency of self-worth in academic competence, perceiving college to be important, and motivation—intrinsic and extrinsic; college-oriented and in-school. On average, individuals who depend more heavily on academic achievement to maintain their self-esteem attribute more importance to college and work harder in school. Notably, while greater emphasis on academic competence correlates more strongly with extrinsic than with intrinsic motivation, the difference is not significant. Similarly, neither intrinsic nor extrinsic motivation is more strongly correlated with college-oriented motivation. Thus, this study has presented the possibility that the prospect of college and desire for academic competence can represent both intrinsic and extrinsic motivators.

Nevertheless, students and educators should be wary of motivation exclusively rooted in college, as it is not sustainable. High schoolers who are motivated solely by college may struggle to find the drive to continue learning and growing after the college admissions process is over. Being motivated primarily by one's need for academic competence in order to maintain a sense of self-worth should also be avoided as it may leave individuals especially susceptible to self-doubt, negative self-perceptions, and high stress resulting from instances of poor performance in school. Instead, to foster in students a genuine interest in learning, educators can highlight the utility of the content taught and the inherent enjoyment of mastering course concepts.

References

An, S. (2020, February 3). *Pushing beyond the idea that college is always the answer*. National Public Radio. https://www.npr.org/local/309/2020/02/03/802228885/pushing-beyond-the-idea-that-college-is-always-the-answer

Crocker, J. (2002). Contingencies of self-worth: Implications for self-regulation and psychological vulnerability. *Self and Identity*, *1*(2), 143–149. https://doi.org/10.1080/152988602317319320

Crocker, J., Luhtanen, R. K., Cooper, M. L., & Bouvrette, A. (2003). *Contingencies of Self-Worth Scale* [Database record]. APA PsycTests. https://doi.org/10.1037/t00082-000"

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108*(3), 593–623. https://doi.org/10.1037/0033-295X.108.3.593

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology, 18*(1), 105–115. https://psycnet.apa.org/doi/10.1037/h0030644

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627–668. https://doi.apa.org/doi/10.1037/0033-2909.125.6.627

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, *85*(3), 541–553. https://doi.apa.org/doi/10.1037/0022-3514.85.3.541

Hansen, M. (2018, December 7). *Showing students college isn't the only option for success.* Illinois Public Media. https://will.illinois.edu/news/story/showing-students-college-isnt-the-only-option-for-success

Kover, D. J., & Worrell, F. C. (2010). The influence of instrumentality beliefs on intrinsic motivation: A study of high-achieving adolescents. *Journal of Advanced Academics*, *21*(3), 470–498. https://doi.org/10.1177%2F1932202X1002100305

Miller, R. B., DeBacker, T. K., & Greene, B. A. (1999). Perceived instrumentality and academics: The link to task valuing. *Journal of Instructional Psychology*, *26*(4), 250–260.

Rieger, E., Prasetya, K., Christensen, B. K., & Shou, Y. (2021). Identifying the contingencies of self-worth associated with eating disorder symptoms: The use of choice-based conjoint analysis. *International Journal of Eating Disorders*, *54*(12), 2167–2179. https://doi.org/10.1002/eat.23634

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.

Skinner, B. F. (1953). *Science and human behavior*. Macmillan.

The Learning Network. (2019, December 19). What students are saying about how to improve American education. *The New York Times.* https://www.nytimes.com/2019/12/19/learning/what-students-are-saying-about-how-to-improve-american-education.html

Vo, T. T., Ullrich-French, S., & French, B. F. (2021). *Adapted Academic Intrinsic Motivation Scale for High School Students* [Database record]. PsycTests.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297–333. https://psycnet.apa.org/doi/10.1037/h0040934

Appendix: *Survey Key*

1 = “strongly disagree” 5 = “neutral” 9 = “strongly agree”

Intrinsic and extrinsic motivation

1. The main reason I study and complete assignments in school is to receive good grades.
2. I study because I want to learn new concepts and skills.
3. I find the topics introduced in my courses to be interesting.
4. I prioritize performing well on tests over fully understanding course content.
5. I work hard in school in order to improve myself.

Academic competence as a contingency of self-worth

1. My academic performance infl puences my self-esteem.
2. I feel better about myself when I know I'm doing well in school.
3. Receiving poor test scores and grades hurts my self-esteem.

Attitudes toward college

Perceived college importance subscale

1. I feel that college is one of the most important milestones in my life.
2. College matters a lot to me.
3. It is important to me that I get into a good college.

College-oriented motivation subscale

1. Thinking about college is what motivates me most.
2. I take challenging classes in order to show colleges that I am a rigorous student.

3. I participate in various activities outside of school in order to impress colleges.
4. I put effort into academics because I want colleges to see that I am a good student.
5. I work hard in school because doing so will help me achieve my college and career goals.

JOURNAL OF SECONDARY PSYCHOLOGICAL STUDIES

Submissions: All submissions to the Journal of Secondary Psychological Studies must be emailed to jofsps@gmail.com. Please visit our website https://jofsps.weebly.com for more information regarding our requirements and formatting preferences for submissions.

Feedback: We would love to hear any feedback on our journal. Please email jofsps@gmail.com to help us improve our publication.

JOFSPS