The Effect of Financial and Food-Based Incentives on Math Test Performance

Thomas Johnson III

Senior Seminar

Mr. Jason Scott

May 2023

Abstract

Previous research has suggested that incentives may be effective in increasing a student's math performance. The purpose of this study was to extend previous research and give the two incentives in the same time period, as well as introduce food as an incentive. Thirty-one students with Algebra 2 experience took incentivized math tests with two different types of incentives. The financial incentive outperformed the control, but the food incentive did not. The relationship between participants' math ability and incentive effectiveness was positive but showed no significance. Therefore, teachers should consider using incentives to increase their students' math test performance. We suggest more research is needed on the precise effectiveness of food as an incentive.

Introduction

Incentives are usually used to help people perform well during a task, and research suggests performance can be increased if used correctly (1). Incentives are common in workplace environments, while incentives are not as common in academic settings (2). In private businesses, 38% of employees want incentives, and 35% prefer cash incentives (3). However, high school students are often only incentivized by their grades (4). What if cash or food could also incentivize performance for high school students? Many variables can alter the effectiveness of incentives, such as the type of incentive, amount of incentive, gender, and types of tasks. If incentives improve performance, it is important to figure out how teachers can effectively use them to increase classroom achievement. In our study, we gave students a math test under three different conditions: a control, a financial incentive, and a food-based incentive, in order to see which one helps students the most.

Incentives can be used in many forms as well as multiple types. A study by Levitt et al. determined that the type of incentive will have different effects on performance (1). The researchers presented students with a standardized test and either a normal incentive or a loss incentive. The normal incentive was $20, $10, or a trophy. Normal incentives were provided if they met the requirements for receiving them. The loss incentive was the same three choices, but the students were initially given them; the students were told that if they did not meet the requirement for their test, they would have to give the incentive back (1). The loss incentive resulted in worse performance than no incentive and the normal incentive outperformed both. Out of all the subjects, the incentives had the largest effect on math. It was also found that financial incentives proved to be more effective for males and older students while it was less effective for younger students (1). This is important to our study because it gives an insight into which incentives are more likely to yield negative and positive results. In the high school context, money, given in a class like math, might produce the biggest effect.

Performance can be influenced by incentives, but some incentives have less of an effect than others. A study by Gneezy determined that the incentive has to reach a certain threshold in order for it to affect performance (4). In their study, 200 students were split evenly into two groups. They were then given $24 and told to choose how much of it they wanted to give to the other participant, between $4-$24, their choice would yield positive or negative results for the person receiving (4). For the researcher's experiment, they had two groups, proposers and responders, and the proposers chose how much of the $24 to give to the responders, and were told about the increasing and decreasing incentives (4). Once they decided on an amount the responders were given either a high positive incentive, a low positive incentive, no incentive (called the Dictator condition in this study), a low negative incentive, or a high negative incentive. What the researchers found was that the low incentives had less of an effect than no incentive at all (4). So it was concluded that no incentive was better than low incentives; this produced "the W effect" (Figure 1). This is important to the current experiment is because it introduces the idea of needing to meet a certain threshold with incentives.

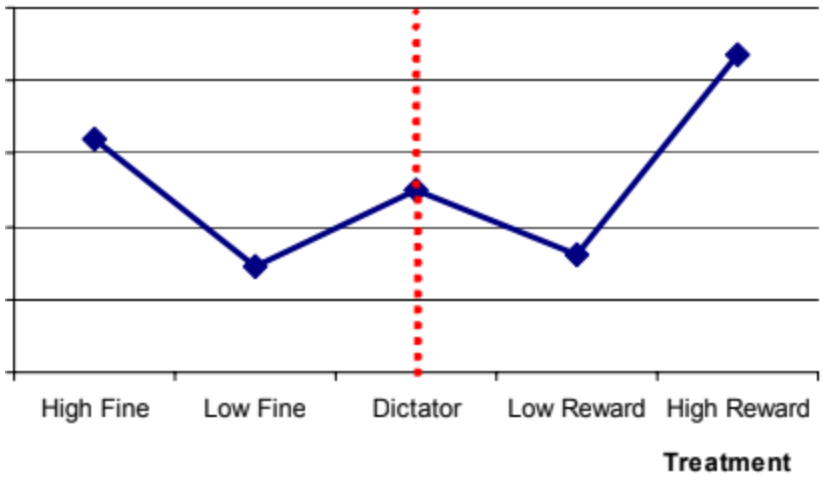

Figure 1. The W Effect. *The low rewards tend to produce lower results than the control, hence the W shape (4). The figure is quoted directly from the prior study (4).*

The reason incentives are given is to help participants perform well, but it is not always clear how someone is supposed to perform. Newell et al. hypothesized that rule discovery leaves performance unchanged by incentives. Rule discovery is when the person receiving the incentive must figure out what they are doing and how they should perform to get the incentive (5). For the researchers' study, they held a series of tests using different variations of cards and had the participants try to figure out the pattern (5). From this experiment, it was determined that no matter the incentive, when rule discovery is dynamic, performance will not be enhanced or diminished. Instead, they found that rule-following, or telling someone what they need to do to receive an incentive, worked better. In school, students are commonly asked to do rule-following projects, so this finding is important to gauge the change in performance with incentives in a school setting.

When introducing incentives to a group, participants' personal characteristics must be considered. Furquan et al. conducted an experiment that did just that, they measured the effects of incentives on gender and grade. In the researchers' experiment, they had two classes: a first-year accounting course, and a fourth-year mathematics class. The two classes were given a test in their respective courses and were split into 3 sections: a control with no change in grade, 5% of their grade (group 1), and 10% of their grade (group 2)(6). The control group was told the test was practice, the 5% were told the test was 5% of their term grade, and the 10% group was told the test was 10% of their term grade (6). The results were that the 10% group outperformed the 5% and nothing group, but gender did not have any role to play in the effect of incentive (6). Regarding grades between the genders, there was a difference in grades, but it was not due to gender. It was because of interest levels in the class (6). This is useful knowledge because as gender does not play a role in incentives, we did not have to factor this into our study design.

In the classroom, incentives have been proven to have some effect, but other researchers have noted they are not commonly used. A study by Fryer proved that incentives did not have much effect because of the student's behaviors. Incentives mentally get students to the point where they are ready to do the work, but they can not physically put it into practice (2). For the researcher's study, they went to schools in Dallas, New York, and Chicago and gave them different academic incentivized tasks. In Dallas,

students were given $2 for every book they read up to 20 books and had to score 80% on their test. In New York, 4th and 7th graders got paid for their performance on 3 reading and 3 math tests, with money for a perfect score, and money for showing up. In Chicago, students were paid for their grades: $50 for A, $35 for a B, $20, for C, and $0 for a D (2). The only positive finding was those native English-speaking Dallas students, who read books and got paid tended to score higher (2).

The reason incentives do not always work in schools is because of multiple variables. Those variables are lack of knowledge, lack of interest, problems outside of the task, and a lack of correlation between the task and how it's measured. (2). For example, the students may have the drive to perform the task for the incentive but they don't know how to achieve it. Other students find themselves wanting the incentive, but not wanting to put in the work for it (2). But sometimes it's not the student's fault, it could be from poor quality of teaching or parenting, leading to the student not being able to perform to the necessary standard (2). It can also be the incentive's fault itself, for not correlating them with the task at hand. An example of this is when students had to score an 80% on a test as their final task, but we're getting an incentive for reading books. Reading books alone does not increase one's reading test abilities (2). These variables contribute to the reason incentives aren't always as effective in school settings.

Incentives come in different forms like financial, rewards, and even academic bonuses, but one of the less commonly accepted is food. Food as an incentive in schools is frowned upon for multiple reasons because it is viewed as "unhealthy" (7). Research by the Connecticut State Department of Education has shown that food as an incentive in the classroom compromises classroom learning, contributes to poor health, encourages "overconsumption of unhealthy foods, poor eating habits, and increases preference for sweets." Even though food as an incentive may bring about positive results in behavior and academics in the short term, it is still viewed as a negative idea and many argue it should not be done (7). Therefore, in this study, we intend to test the effect of food on performance, to see if its benefits outweigh the drawbacks.

Overall, research has suggested that incentives have a positive effect on performance. There are variables to take into account such as type of incentive, amount of incentive, type of task, participant interest, skill, and academic subject (1,4,6). This study contributes to the study of incentives because, to our knowledge, no one has done a study on both food and monetary incentives. This provides a direct comparison between the two types of incentives and which one is better in an academic setting. For our study, a group of students was given three separate math tests over a period of time.

First, we hypothesized that scores on a test would differ significantly by incentive type: money, food, or none. This is because research suggests that the amount the incentive is valued will affect performance (4). However, no studies, to our knowledge, have tested food, so whether or not it helps is unknown. Second, we hypothesize that there is a positive relationship between math ability and how much the incentives increase one's performance. This is because research shows that math is the subject most affected by incentives in general, and performance is dependent on the level of one's ability to score higher and a person's interest in completing the task (6, 2).

Method

A total of 31 students at The Neighborhood Academy participated in the study. The students were between the ages of 15 and 17, and the group was 69% male and 31% female. The students were selected based on their experience in math; Honors Algebra 2 (semester 1 average = 86%), Standard Algebra 2 (89%), and Honors Trigonometry (87%) classes were used for this study as they had finished or were currently taking Algebra 2.

For this project to have conclusive data, we needed to gather incentives; they were food (chocolate bars or assorted Flamin' Hot chips), money ($2), and a control with no incentive. Based on our personal knowledge of the students, we knew Flamin' Hot chips were the favorite, in addition to candy. The financial incentive was approximately equal in value, based on convenience store prices. For our tests the primary author reviewed practice SAT tests and selected questions that Algebra 2 students would be able to do (**8**). Then, the primary author made three 10-question tests and had 6 volunteer students (not in the study) take them to make sure the test wasn't too easy or too difficult. One of the tests was too difficult (with a lower average score), so a question from that test was replaced with an easier question. The average scores of these tests were 5.5, 5.58, and 5.83 out of 10 on the three tests, so we made a cut-off of 5+, or "average, or better than average" for the incentive. Also, a math teacher reviewed the tests, to make sure they were appropriate for Algebra 2 and were of similar difficulty. The participants were assigned random orders of the tests, but they had the same conditions every time.

Incentives were distributed in the following order: food, control, then money. The reason was that we found a lot of evidence that money would have a large effect so we saved it for last. We didn't want to start with no incentive, because the participants may not care for future tests if they were not compensated right away. When distributing the tests, it was explained to each math class that they would have 10 minutes to complete the tests, and if they got 5 or more questions correct they received the incentive for that day. The tests were graded and incentives were handed out during the same period. While giving the tests we decided to leave some time between each test so they were given every other weekday. All participants who were present for two or more of the tests completed them and were included in the study. Students absent for one test made it up on a later day.

Results

In order to find out the effect of incentives on performance, the primary researcher constructed three different tests, with three different incentive types for a group of student volunteers. The students were assigned one of the three tests randomly to take on each day, with each day being a different incentive. Then students were scored and received an incentive the same period if they got 5 or more out of ten they received an incentive. We expected to find that incentives have a positive effect on performance. Seen below in figures 2-4 is data on the different incentive's effectiveness for the tests. The data is generally clumped between 4 and 6.

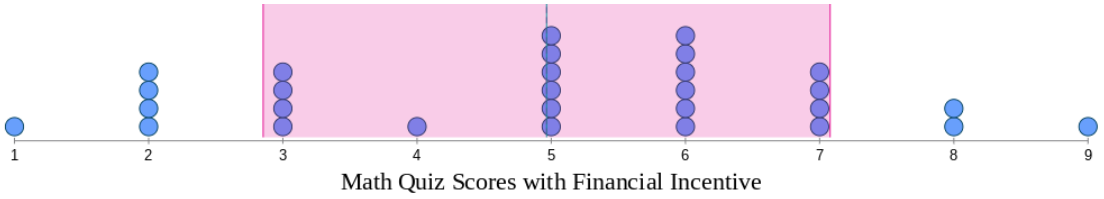

Figure 2. Math quiz scores with food incentives

The picture above shows the scores of students who took a math quiz with a food incentive, their scores had a mean of 5 and a standard deviation of 5 +- 2.1. The data is generally symmetrical and mound-shaped.

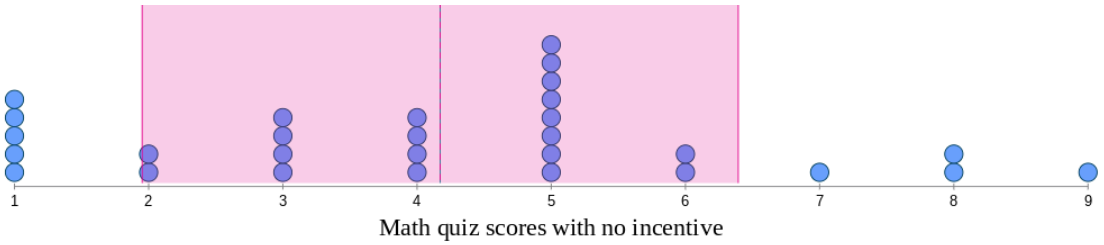

Figure 3. Math quiz scores with no incentive (control)

The picture above shows the scores of students who took a math test with no incentive (control), their scores had a mean of 4.2 and a standard deviation of 4.2 +- 2.2. The data is mound shaped with an unusual cluster on the left.

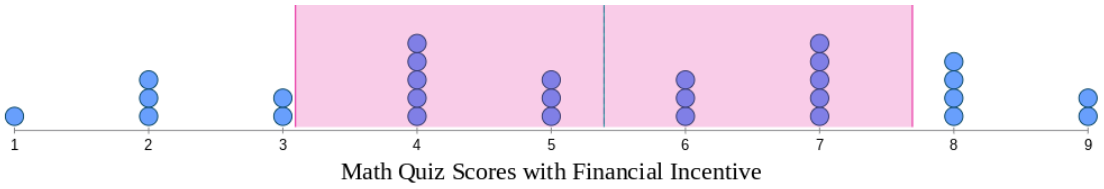

Figure 4. Math quiz scores with financial incentives

The picture above shows the scores of students who took a math test with a financial incentive, their scores had a mean of 5.4 and a standard deviation of 5.4 +- 2.3. The data is generally bimodal.

Our first hypothesis is that test scores will differ by incentive. Test scores are based on three different practice SAT tests, each with 10 questions. Three incentives were used: no incentive, food, and money. A one-way ANOVA for correlated samples found a significant effect on incentive type ($F(2,56)=3.47$, $p=0.038$). A Tukey post hoc test analysis found that the money incentive ($M=5.4$ $SD=2.3$) outperformed the control ($M=4.25$ $SD=2.2$, $p<0.05$). All other differences, including food ($M=5.0$ $SD=2.1$), were not significant.

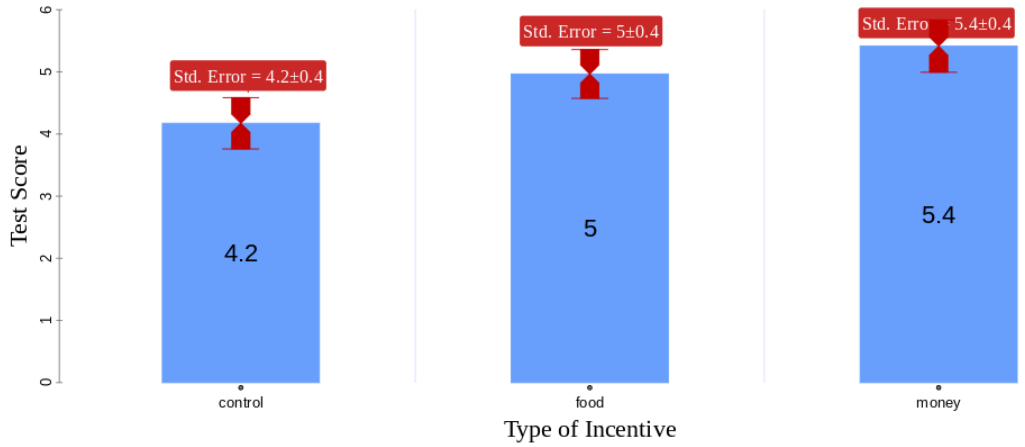

Figure 5. Test scores by incentive type. *The horizontal axis shows the 3 different incentives that the 29 people experienced across 3 different days. The vertical axis shows their test scores. The money incentive outperformed the control incentive ($p<0.05$).*

Our second hypothesis was a positive relationship between math ability and incentive effectiveness. We used students' semester one math grades to measure their math ability. To measure the effect of food and money, we subtracted the two incentive's results from the results of the control group. A correlation coefficient r-test found a positive non significant relationship between the effect of money and math scores ($r(27)=0.29$, $p=0.07$). People who score better in math may be more incentivized by money, than those who are less skilled at math, but the p-value was not statistically significant. However, we did find a positive significant relationship between the effect of food and math scores ($r(27)=0.45$, $p=.007$). People who score better in math respond better to food incentives than people who aren't as good at math.

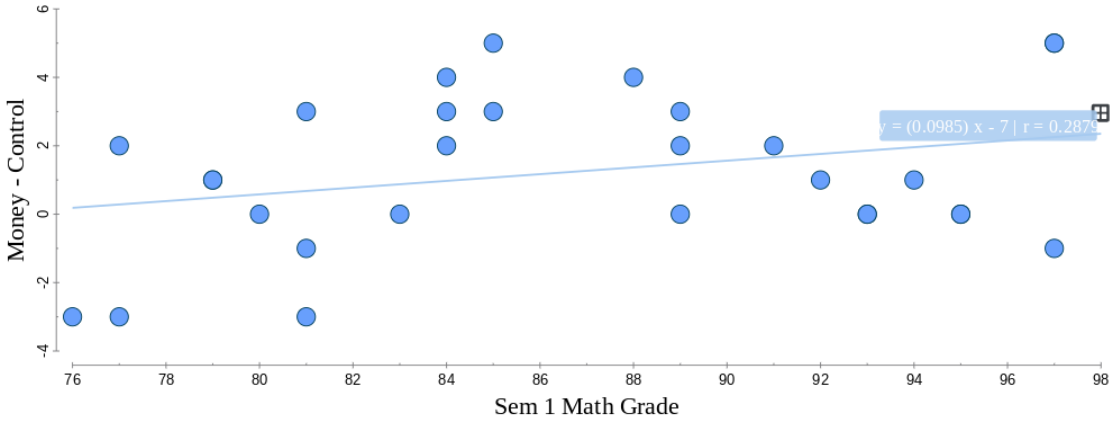

Figure 6. Money effect on semester 1 grades. *The picture above shows students' semester 1 grades and the effect of the financial incentive compared to no incentive. The relationship between the variables is not significant ($p=0.07$).*

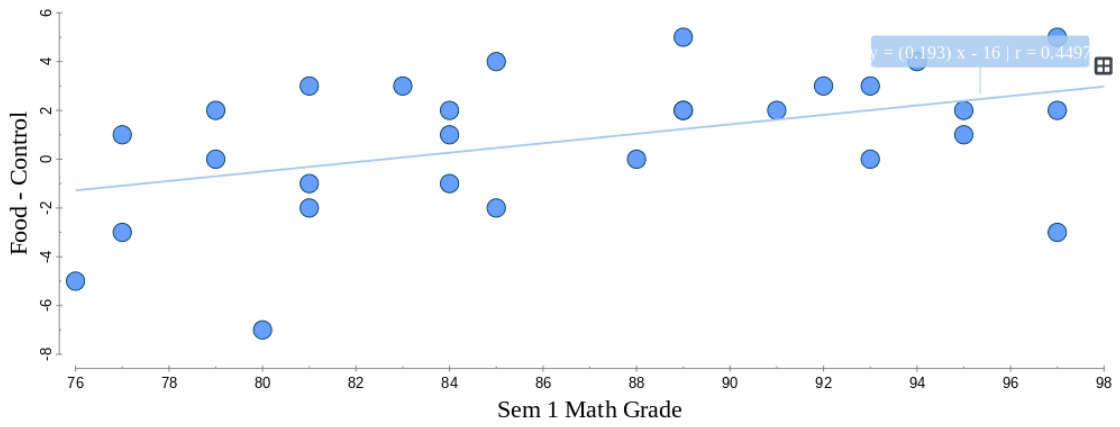

Figure 7. Food effect on semester 1 grades. *The picture above shows students' semester 1 grades and the effect of a food incentive compared to no incentive. The relationship between the variables is significant (p=0.007).*

Discussion

In this study, we wanted to find out if incentives improved students' performance. Our first hypothesis was that performance would differ by incentive type. This was partially supported because the money incentive on average scored higher than the no incentive average, and there was no significant difference between the food incentive average and the no incentive average (Figure 5). Our second hypothesis was that there was a positive relationship between incentives and math performance. This was partially supported because students' math performance was affected by the food incentive, but not the money incentive (Figures 6 and 7).

Our results are consistent with other studies by Levitt et al, and Gneezy. Gneezy found high financial rewards caused better performance and Levitt found that financial incentives helped males and older students perform better on standardized math and reading tests. Our study found that financial incentives resulted in higher performance than the control, no matter the circumstances. Whether it be a game for fun or a test for a grade, money can help to boost performance. These studies help strengthen the claim that money helps people perform better. Additionally, the amount of money must be taken into account. Gneezy's idea of the W-effect supports this claim that the correct amount of money must be given to yield positive results. Our results suggest that we had the correct amount of $2, but $2 is not the magic number for all groups. The environment and recipients of the incentives must be taken into consideration because the amount required will be different for every group. Having knowledge of context is useful, we chose $2 because it has immediate use for our participants. They could walk out and use the $2 and purchase something from the school store, giving the money immediate value.

There is not a lot of research on food as an incentive, however, we agree with other researchers that food incentives shouldn't be used (**7**). The first problem is incentivizing unhealthy habits. Giving food as an incentive is the same as telling someone it's okay to eat unhealthy foods for doing well. This could lead to poor health, especially if used in schools because they'll be forming these habits so young (**7**). Not only is it an unhealthy habit, but it's also a hit or a miss for its effectiveness as an incentive. It's best to just not do a food incentive and just use money because as our study has found money is effective as an incentive. Money's effectiveness comes from the idea of spending however the participant wishes. Instead of buying food as an incentive and promoting unhealthy habits, giving someone the money to buy

whatever they want is better than deciding for them. The students also are going to be more incentivized because they are buying what they want, not what the instructor picks for them. This way if they decide to buy unhealthy foods, it won't be anyone's fault but their own.

Similar to Levitt et al, we found that math as a subject is effective when using incentives. The reason for this is because of the simplicity of math, there is only one right answer. The incentives push students to work a little harder, and they put in a little more work to get the right answer. This is unlike a reading test, where no matter how much the answer is desired if the reading can't be understood, no answer will be found. Also, the value placed on the student's relationship with math must be taken into account (6). For our study, we used students who have previously taken Algebra 2 or are currently taking Algebra 2, with mostly honors classes to make sure they knew the concepts. So these students are similar to other studies where students had high interest and knowledge of their subject (6). All of these students have more knowledge of math and more interest because of their courses, making them more interested in math itself and more likely to be incentivized. Notably, we found that students with lower math ability did about the same as the control. While students with a higher math ability were more affected by incentives.

Our study had multiple limitations. Our first limitation was the number of people; it was hard to prove the effectiveness of food because we didn't have a large number. Our suggestion for other researchers would be to use more people. Another limitation would be the amount of money we had. If we had more money the results could have more significant effects than those already consistent with Gneezy. Our suggestion for other researchers would be to use the highest value incentives for maximum effect. Another limitation was our participants, they all came from different age groups. Some took Algebra 2 years ago and some were taking it at the time. Our suggestion would be for researchers to find people at the same level and experience when it comes to the math being tested. One future improvement that could be added to this project would be using real food, like pizza and not snacks, as an incentive. This may boost the effectiveness of food as an incentive. Also finding if there is some sort of emotional difference between receiving and missing the incentives could let researchers see if the possible negatives outweigh the positive benefits of incentives.

According to our research, incentives, especially monetary incentives, are likely effective in boosting academic performance on math tests. However, the amount and type of incentives used may cause the effect to vary. Therefore, incentives might be able to be implemented in academic settings to increase the overall performance of students. Teachers and administrators should focus on finding the right kind of incentive that works in their environment because incentive types will vary from classroom to classroom.

Acknowledgments

I would like to thank Mr. Max Cookingham, Ms. Erin Simpson, and Mr. Jason Scott for lending their math classes. As well as the group of students who took part in helping provide feedback on my test.

Works Cited

1. Levitt, Steven D., et al. “The Impact of Short-term Incentives on Student Performance.” University of Chicago, September 2011, https://mfidev.uchicago.edu/events/20111028_experiments/papers/Levitt_List_Neckermann_Sadoff_Short-Term_Incentives_September2011.pdf. Accessed 14 November 2022.
2. Fryer Jr, Roland G. “Financial Incentives and Student Achievement: Evidence from Randomized Trials.” *Quarterly Journal of Business*, vol. 126, no. 4, 2011, pp. 1755-1798.
3. “14 Employee Rewards Program Statistics.” *Applauz*, https://www.applauz.me/resources/employee-rewards-program-statistics. Accessed 14 November 2022.
4. Gneezy, Uri. “The W effect of incentives.” 13 October 2003, http://gsbwww.uchicago.edu/fac/uri.gneezy/vita/The ... f%20incentives37.pdf. Accessed 14 November 2022.
5. Osborn Popp, Pamela J., et al. “Rule discovery performance unchanged by incentives.” *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 44, 2022, pp. 840-846.
6. Furquan, Mehreen A., et al. “The Impact of Grade Incentives and Gender on Student Performance: An Experiment.” *The Lahore Journal of Business*, vol. 3, no. 2, 2015, pp. 79–92.
7. Connecticut State Department of Education. “Alternatives to Food as Reward.” September 2004, https://district.schoolnutritionandfitness.com/laurelsd/files/Alternatives_to_Using_Food_Rewards.pdf. Accessed 14 November 2022.
8. “Paper SAT Practice Tests – SAT Suite | College Board.” *SAT Suite of Assessments*, https://satsuite.collegeboard.org/sat/practice-preparation/practice-tests/paper. Accessed 26 January 2023.