

**Investigating the Efficacy and Mechanisms of a Middle Warning in Online Cheating
Behaviors**

by

Alissa M. Zawacki

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
May 2, 2020

Keywords: personnel selection, cheating,
middle-warning, online environment

Copyright 2020 by Alissa M. Zawacki

Approved by

Jinyan Fan, Chair, Professor of Psychology
Ana Franco-Watkins, Professor of Psychology
Jesse Michel, Associate Professor of Psychology

Abstract

Although online selection tools provide several benefits to both organizations and applicants, one large concern of such tools is the ability for applicants to potentially cheat. Thus, it is essential that a strong method be identified for reducing these unethical behaviors. Warning messages have shown to be effective at reducing these behaviors, but one concern is that their effectiveness may wear off over time or applicants may become habituated to the instructions/warnings given at the beginning of tests. The current study tested the efficacy of a warning message delivered in the *middle* of an online knowledge test in reducing cheating behaviors, as well as examining cognitive (moral disengagement) and emotional (fear and guilt) mechanisms that may be responsible for the above effects. Results showed that while neither cognitive or emotional mechanisms were found to significantly mediate middle-warning messages and cheating behaviors, results did show that a middle-warning message can significantly decrease cheating behaviors. Additionally, it was found that cheater status moderates the relationship between warning messages and knowledge test scores, such that cheaters who do not receive a warning message have the highest scores overall, and cheaters who receive the warning message have the lowest scores overall.

Table of Contents

Abstract	ii
List of Tables	v
List of Figures	vi
Introduction.....	1
Literature Review	2
Online Selection Tools.....	2
Identifying and Deterring Cheating in Online Tests.....	4
Current Study	6
Cognitive Mechanisms (Moral Disengagement)	7
Affective Mechanisms (Negative Emotions)	9
Method	13
Participants.....	13
Procedure	13
Measures	14
Analytic Strategy	15
Results.....	16
Hypothesis Testing	18
Discussion.....	24
Limitations and Directions for Future Research	27
Practical Implications	29
References	31
Appendix 1 (Warning and Control Messages).....	36

Appendix 2 (Measures)..... 37

List of Tables

Table 1 17

Table 2 21

List of Figures

Figure 1	12
Figure 2	19
Figure 3	20
Figure 4	24

Introduction

As technology has continued to evolve, so has the usage of internet-based selection tools. Companies are moving paper-pencil testing to a computer-based system in order to reap a variety of benefits (Wiechmann & Ryan, 2003). Although these online tools provide benefits to both organizations and applicants, one large concern of administering online-based tests is the unproctored environment in which they take place (Tippins, Beaty, Drasgow, Gibson, Pearlman, Segall, & Shepherd, 2006). Being that the environment is both unproctored and a high-stakes situation, due to the nature of the selection context, there is an increased chance for cheating to occur (Pearlman, 2009). While there have been some attempts to mitigate this problem, there are still concerns with the current methods and a lack of understanding as to why certain methods may be more effective than others. One promising method that researchers have identified is the usage of a warning message to deter cheating. While studies have shown positive results of this method, one large concern of pre-warning messages is that their effectiveness may wear off over time or that applicants may become habituated to the instructions/warnings that are oftentimes given at the beginning of online tests (Corrigan-Gibbs, Gupta, Northcutt, Cutrell, Thies, 2015). Thus, additional approaches and mechanisms need to be examined in order to determine the most effective deterrent of online cheating behaviors.

The current study looks to investigate two crucial research questions regarding online selection tools and cheating behaviors that may occur in online environments. The first research question regards the efficacy of a warning message delivered in the *middle* of the test in reducing cheating behaviors. The second research question concerns cognitive (moral disengagement) and emotional (fear and guilt) mechanisms that may be responsible for the above effects. These questions will be answered utilizing results from an online knowledge test taken

by an applicant pool from a large university in China. To begin, a discussion of online selection tools and various anti-cheating methods that are currently being utilized will be examined. Next, cognitive and emotional mechanisms responsible for the effects of middle warnings on cheating behaviors and test scores are elaborated on, leading to a proposed moderated mediation model.

Literature Review

Online Selection Tools

Although there are a wide variety of selection tools utilized in the workplace, some are more prevalent than others in terms of usage online. Assessments such as personality tests, integrity tests, cognitive ability tests, knowledge tests, and emotional intelligence tests are some of the more prominent ones utilized by organizations. Some of these tests are susceptible to faking, such as personality tests or integrity tests (Fan, Gao, Carroll, Lopez, Tian, & Meng, 2012), while others are more susceptible to cheating, such as cognitive ability tests and knowledge tests, where test items have clear correct or incorrect answers. Cheating has previously been defined as, “a rule-breaking behavior that is exhibited with the intention of gaining an unfair advantage over a party or parties with whom the cheater is associated through a norm-governed relationship” (Bing, Davison, Vitell, Ammeter, Garner, & Novicevic, 2012, p. 29). This study in particular focuses on the usage of a general knowledge test in the selection context. Knowledge tests have been a widely utilized tool in selecting applicants, as they have continually found to be predictive of future performance (Kuncel, Hezlett, & Ones, 2001).

In comparison with traditional paper-pencil tests, organizations may choose to use internet-based testing for a variety of reasons. Not only does online testing reduce costs, it also increases the speed of assessment. Additionally, test instructions and timing are able to be consistent across applicants, and the computers are able to produce scores and results in a timely

manner. Other advantages also include reaching a wider applicant pool, as companies are able to recruit individuals without needing them to travel all the way to the company or a testing site. Applicants are also able to take the test at a convenient time for themselves, which might not be during regular business hours, as would be needed with traditional paper-pencil testing. Lastly, there are administrative advantages to online testing as companies are able to quickly alter the tests being given and make changes to questions or cut-off scores, without needing to print and redistribute an entirely new test (Tippins et al., 2006).

Although there may be several benefits for employers choosing to use internet-based tests, specifically for administering knowledge tests, there are also a number of disadvantages. Some of these disadvantages include a potential for equipment or software malfunctioning, lack of accommodations that might be feasible at a testing center, and the negative reactions from applicants that might come along with such an impersonal environment. One other major disadvantage is the potential for cheating to occur (Tippins et al., 2006). Although studies have begun to look at the effects of cheating in unproctored vs. proctored environments, these studies have yielded mixed results (Do, 2009). Some believe that unproctored internet testing (UIT) should not be used in a selection context due to the nature of the situation. The selection context can be considered a high-stake testing environment in which judgements or decisions with high consequences are made regarding the individual being tested (Pearlman, 2009), and as found in Gneezy (2005), individuals lie or cheat more when the stakes and expected benefits are high. Additionally, evidence has been found that cheating is more likely to occur in certain Asian countries such as Japan (e.g., Diekhoff, LaBeff, Shinohara, & Yasukawa, 1999) and China (e.g. Bernardi, Baca, Landers, & Witek, 2008), where social pressures and expectations around

cheating vary in comparison to Western countries. Therefore, our applicant pool from China is at an even higher risk to exhibit cheating behaviors.

Identifying and Deterring Cheating in Online Tests

Various methods have been used in the past to identify cheating in an online environment. For instance, scholars have examined mean differences in test scores across high-stakes and low-stakes settings utilizing either a between-subjects design (e.g., Harmon & Lambrinos, 2008) or a within-subjects design (e.g., Kantrowitz & Dainis, 2014). Substantially higher mean scores in the high-stakes settings are taken as indirect evidence that cheating has occurred. A recent meta-analysis also determined that there were statistically significant differences between proctored and unproctored environment test scores, whereas scores in the unproctored environment tended to be higher (Steger, Schroeders, & Gnambs, 2018), again indicating that cheating may be more likely to occur in an online unproctored environment.

An additional method used to identify cheating is keystroke analytics. Keystroke analytics is defined as, “an online biometric authentication system that uses a test taker’s pattern of typing a specific 15-character phrase to uniquely authentic him or her in order to launch a high-stakes test” (Foster, 2009, p. 31). This pattern is one that is initially captured when the applicant first enrolls to take the test. The enrollee will be prompted to type out a 15-character phrase 12-15 times. This is then matched up with the phrase being typed before the applicant launches the exam to determine if cheating by means of a fake test taker is occurring.

One of the more creative approaches to identify cheating was done using the examination of a free-response question and a “honey pot” website. For the free-response question participants were asked to answer the question by typing in a response, most of which were a few sentences long. Once tests were complete, these answers were then re-copied into an internet

browser and responses that matched the text from a return page were labeled as cheaters. Secondly, they also compared participants answers to each other's to determine if students had shared answers. As for the honey pot website, prior to the start of the exam, the researchers posted all of the exam questions to a public-facing website that was indexed by Google. If a participant searched for an exact exam question or part of the question, the "honey pot" website would be the first return on Google. The answers were not provided; however, participants could click a "show answer" button that, without them knowing, would flag that individual participant as a cheater (Corrigan-Gibbs et al., 2015). Although these methods may be more robust than others, they do require several technical aspects that might be difficult to achieve in a way that could be generalizable to other online tests (i.e. creating a honey pot website for each new test).

There have been some attempts to mitigate or deter the potential cheating behaviors that may occur in an online internet-based test. Previous methods include a speeded timed test (Arthur Jr., Glaze, Villado, & Taylor, 2010), video monitoring (e.g., Karim, Kaminsky, & Behrend, 2014), randomizing test items, selecting from a large pool of questions, and constructing tests from Item Response Theory item pools (e.g., Burke, 2009). One last method that has been used to reduce cheating is the implementation of a pre-warning. One study found that in two different experiments, the usage of a stern warning, in particular one that warns of legitimate, negative outcomes if caught cheating, was more effective at reducing cheating behaviors than the usage of an honor code (Corrigan-Gibbs et al., 2015). A second study found similar results. Participants were placed into four groups using a 2 (honor code reminder present or honor code reminder not present) x 2 (realistic course warning present or realistic course warning not present) design. The authors found that those who received no reminder of the

honor code and no realistic course warning cheated the most, whereas individuals who received both a reminder of the honor code and a realistic course warning cheated the least amount overall (Bing et al., 2012).

Although these studies show the strong impact that a warning can have, both studies' warnings are utilized at the onset of the exam or course. One limitation of these studies thus, is the effect of time on the warning. Some fear that the usage of a pre-warning may wear off over time or that participants may become habituated to the instructions/warnings that are usually present at the beginning of exams. Corrigan-Gibbs et al. (2015) calls for future research to study the effects of how online warnings change over time. In the current study, a warning message is implemented in the *middle* of an online knowledge test (i.e., middle-warnings) to examine whether and how middle-warnings reduce cheating behaviors and affect test scores.

Current Study

The usage of middle warnings in a selection context is not an entirely new idea. This method has been utilized previously, not with cheating behaviors, but rather with faking behaviors in personality tests. It has been reported that middle warnings are effective in reducing faking on personality tests, although they may also lead to under-reporting personality scores (e.g., Ellingson et al., 2012; Fan et al., 2012). The reason for under-reporting is that there are no right or wrong answers on a personality test, thus there is room for the scores to dramatically decrease after a warning is received. Interestingly, this potential concern does not apply to cheating on knowledge tests, because there are correct answers in these tests and there is no room to further lower knowledge test scores. However, if test scores are highly inflated from the average due to cheating behaviors or the unproctored environment, scores may in fact decrease after a warning is received. Previous research has found that unproctored environments lead to

higher test scores, which has been attributed to potential cheating behaviors (e.g., Harmon & Lambrinos, 2008; Steger et al., 2018). As noted earlier, warning messages can help decrease these cheating behaviors (Bing et al., 2012), especially when they contain a negative consequence component, as exhibited in the current study's warning message (see Appendix A).

Thus,

Hypothesis 1a: Applicants receiving the middle-warning message will display fewer cheating behaviors than those receiving the control message.

Hypothesis 1b: Applicants receiving the middle-warning message will have lower knowledge test scores than those receiving the control message.

Drawing on theories from moral psychology, we propose that there are two types of mechanisms that may explain how and why middle warnings reduce cheating in online knowledge tests. These include cognitive mechanisms (moral disengagement) as well as emotional mechanisms (fear and guilt). Figure 1 depicts our conceptual model.

Cognitive Mechanisms (Moral Disengagement)

The theory of moral disengagement stems from the social cognitive theory and self-regulatory systems. Bandura (1986) claims that when unethical situations arise, individuals' moral standards and self-regulatory mechanisms prevent the individual from engaging in dishonest behaviors. These self-regulatory mechanisms include self-monitoring, judgmental functions, and self-reactive functions. Thus, individuals will judge situations they are in based off of their internal standards and what potential negative or positive self-reactions would occur if those standards were violated or upheld (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996).

These self-regulatory mechanisms usually lead individuals to maintain their moral standards, however these functions only operate when they are activated. Thus, it is possible that

certain situations that may result in large personal gains, such as a high-stakes selection context, could lead someone to disengage from these morally self-regulating systems. Bandura and colleagues (1996) identified ways that individuals may engage in moral disengagement, in order to act unethically but not feel the guilt or shame that would usually come along with breaking one's moral standards. These include things such as justifying one's actions in comparison to others, focusing on the ultimate gains, or minimizing the perceived consequences of their actions (Kish-Gephart, Detert, Trevino, Baker, & Martin, 2014). For example, one previous study found that when students perceived their peers to be cheating more frequently, they also engaged in further cheating behavior, thus attempting to further justify their behaviors (Farnese, Tramontano, Fida, & Paciello, 2011). Another study reported a strong link between personal gain and moral disengagement, such that when personal gain was high, moral disengagement was also high (Kish-Gephart, et al., 2014). The selection context can be thought of as one of the largest places for personal gain, and thus an increase in moral disengagement may be likely. These morally disengaging behaviors are likely to occur unless the individuals are given a reason to reexamine their unethical behaviors.

Previous work has shown that reminding individuals of their moral obligations, by having them sign an honor code (e.g., Mazar, Amir, & Ariely, 2008) or sending them a stern warning at the beginning of an online test (e.g., Corrigan-Gibbs et al., 2015), can prevent or reduce moral disengagement and cheating from occurring. If an individual is given no reminder of their moral obligation, they will have an easier time morally disengaging, and thus an increased chance of cheating on the exam. However, if an individual receives a middle-warning message, they are more likely to be reminded of their moral standards and have a harder time morally disengaging through cheating behaviors. Thus, we hypothesize:

Hypothesis 2: Moral disengagement will mediate the effect of a middle-warning message on reducing cheating behaviors and knowledge test scores.

Affective Mechanisms (Negative Emotions)

Another factor that may contribute to the effectiveness of warning messages is emotion. Several theories have examined the ways in which certain stimuli/situations lead to emotional reactions which then further play a role in deciding whether individuals act ethically or unethically (e.g. Jones, 1991; Harrington, 1997; Gaudine & Thorne, 2001). Researchers in moral psychology have often focused on “moral emotions”, which typically include guilt and shame. However, more recently researchers have turned their focus to non-moral emotions that might also play a role in moral and ethical decision making (Yacout & Vitell, 2016). These include emotions such as fear, anger, or happiness. Individuals will have different emotional reactions to middle warnings depending on whether or not they are acting unethically. These subsequent emotions will change applicants’ behaviors, but again only to the extent that they are actually cheating. Due to the negative consequences that can occur from unethical behaviors, thinking about or being reminded of these consequences may elicit negative emotions such as fear, guilt, or shame (Connelly, Helton-Fauth, & Mumford, 2004).

Although research on moral emotions has largely focused on shame and guilt, in the current study we will focus on the role of guilt in mediating behavioral change. This is because, although shame can lead an individual to focus inward on one’s identity and flaws, guilt can lead an individual to focus outward on the behavior that aroused the emotion and focus is given to steps that can be taken to fix that behavior (Eisenberg, 2000). As for negative non-moral emotions, fear and anger have both been found to impact behavior. One study in particular found that anger led participants to report a greater desire for retaliation and lower levels of

ethical decision making, whereas fear led to higher ethical decision making, above and beyond the control and anger group (Kligyte, Connelly, Thiel, & Devenport, 2013). Although it is reasonable to assume participants in the current study may report anger in response to receiving the warning message, due to the nature of the situation, it is unlikely that anger will further impact their behavior, as retaliation in this case would only result in lower test scores. Fear on the other hand may lead to positive behavioral change. Thus, below we argue that fear and guilt are the emotions most likely to play a role in mediating warning messages and the outcomes of cheating behaviors and knowledge test scores.

Guilt. Guilt can arise when individuals feel they have violated a moral, ethical, or cultural standard (Ausebel, 1955; Izard, 1977; Lazarus, 1991; Tangney, 1999). If individuals receive a message that leads them to believe they have violated these standards, especially ones that they had previously been informed of, feelings of guilt can grow. Once an individual feels guilt, it can cause them to fixate on the situation/behavior that caused the guilt in the first place. This can then further motivate them to change their behaviors in hopes to make amends for the dishonest behavior (Tangney, 1999). Thus, if an individual is cheating and receives a warning message, feelings of guilt will likely ensue which can lead to a decrease in the unethical cheating behaviors. Thus, we propose:

Hypothesis 3: Guilt will mediate the effect of middle warnings on reducing cheating behaviors and knowledge test scores.

Fear. Fear can be a driving force in behavioral change. It has been found to positively relate to ethical intentions (Yacout & Vitell, 2016) as well as ethical decision making (Kligyte, et al., 2013). If an individual feels they have a reason to be fearful, their response will be to escape or avoid the threatening agent (Nabi, 1999). Test takers who receive middle warnings may feel

fearful and attempt to avoid punishment by reducing cheating behaviors on the subsequent part of the test. Previous researchers have also found that fear of punishment has a strong effect on cheating behaviors, such that individuals who exhibit cheating behaviors report fear of punishment as the number one deterrent of cheating (Diekhoff et al., 1999). This fear plays a role, especially when the threat of punishment, as exhibited in our warning message, provides a realistic negative consequence (Akeley & Miller, 2012). Thus, we propose:

Hypothesis 4: Fear will mediate the effect of middle warnings on reducing cheating behaviors and knowledge test scores.

Cheaters and non-cheaters are expected to react differently to receiving a warning message. While, a warning message may change behaviors for those who are cheating, the same is unlikely to occur for those who are not cheating. As stated previously, knowledge test scores can only go down if they were incorrectly inflated in the first place, and we would only expect to see a decrease in cheating behaviors, if those behaviors were being displayed prior to the warning message. Thus, a decrease in knowledge test scores and cheating behaviors are only likely to occur if the individual was cheating in the first place. Thus, we propose:

Hypothesis 5a-b: Cheater status will moderate the relationship between the warning condition and (a) cheating behaviors and (b) knowledge test scores.

Receiving a warning message may also impact both cognitive and emotional mechanisms differently depending on whether or not the individual was cheating in the first place. Cheaters who receive the warning message, will have a harder time morally disengaging, as they will be reminded of their ethical standards (Bandura et al., 1996), whereas cheaters who receive the control message will have an easier time displacing those standards and thus morally disengage. It is possible that both cheaters and non-cheaters will experience fear if they receive the warning

message, due to a credible threat of punishment. However, those who are cheating are even more so likely to feel as though they have a reason to be fearful due to their prior unethical behavior and may report higher levels of fear (Nabi, 1999). Unlike fear, guilt usually only occurs when an individual feels they have violated a moral standard (Tangney, 1999), thus non-cheaters who receive a warning message will most likely not experience as high of levels of guilt in comparison to those who believe they have been caught acting unethically.

Hypothesis 6a-c: Cheater status will moderate the relationship between the warning message condition and (a) moral disengagement, (b) guilt, and (c) fear.

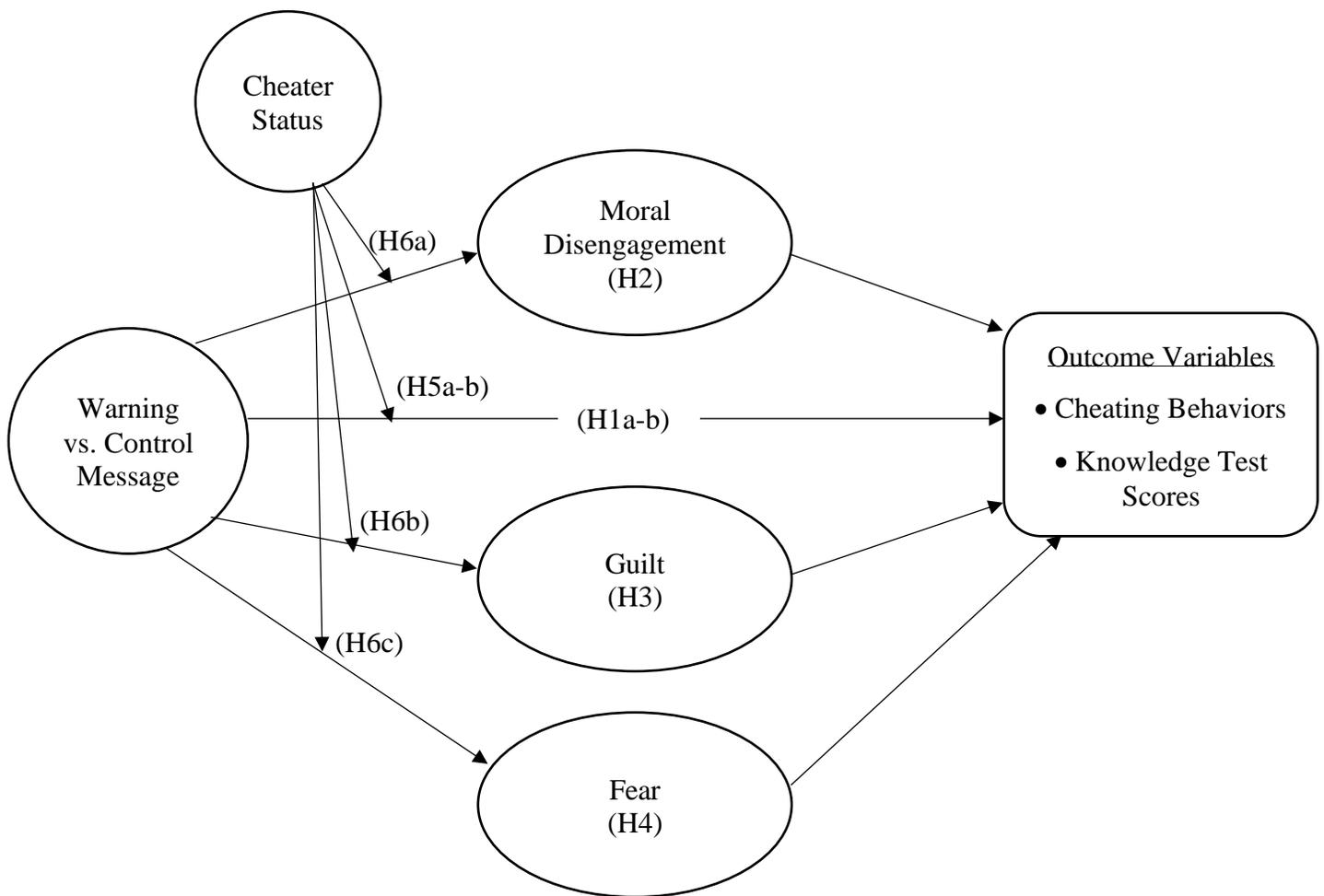


Figure 1. Proposed conceptual model

Method

Participants

Participants were 1,077 undergraduate students who were enrolled at a large university in China. They were applying for a very prestigious two-year academic honors program. Applicants were required to complete the online knowledge test as part of their initial screening for the application process. Upon completion of the test they were given a monetary compensation for their participation. Of those who took the test, 79.7% were female, with an average age of 20.49 years.

Procedure

Upon initial application, candidates received an email with information regarding the required online knowledge test as well as a personal link to the testing site. The email also included two reminders, the first being that during the test applicants were to be recorded via video-monitoring and the second is that they were not allowed to use any external resources while completing the test. The knowledge test contained 90 questions in total, divided into three equal parts (30 items in each part), with items from various general subjects such as science, social studies, and liberal arts. For the purpose of this study, only responses from the first two sections were examined. Applicants had 20 minutes to complete each section of the test, for a total testing time of 60 minutes. Applicants were randomly assigned to either a control condition or a warning condition. Applicants in the control condition received a control message at the end of the first section (after Question No.30). Applicants in the warning condition received a warning message also after Question No.30, indicating that the system has identified unusual answering behaviors and that their test had been flagged. Throughout the test, participants were also asked to complete emotion surveys as well as a moral disengagement survey.

Two methods were used to identify cheating behavior, our main outcome variable. The first method entailed the applicant's computer screen automatically maximizing at the start of the test (similar to a lockdown browser, except applicants were able to exit if they attempted to). With this maximization, the system was then able to record the number of times an applicants exited the testing screen, which was counted as an attempt to use external resources (e.g. an additional internet browsing tab). The second method was video recording. This particular video recording system captured images (photos) from the applicant's webcam every 5 seconds. Coders were trained to identify what appears to be cheating behaviors (extra people in image, continually looking in different directions, holding external resources, etc.). Applicants were categorized as a cheater according to one of two criteria: (a) if an applicant is caught exiting the testing screen one or more times for 5 or more seconds or (b) if an applicant is flagged by coders for two or more cheating behaviors in the video images.

Measures

Fear and Guilt. Applicant emotions were measured using sub-scales from the original PANAS-X Scale (Watson & Clark, 1994). Six items were used to measure fear, and six items were used to measure guilt. Applicants rated how each given emotion represented how they were currently feeling, sample items for fear are, "afraid" or "scared". Alpha for the fear scale was good at $\alpha = .95$. Sample items for guilt are, "ashamed" or "dissatisfied with self". Alpha for the guilt scale was also good at $\alpha = .95$. Responses were scored on a 5-point Likert scale, with 1 being *Very slightly or not at all*, and 5 being *Extremely* (See Appendix B). The emotions were measured at two different time points. Time 1 was 15 questions after the start of the test in section one, and Time 2 was 15 questions after receiving the warning/control message in section two.

Moral Disengagement. Applicant moral disengagement was measured using a 6-item scale (Shu, Gino, & Bazerman, 2011). Applicants were asked to respond to how they felt during the test. A sample item is, “I felt that cheating was an appropriate behavior because no one got hurt.” Alpha for the moral disengagement scale was low at $\alpha = .56$. Applicants were asked to respond on a 5-point Likert scale, with 1 being *Strong Disagree* and 5 being *Strongly Agree*. This scale was recoded, as done in the original article, with 1 being recoded to -2, 5 being recoded to 2, and 2, 3, and 4 being recoded to zero (See Appendix B). Due to the directness of this measure in asking about cheating behaviors, and so not to raise suspicion, this measure was collected at the end of the knowledge test.

General Knowledge Test. This study focused on the first two sections, or 60 items, from the 90-item general knowledge test. The test was created based on a large general knowledge test question pool in China, with questions being pulled from varying levels of difficulty. The same 30 questions were utilized in each section, however their order was randomly assigned, to ensure similar levels of difficulty for all participants in each section while still maintaining some test security. Items on the test included 15 questions from six different subjects including natural sciences, social sciences, literature and art, computer science, and sports. Participants scores from the first section represent Time 1 knowledge test scores and their scores for the second section represent Time 2.

Analytic Strategy

The predictor variable of condition was split into control group and warning group, with “0” representing the control message and “1” representing the warning message. The moderator variable of cheater status was defined as “0” representing non-cheaters, and “1” representing cheaters. The mediators of guilt and fear, as well as both outcome variables, knowledge test

scores and cheating behaviors, were measured at Time 1 and Time 2. Therefore, to observe the changes that occurred both over time and between groups, Hypotheses 1b, 6a, 6b, and 6c were first examined via a two-way mixed ANOVA. Simple effects from the significant interactions were further examined via ANOVA. Hypothesis 1a was examined first with a two-way mixed MANOVA and followed up with individual two-way mixed ANOVAs for the separate cheating behaviors. The direct effect of condition on moral disengagement will be examined via an independent samples t-test. Had, the three mediators been effective, they would have been ran in a full mediation model using SPSS Process Model 4 (Hayes, 2013). Analyses for the moderation of cheater status was tested in SPSS Process using Model 1 and controlling for Time 1 as a covariate where applicable.

Results

Descriptives and correlations for study variables are displayed in Table 1. The outcome variable of “cheating behaviors” was proposed to be made up of jump-out behaviors as well as the number of cheating photos that were identified for each individual. To start, correlations between the cheating behaviors were examined. As can be seen, the two types of cheating behaviors, jump-out behaviors ($r = .652, p < .01$), and cheating photos ($r = .719, p < .01$) correlated highly with themselves at the two different time points. However, when examined based on time point and not type of behavior, although still significant, jump-out behaviors and cheating photos from Time 1 ($r = .128, p < .01$) and jump-out behaviors and cheating photos from Time 2 ($r = .096, p < .01$) had much lower correlations. To begin, a two-way mixed MANOVA was conducted to examine the cheating behaviors together, then they were later observed individually.

Table 1
Descriptive Statistics and Correlations for Study Variables

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Condition Group	0.50	0.50	—												
2. Cheater Status	0.28	0.45	.06	—											
3. Guilt Time 1	1.59	0.64	.012	-.005	—										
4. Guilt Time 2	1.52	0.64	-.046	-.015	.831**	—									
5. Fear Time 1	1.51	0.59	-.001	.003	.800**	.729**	—								
6. Fear Time 2	1.44	0.60	-.002	.007	.710**	.829**	.820**	—							
7. Moral Disengagement	-1.02	0.62	.058	.072*	.029	.051	.042	.083**	—						
8. Jump-out Behaviors Time 1	0.94	4.13	.026	.367**	-.009	-.020	-.026	-.016	.088**	—					
9. Jump-out Behaviors Time 2	0.59	3.62	-.103**	.223**	-.027	.014	-.032	-.011	.055	.652**	—				
10. Cheating Photos Time 1	0.71	2.30	-.041	.496**	.008	.010	.019	.035	.065*	.128**	.080**	—			
11. Cheating Photos Time 2	0.57	2.21	-.156**	.329**	-.021	-.009	-.010	.017	.077*	.090**	.096**	.719**	—		
12. Knowledge Test Scores Time 1	16.1	4.66	.047	.312**	-.105**	-.103**	-.048	-.052	.102**	.310**	.199**	.371**	.268**	—	
13. Knowledge Test Scores Time 2	13.6	4.74	-.091**	.193**	-.086**	-.078*	-.044	-.054	.069*	.203**	.286**	.250**	.323**	.620**	—

Notes. *n* = 1049-1077. For Condition Group, 0 = Control, 1 = Warning. For Cheater Status, 0 = Non-cheater, 1 = Cheater. Moral disengagement reflects recoded values.

p* < .05. *p* < .01.

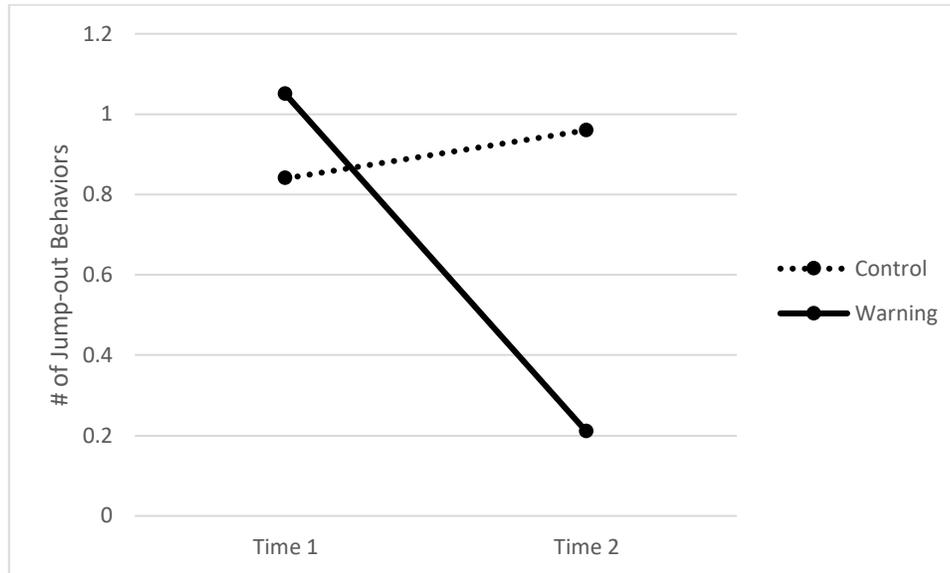
Hypothesis Testing

The first hypothesis regarding the effect of a middle-warning message on the number of cheating behaviors from Time 1 to Time 2 was tested using a 2 (Time) x 2 (Condition) mixed-model MANOVA, with Time being the within-subjects factor and Condition being the between-subjects factor. Results from the MANOVA show that there was a significant Time x Condition interaction effect on the combined dependent variables of jump-out behaviors and photo behaviors $F(2, 1046)=20.94, p<.001, \text{Wilks' } \Lambda = .961$. Thus, the number of cheating behaviors that were performed from Time 1 to Time 2 depended on which message condition an applicant was in. To further explore this relationship, individual two-way mixed ANOVAs were examined.

Cheating behaviors were next examined separately, as jump-out behaviors and cheating photos. Results for these analyses can be found in Table 2. Jump-out behaviors were found to have a significant Time x Condition interaction. Descriptives further show that while the warning group had a slightly higher number of jump-out behaviors at Time 1 (Control, $M=0.84, SD=3.89$; Warning, $M=1.05, SD=4.36$), jump-out behaviors significantly decreased for the warning condition at Time 2 ($M=0.21, SD=2.21$) compared to the control condition ($M=0.96, SD=4.58$) which slightly increased as can be seen in Figure 2. The simple effects were further examined, and it was found that the difference from Time 1 to Time 2 for the warning group was significant ($F(1, 1075)=35.49, p<.001, \eta_p^2 = .032$, and while the difference between groups at Time 1 was not significant, the difference between the warning group and control group at Time 2 was significantly different ($F(1,1075)=11.42, p=.001, \eta_p^2 = .011$).

Figure 2

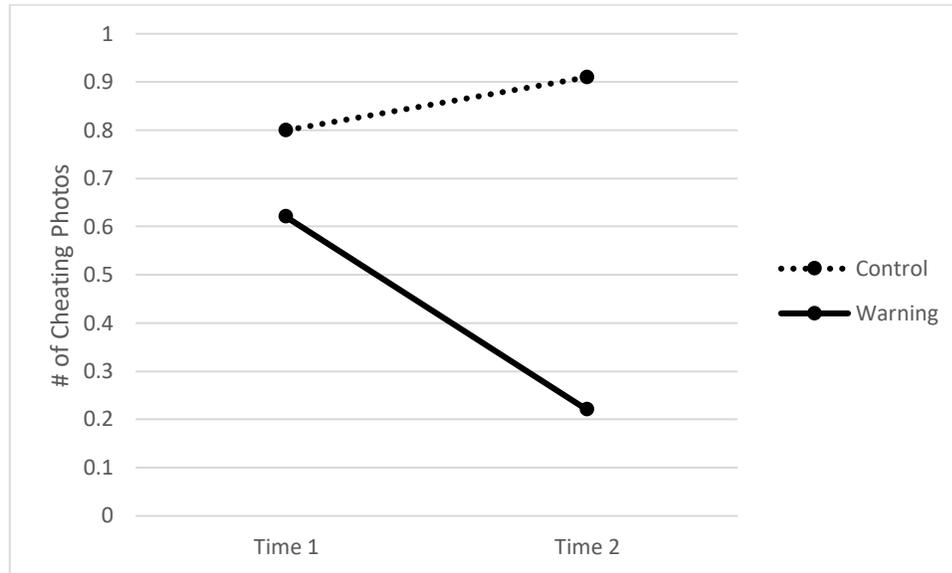
Mixed ANOVA Jump-out Behaviors Interaction



Cheating photos also were found to have a significant Time x Condition interaction, such that the number of cheating photos in Time 1 were fairly similar for the warning condition ($M=.62, SD=1.88$) and the control condition ($M=.80, SD=2.65$), but significantly deviated come Time 2 such that the warning condition ($M=.22, SD=.90$) decreased significantly while the control condition ($M=.91, SD=2.94$) again slightly increased. This interaction is plotted in Figure 3. An examination of the simple effects found that there while there was no significant difference between the groups at Time 1, the difference observed at Time 2 were significantly different ($F(1,1047)=124.86, p<.001, \eta_p^2 = .024$). The decrease for the warning group from Time 1 to Time 2 was also found to be significant ($F(1,1047)=29.13, p<.001, \eta_p^2 = .027$).

Figure 3

Mixed ANOVA Cheating Photos Interaction



Hypothesis 1b examined if there were differences in knowledge test scores due to the warning message. As originally hypothesized, we expected there to be a decrease in knowledge test scores for those who received a warning message. As seen in Table 2, this hypothesis was also supported with a significant Time x Condition interaction. Knowledge test scores at Time 1 for those in the warning condition ($M=16.40$, $SD=4.80$) were slightly higher than those in the control group ($M=15.96$, $SD=4.52$), however they were significantly lower at Time 2 for the warning group ($M=13.16$, $SD=4.51$) than for the control group ($M=14.03$, $SD=4.93$). Thus, Hypothesis 1b was supported for knowledge test scores. Further examination of the simple effects found that scores between the groups significantly differed at Time 2 ($F(1,1074)=9.05$, $p=.003$, $\eta_p^2=.008$). However, it should also be noted that both groups were found to have significantly lower test scores from Time 1 to Time 2.

Table 2*Two-Way Mixed ANOVA Results*

Variable	Level	<i>df</i>	MS	F	<i>p</i>	Effect Size
Jump-Out	Time	(1,1075)	68.56	13.12	<.001	.012
	Condition		38.04	1.533	.216	.001
	Time*Condition		122.27	23.40	<.001	.021
Cheating Photos	Time	(1,1047)	10.91	7.79	.005	.007
	Condition		100.65	11.65	.001	.011
	Time*Condition		33.29	23.77	<.001	.022
Knowledge Test	Time	(1,1074)	3583.90	436.41	<.001	.289
	Condition		24.53	.685	.408	.001
	Time*Condition		229.41	27.94	<.001	.025
Guilt	Time	(1,1075)	2.89	42.185	<.001	.038
	Condition		.250	.335	.563	.000
	Time*Condition		.724	10.59	.001	.010
Fear	Time	(1,1075)	2.78	43.61	<.001	.039
	Condition		.003	.004	.948	.000
	Time*Condition		.000	.003	.956	.000

Notes. Effect size = partial eta squared.

The second hypothesis examined the mediating effect of moral disengagement on reducing cheating behaviors and knowledge test scores. This hypothesis was broken down, and thus the direct effect of Condition on Moral Disengagement was first observed directly. Results from an independent samples t-test showed that there was no significant effect of Condition on Moral Disengagement ($t(1075)=-1.89, p=.059$). However, the means showed that the control group ($M=-1.05, SD=.63$) did have higher levels of moral disengagement than the warning group

($M=-.99$, $SD=.58$). Mediation results show no significant effect of moral disengagement as a mediator of condition with knowledge test scores $.0064$ [$-.0218$, $.0386$] or cheating behaviors $.0049$ [$-.0017$, $.0160$]. Thus, Hypothesis 2 was not supported.

Hypothesis 3 and 4 were tested in a similar manner. Due to the complexity of having both mediators and outcomes measured at two time points, the direct effects of the independent variable were tested on the mediators first using a two-way-mixed ANOVA to gain a clear picture of the relationship occurring. It was hypothesized that those in the warning group would have an increase in guilt and fear from Time 1 to Time 2, due to the warning message. Guilt was examined first, and no significant effect of Condition on Time was found as seen in Table 2. Inconsistent with expectations, guilt decreased for both groups from Time 1 to Time 2. Such that those in the warning group at Time 1 ($M=1.60$, $SD=.64$) actually had decreased guilt at Time 2 ($M=1.49$, $SD=.62$) more so than the decrease seen in the control group from Time 1 ($M=1.58$, $SD=.64$) to Time 2 ($M=1.55$, $SD=.65$). Therefore, due to the insignificant findings, the full mediation was not tested. Thus Hypothesis 3 was not supported.

Hypothesis 4, as seen in Table 2, was also inconsistent with our expected results. Fear was found to decrease for both groups at a similar rate from Time 1 to Time 2. Those in the warning group at Time 1 ($M=1.51$, $SD=.60$) started with the same level of fear as those in the control group ($M=1.51$, $SD=.58$), and they both experienced decreased in fear at Time 2 (Warning, $M=1.44$, $SD=.58$; Control, $M=1.44$, $SD=.62$). Thus, Hypothesis 4 was not supported.

The fifth hypothesis examined cheater status as a moderator of cheating behaviors and knowledge test scores. Hypothesis 5a, was expected to be significant due to the fact that cheater status was created from Time 1 cheating behaviors. An interaction term was created for cheater status and condition. Results from a linear regression show that the interaction term was

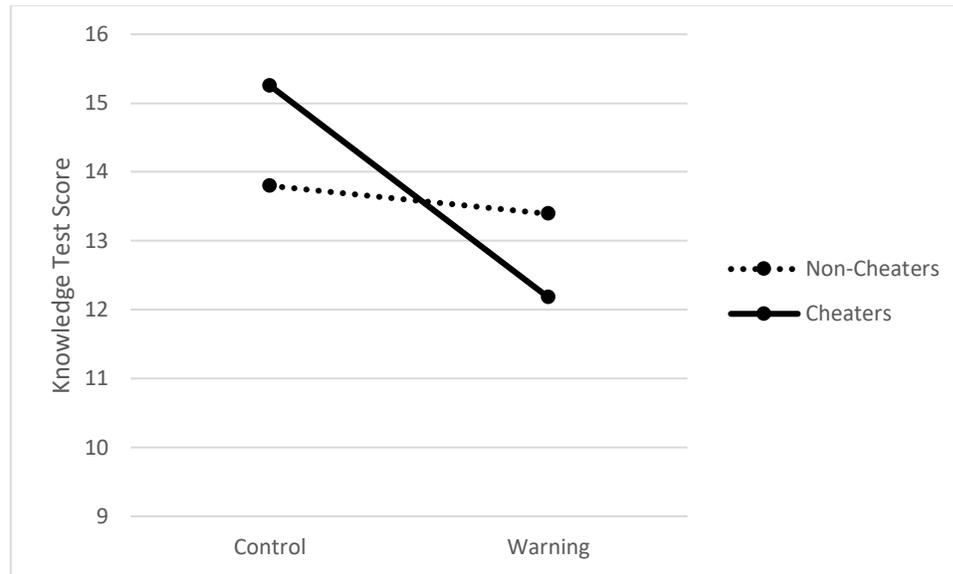
significant, $\beta = .445$, $t(1048)=14.37$, $p<.001$. Thus, cheater status did moderate the relationship between condition and cheating behaviors, supporting hypothesis 5a. To examine the behaviors individually, Model 1 of the PROCESS Macro (Hayes, 2013) was used to regress Time 2 cheating behaviors on Condition and the cheater status interaction term, with Time 1 cheating behaviors as a covariate. The moderating effect of cheater status on jump out behaviors was seen through the increase in ΔR^2 for the interaction term ($\Delta R^2 = .034$, $F(1,1072)=68.55$, $p<.001$). This same effect was again seen for cheating photos. With cheater status again found to be a significant moderator ($\Delta R^2 = .027$, $F(1,1044)=64.19$, $p<.001$), thus supporting Hypothesis 5a.

Hypothesis 5b was examined in a similar manner, with cheater status moderating the effect of Condition on Knowledge Test Scores. Knowledge test scores were again controlled for at Time 1. Cheater status was found to significantly moderate this relationship ($\Delta R^2 = .0158$, $F(1,1071)=28.96$, $p<.001$). This interaction is plotted in Figure 4. As can be seen below, cheaters in the control group had significantly higher test scores than non-cheaters, and cheaters in the warning group had significantly lower test scores than the non-cheaters. Although there was a small decrease in knowledge test scores for non-cheaters who received the warning message, their test scores at Time 2 were not significantly different from non-cheaters who received the control message as was seen in the conditional effects ($b = -.41$, 95% CI $[-.920, .106]$, $t = -1.56$, $p = .12$). Thus, Hypothesis 5b was supported.

Lastly, the moderating effect of cheater status was tested on the three mediators for Hypothesis 6a-c, again controlling for Time 1 guilt and fear. It was found that cheater status did not significantly interact with condition to impact moral disengagement ($\Delta R^2 = .0023$, $F(1,1073)=2.47$, $p=.12$), guilt ($\Delta R^2 = .000$, $F(1,1072)=.149$, $p=.70$), or fear ($\Delta R^2 = .0008$, $F(1,1072)=2.70$, $p=.10$). Thus, Hypothesis 6a, 6b, and 6c were not supported.

Figure 4

Knowledge Test scores based on Cheater Status



Discussion

The current study looked to examine the effectiveness of a middle-warning message in reducing cheating behaviors in an online knowledge test. Along with this, cognitive (moral disengagement) and emotional (fear and guilt) mechanisms were observed, in addition, the impact of a middle-warning message on test scores was examined. Previous research has shown the effectiveness of pre-warning messages on reducing cheating behaviors (Bing et al., 2012; Corrigan-Gibbs et al., 2015) and support has been found for the usage of middle-warning messages to reduce faking behaviors in personality tests (Fan et al., 2012). However, no study to date has utilized a middle-warning message to deter cheating behaviors. Thus, this study looked to fill this gap.

As was hypothesized, those who received a middle-warning message did in fact show a significant reduction in cheating behaviors from Time 1 to Time 2 and this reduction was significantly different from the small changes that were seen in those who did not receive a

warning message. This supports previous literature that suggests giving a serious warning message, one that cautions of realistic consequences, can reduce cheating behaviors (Corrigan-Gibbs et al., 2015). Being that this is a competitive honors program, and the warning message cautioned that the college would be notified if cheating behaviors continued, it makes sense that our message may have been effective. This provides further support for the idea that without some type of deterrent, cheating behaviors can and do occur in online environments.

It has also been previously hypothesized that test scores in online environments are usually higher due to possible cheating behaviors that may occur (Steger et al., 2018). As was seen in Hypothesis 1, those who received a control message, had higher knowledge test scores than those who received a warning message. However, when examining this closer in Hypothesis 5, it was seen that for cheaters who received the control message, their scores were the highest of any group, while cheaters who received the warning message had the lowest scores of any group. This goes to show, that unless given some stipulation as to what may occur if cheating takes place, applicants may abuse the online environment and attempt to cheat in order to increase their test scores.

The second hypothesis was based on the premise of the Theory of Moral Disengagement (Bandura, 1986), such that, individuals have self-regulatory systems that prevent them from engaging in dishonest behaviors. Previous research has found support for this idea, and while the current study could not fully support this claim, it did come close. Moral disengagement scores were higher for those who received the warning message; thus, they had a harder time acting unethically, potentially due to the warning message. The difference from the warning message group and the control group was not statistically different, however as discussed below, this may have been due to potential translation issues in the scale. Another plausible explanation

is a cultural difference that may be prevalent in terms of cheating behaviors. As previously mentioned, evidence has been found to support the notion that certain Asian countries, such as Japan and China, have increased rates of cheating and norms surrounding cheating behaviors (Diekhoff et al., 1999; Bernardi et al., 2008). Given that one premise of the theory of moral disengagement is justifying one's actions by comparing themselves to others, the lack of moral disengagement found in the current study may be due to an increased acceptance of cheating behaviors in certain Asian countries. Thus, if the applicants perceived that their peers may also be taking advantage of the online environment and attempting to cheat, they may have also cheated but not reported a decrease in moral disengagement due to the internal justification they are making for themselves.

There has been limited research examining emotional mechanisms related to warning messages and cheating behaviors. The current study looked to examine the effectiveness of moral and non-moral emotions in mediating the warning message and cheating behaviors, however no significant effect was found for guilt or fear. In fact, both guilt and fear decreased from Time 1 to Time 2 for both those who received a control message as well as a warning message. Therefore, these particular emotional mechanisms were not the reasons for the seen behavioral changes as was predicted. Although guilt has been examined in the morality literature, perhaps again a cultural difference is playing a role in its impact. For instance, maybe those who receive a warning message were able to justify their actions and therefore did not report feelings of guilt.

One explanation for the decrease in fear from Time 1 to Time 2 can be found in test anxiety literature. A previous study found that students' levels of fear/anxiety decreased significantly from pre to post exam on three different testing occasions (Yih, Kirby, Spitzer, &

Smith, 2019), an earlier study also reported similar decreases in levels of fear from pre to post examination (Smith, & Ellsworth, 1987). Although, it still would have been expected that once a warning message was received fear may have increased, especially for those who were cheating (as hypothesized in Hypothesis 6c), this was not found to be true in our study. Perhaps, the level of heightened fear that comes along with the testing environment for an admissions decision, is even higher than that in a regular academic testing environment. Therefore, once students began the exam, they may have relaxed, and their levels of fear may also have decreased, regardless of additional factors.

Limitations and Directions for Future Research

One limitation of this study came from the reliability of the moral disengagement scale. The original authors reported an alpha of .90 in their pilot study and an alpha level of .80 in their actual study (Shu, Gino, & Bazerman, 2011). Our study however only had an alpha level of .56, which is often considered below acceptable (Cho & Kim, 2015). A potential explanation for this large discrepancy may be due to translation error. The initial scale was only ever used on an English-speaking sample. Therefore, in the translation from English to Chinese, there may have been rephrasing that did not clearly translate. The other explanation is that the scale itself needs further work. Shu and colleagues (2011) created and utilized this specific scale for a unique setting. Given that the scale is so specific to moral disengagement in a cheating environment, it has received little additional usage. Therefore, future studies should look to add support or to modify the current scale to better understand the role that morality may play in unethical cheating behaviors.

A second limitation of this study is the generalizability of the results across cultures. Given that some Asian cultures normalize and/or report higher levels of cheating behaviors

(Diekhoff et al., 1999; Bernardi et al., 2008), our Chinese sample's cheating behaviors may have been higher than other cultures. Therefore, applicants in other countries may not only utilize different cheating behaviors or be found to have fewer cheating behaviors, they may also react differently to receiving a warning message. Perhaps, fear or guilt may mediate the relationship between warnings and behaviors in other cultures where cheating is not so prevalent.

Additionally, if fear increases, there is the potential for there to be a larger impact on knowledge test scores, which should ultimately try to be avoided. We found that regardless of message condition, knowledge test scores did decrease from Time 1 to Time 2. However, researchers should verify that middle-warning messages do not significantly negatively impact applicants in other cultures so much so that it test scores are inadvertently lowered. Therefore, future research should examine middle-warning messages in applicants from other cultures to better understand the generalizability of our findings.

This study was able to find support for the usage of middle-warning messages to deter cheating behaviors in an online environment. Previous studies only utilized warning messages at the beginning of tests, however it is unclear if one method is better than the other. Future studies should look to compare the two methods to determine if one is more effective at reducing overall cheating behaviors. Although we would still hypothesize that a middle-warning message would be more effective, due to the habituation of warning messages at the beginning of tests, empirical evidence is needed to support such claims. Additionally, research has started to show the possibility of being able to catch cheaters in action using machine-learned video techniques (Li, Chang, Yuan, Hauptmann, 2015). This would be similar to the photo coding done in this study; however, the new method would be able to determine cheating behaviors as they happen. If this becomes a more common practice, it is necessary to fully understand how receiving a middle-

warning message, as cheating behaviors occur, would affect applicants. It should also be noted that the interaction effect of condition and cheater status on knowledge test scores was much more informative than solely looking at the effect of condition. As it was made clear that it is specifically cheaters who are largely impacted by the warning message, rather than non-cheaters, and therefore future research should continue to utilize cheater status as a moderator when looking at these relationships.

Practical Implications

The current study showed the importance of detecting and deterring cheating in an online selection test. As companies continue to move to online environments, they need to be sure that their online platforms are creating a fair and even playing field for all applicants. Thus, if some individuals attempt to utilize the online environment as a means to cheat, companies need to be sure that they are able to identify the applicants who are acting unethically or at least attempt to decrease their dishonest behaviors. The current study utilized a real applicant pool to show the effectiveness of a middle-warning message. Results from this study should be used to continue to drive selection practices, such as implementing warning messages in tests that may allow for cheating. Previous studies have largely relied on test scores to determine if cheating has potentially occurred (e.g., Harmon & Lambrinos, 2008; Kantrowitz & Dainis, 2014), however new IT systems, like the one utilized in this study, are allowing for more robust methods of identifying unethical behavior. Companies should move towards these more advanced methods, as it is clear that without some type of deterrent, cheating behaviors can and do occur.

Despite numerous attempts to prevent cheating, there is still a need to better understand, from a practitioner and researcher perspective, which of these methods are most effective at reducing cheating and why. Although this study was unable to answer the question of why

certain methods may work, it did find that middle-warning messages can, and do, significantly decrease cheating behaviors. Thus, researchers should continue working to understand why such methods, like middle-warning messages, are effective at reducing unethical behaviors in a selection context.

References

- Akeley, S. J., & Miller, A. (2012). The effects of instructor fear appeal and moral appeals on cheating-related attitudes and behavior of university students. *Ethics and Behavior, 22*, 196-207.
- Arthur Jr., W., Glaze, R. M., Villado, A. J., & Taylor, J. E. (2009). The magnitude and extent of cheating and response distortion effects on unproctored internet-based tests of cognitive ability and personality. *International Journal of Selection and Assessment, 18*, 1-16.
- Auseubel, D. (1955). Relationships between shame and guilt in the socialization process. *Psychological Review, 62*, 378-390.
- Bandura, A. (1986). *Social foundations of thought and action*. Upper Saddle River: Prentice Hall.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*, 364-374.
- Bernardi, R. A., Baca, A. V., Landers, K. S., & Witek, M. B. (2008). Methods of cheating and deterrents to classroom cheating: An international study. *Ethics and Behavior, 18*(4), 373-391.
- Bing, M. N., Davison, H. K., Vitell, S. J., Ammeter, A. P., Garner, B. L., & Novicevic, M. M. (2012). An experimental investigation of an interactive model of academic cheating among business school students. *Academy of Management Learning and Education, 11*(1), 28-48.
- Burke, E. (2009). Preserving the integrity of online testing. *Industrial and Organizational Psychology, 2*, 35-38.

- Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well-known but poorly understood. *Organizational Research Methods, 2*, 207-230.
- Connelly, S., Helton-Fauth, W., & Mumford, M. D. (2004). A managerial in-basket study of the impact of trait emotions on ethical choice. *Journal of Business Ethics, 51*, 245-267.
- Corrigan-Gibbs, H., Gupta, N., Northcutt, C., Cutrell, E., Thies, W. (2015). Deterring cheating in online environments. *ACM Transactions on Computer-Human Interaction, 22*, 1-23.
- Diekhoff, G. M., LaBeff, E. E., Shinohara, K., & Yasukawa, H. (1999). College cheating in Japan and the United States. *Research in Higher Education, 40*, 343-353.
- Do, B.-R. (2009). Research on unproctored internet testing. *Industrial Organizational Psychology, 2*, 49-51.
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology, 51*, 665-697.
- Ellingson, J. E., Heggstad, E. D., & Makarius, E. E. (2012). Personality retesting for managing intentional distortion. *Journal of Personality and Social Psychology, 102*, 1063-1076.
- Fan, J., Gao, D., Carroll, S. A., Lopez, F. J., Tian, T. S., & Meng, H. (2012). Testing the efficacy of a new procedure for reducing faking on personality tests within selection contexts. *Journal of Applied Psychology, 97*, 866-880.
- Farnese, M. L., Tramontano, C., Fida, R. & Paciello, M. (2011). Cheating behaviors in academic context: Does academic moral disengagement matter? *Social and Behavioral Sciences, 29*, 356-365.
- Foster, D. Secure, online, high-stakes testing: Science fiction or business reality? *Industrial and Organizational Psychology, 2*, 31-34.
- Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review, 9*,

384-394.

Gaudine, A. & Thorne, L. (2001). Emotion and ethical decision-making in organizations.

Journal of Business Ethics, 31, 175-187.

Harmon, O. R. & Lambrinos, J. (2008). Are online exams an invitation to cheat? *Journal of*

Economic Education, 39, 116-125.

Harrington, S. J. (1997). A test of a person-issue contingent model of ethical decision making in

organizations. *Journal of Business Ethics, 16*, 363-375.

Hayes, A. F. (2013). Introduction to Mediation, Moderation, and Conditional Process Analysis:

A Regression-Based Approach. New York, NY: The Guilford Press. *Journal of Educational Measurement, 51*(3), 335-337.

Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.

Jones, T. M. (1991). Ethical decision making by individuals in organizations. *Academy of*

Management. The Academy of Management Review, 16(2), 366-395.

Kantrowitz, T. M. & Dainis, A. M. (2014). How secure are unproctored pre-employment tests?

Analysis of inconsistent test scores. *Journal of Business Psychology, 29*, 605-616.

Karim, M.N., Kaminsky, S. E., & Behrend, T. S. (2014). Cheating, reactions, and performance in

remotely proctored testing: An exploratory experimental study. *Journal of Business*

Psychology, 29, 555-572.

Kish-Gephart, J., Detert, J., Trevino, L. K., Baker, V., & Martin, S. (2014). Situational moral

disengagement: Can the effects of self-interest be mitigated? *Journal of Business Ethics, 125*, 267-285.

Klignite, V., Connelly, S., Thiel, C., & Devenport, L. (2013). The influence of anger, fear, and

emotional regulation on ethical decision making. *Human Performance, 26*, 297-326.

- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). Comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, *127*, 162-181.
- Lazarus, R. S. (1991). *Emotion and adaption*. New York: Oxford University Press.
- Leijten, M. & Waes, L. V. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communicaiton*, *30*(3), 358-392.
- Li, X., Chang, K., Yuan, Y., & Hauptmann, A. (2015). Massive open online proctor: Protecting the credibility of MOOCs certificates. *Collaroration in the Open Classroom*.
<https://doi.org/10.1145/2675122.2675245>.
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *American Marketing Association*, *45*, 633-644.
- Nabi, R. L. (1999). A cognitive-functional model for the effects of discrete negative emotions on information processing, attitude change, and recall. *Communication Theory*, *9*, 292-320.
- Pearlman, K. (2009). Unproctored internet testing: Practical, legal, and ethical concerns. *Industrial and Organizational Psychology*, *2*, 14-19.
- Shu, L. L., Gino, F., Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin*, *37*(3), 330-349.
- Smith, C. A., & Ellsworth, P. C. (1987). Patterns of appraisal and emotion related to taking an exam. *Journal of Personality and Social Psychology*, *52*(3), 475-488.
- Steger, D., Schroeders, U., & Gnambs, T. (2018). A meta-analysis of test scores in proctored and unproctored ability assessments. *European Journal of Psychological Assessment*, *36*, 174-184.

- Tangney, J.P. (1999). The self-conscious emotions: Shame, guilt, embarrassment, and pride. In T. Dalgliesh & M. J. Power (Eds.), *Handbook of cognition and emotion*. (pg. 541-568). Chichester, England: Wiley.
- Tippins, N. T., Beaty, J., Drasgow, F., Gibson, W. M., Pearlman, K., Segall, D. O., & Shepherd, W. (2006). Unproctored internet testing in employment settings. *Personnel Psychology*, *59*, 189-225.
- Watson, D., & Calrk, L. A. (1994). *The PANAS-X: Manual for the Positive and Negative Affect Schedule-Expanded Form*. Ames: The University of Iowa.
- Wiechmann, D., & Ryan, A. M., (2003). Reactions to computerized testing in selection contexts. *International Journal of Selection and Assessment*, *11*(2), 215-229.
- Yacout, O. M., & Vitell, S. (2018). Ethical consumer decision-making: The role of need for cognition and affective responses. *Business Ethics: A European Review*, *27*, 178-194.
- Yih, J., Kirby, L. D., Spitzer, E. G., & Smith, C. A. (2019). Emotion as a process: Appraisal, emotion, and coping patterns across time. *Motivation Science*. <http://doi-org.spot.lib.auburn.edu/10.1037/mot0000144.supp>

Appendix A: Warning and Control Messages

Middle-warning Message (to be translated into Chinese):

Attention!

Thank you for your participation in this assessment. The page you currently see is the system information. This page has appeared because the system found that you may be using external resources during the testing process, but we are not sure if it is caused by a system malfunction or if you have actually used external resources. You have been flagged once. The system will judge whether the test is valid or not based on your subsequent testing behavior. If you are no longer marked during the subsequent answers, your results will not be affected. Please remember, you should complete all of the questions independently, usage of external resources, asking others for help, or sharing questions with other test takers is not allowed. If you continue to be flagged by the system, your test score will be considered invalid and will be reported to the university. Please select “Resume” to continue.

Control Message (to be translated into Chinese):

Thank you for participating in this portion of the selection process. The page you currently see is the system information, which indicates that the testing system is undergoing a random system check to ensure that the system is working properly, please be patient. There is a countdown timer at the bottom of the screen, when the timer becomes zero, this indicates that the random system check is complete. Please select “Resume” to continue.

Appendix B: Measures

Moral Disengagement

Shu, L. L., Gino, F., Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin*, 37(3), 330-349.

[To be translated into Chinese]

Instructions: The following statements are about how you felt during the knowledge test. Please report to what extent you agree with these statements. Your answers in this section are for us to improve the test system and do not affect your admission.

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

1. I felt that sometimes getting ahead of the curve was more important than adhering to rules.
2. I felt that rules should be flexible enough to be adapted to different situations.
3. I felt that cheating was an appropriate behavior because no one got hurt.
4. I felt that if others engaged in cheating behavior, then it was morally permissible for me to engage in cheating behavior as well.
5. I felt it was appropriate to seek short-cuts as long as it was not at someone else's expense.
6. I felt that the end result was more important than the means by which I pursued those results.

