JOURNAL OF
EMERGING
INVESTIGATORS

VOLUME 3, ISSUE 1 | JANUARY 2020
emerginginvestigators.org

Into perspective

A mathematical analysis of perspective in paintings

Chicken confinement

Does chicken captivity affect their associative learning?

Dehydration detection

An early detection sensor based on green fluorescent protein

Antibiotic-resistant bacteria

How prevalent are dangerous strains in everyday locations?

JOURNAL OF EMERGING INVESTIGATORS

The Journal of Emerging Investigators is an open-access journal that publishes original research in the biological and physical sciences that is written by middle and high school students. JEI provides students, under the guidance of a teacher or advisor, the opportunity to submit and gain feedback on original research and to publish their findings in a peer-reviewed scientific journal. Because grade-school students often lack access to formal research institutions, we expect that the work submitted by students may come from classroom-based projects, science fair projects, or other forms of mentor-supervised research.

JEI is a non-profit group run and operated by graduate students, postdoctoral fellows, and professors across the United States.

Contents

VOLUME 3, ISSUE 1 | JANUARY 2020

JOURNAL OF
EMERGING INVESTIGATORS

Article

The effect of varying training on neural network weights and visualizations

Emmett Fountain, Joe Rasmus
Williamston High School, Williamston, Michigan

SUMMARY

Neural networks are used throughout modern society to solve many problems commonly thought of as impossible for computers. As their use becomes more widespread, an issue of measurement arises. In order to create metrics with which to measure the ability of neural networks, new techniques must be constantly developed. The purpose of this research is to determine whether varying training produces measurable and consistent patterns in the visualizations and weights of Convolutional Neural Networks. In order to carry out this investigation, a convolutional neural network was designed and run with varying levels of training to see if consistent, accurate, and precise changes or patterns could be observed. To determine if such a change or pattern existed, the weights and filters were analyzed through visualizations, both qualitatively and quantitatively. Several patterns were discovered in the visualizations, but they were inconsistent across layers and the quantitative models were only consistent in specific circumstances. This indicated that while training introduced and strengthened patterns in the weights and visualizations, the patterns observed may not be consistent between all neural networks.

INTRODUCTION

Neural network technology, and machine learning more generally, is present throughout much of our modern world, from self-driving cars to text recognition (1). Machine intelligence is becoming a crucial part of our society, most notably in healthcare, education, the automotive industry, and general safety (2). Neural networks, machine learning implementations which use systems of weights and biases to make decisions, are often part of these intelligent systems. Due to the integration of machine learning in highly important fields of society, deficiencies in our understanding of its function can often be devastating. If an automated vehicle fails to notice and recognize an obstacle, it could result in death, similarly if an automated doctor in the not-so-distant future were to misdiagnose a patient it could be equally fatal. Understanding the effect of training on components of a neural network deepens our insight into how the weights affect predictions of the neural networks. This deeper insight could open up the possibility for new techniques of measuring competency of a network other than test accuracy, which is often heavily situational and can overrepresent functionality due to overfitting (that is, when a neural network works only on the specific training set or test set). As the field of machine learning grows, it becomes more important to understand the effects of training on neural networks and their weights - numbers that influence the output of a neural network that are crucial to their operation.

This investigation focuses on neural networks. At its simplest, a neural network works very similarly to a mathematical function, with weights modifying input values multiplicatively and biases modifying them additively. One way of "learning" new mappings from input to output is supervised learning where the network is fed both an input output pair which is known to be correct. For example, you could provide an image of a cat as input and the classification "cat" as the output. With this input and output the network calculates loss, a measure of how inaccurate the network's prediction was. The network then uses an algorithm, to modify the weights and biases to better map the input to the correct output. This kind of basic neural network is often called a fully connected neural network.

Convolutional Neural Networks (CNNs) are neural networks which contain at least one, but usually more, convolutional layers. Layers at their most basic are steps in a neural network which data must pass through, usually being modified in some way before being output to the next layer or as an output. Convolutional layers in particular, are layers which apply convolutions to an image. A convolution is an image analysis technique in which a filter - an ordered set of weights - is iterated over the input, modifying it to create a new output. Due to convolutions being the core of this type of neural network, these networks are often used to analyze and make predictions based off of images.

A simple CNN was created using Python and then trained with Cifar-10 (3), an image database used for machine learning. The filters used in the convolutions were then analyzed using techniques similar to those outlined during the 1989 Neural Information Processing Systems conference (4). This analysis generated qualitative and quantitative data used to compare the varying levels of training. My research aims to contribute to an understanding of the effect of training on weights by determining whether consistent and measurable patterns appear in visualizations of weights at varying degrees of training. To accomplish this goal, visualizations of weights were analyzed to look for patterns at varying levels

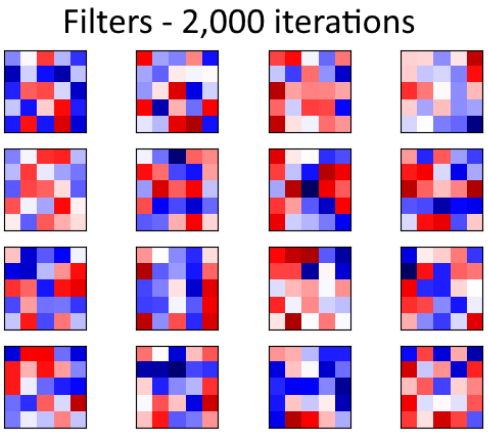

Figure 1. Weights of filters with little training are randomly distributed. All 16 filters in layer 1, after 2,000 iterations. Each panel is one of the filters of layer 1, the colors represent the value of the weights with blue being low and red being high. The exact values of the weights are not important in this situation because all we need to see are the patterns in the intensity and sign. Note the prevalence of intense weights throughout the filters, mixed seemingly randomly.

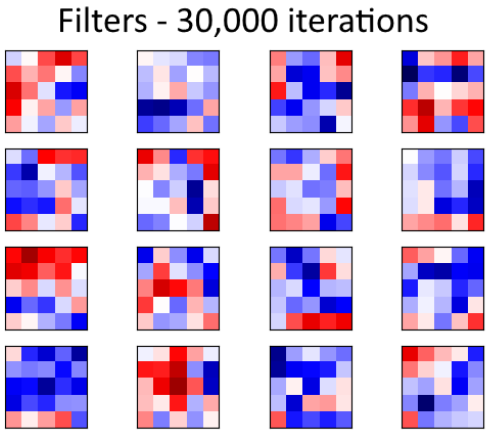

Figure 2. All 16 filters in layer 1, after 30,000 iterations. Weights of filters with more training form clusters Each panel is one of the filters of layer 1, the colors represent the value of the weights with blue being low and red being high. The exact values of the weights are not important in this situation because all we need to see are the patterns in the intensity and sign. There are far fewer intense weights, which appear mostly in clusters with a similar sign.

of training. In the analysis, a pattern was found and termed "softness", it was tracked both as a quantitative variable and as a qualitative variable. We predicted that quantifiable patterns would be detected, but they would most likely not be consistent.

RESULTS

The first qualitative trend noticed was a change in the distribution of strong weights among the filters. When there was little training, the weights were distributed with a random spread. As training increased, clusters of weights began to form. The distribution of weights across the filters changed from a random spread with little training to a spread with clusters of strong weights when there were higher training levels (**Figures 1 & 2**). Two other visualizations were created to analyze the weights. Both visualizations contained a number of panels, each with an image which has had a unique filter applied to it. The visualizations differed in the base images that were used; one used an image of a bird taken from the Cifar-10 dataset, while the other used an image of black and white static noise. A prominent pattern in these visualizations was the increasing prevalence of "soft" images when filters are better trained, seen in both the images generated from static and the bird images in both the first and second layer of both of these image sets (**Figures 3 & 4**). Images which are soft have pixels where close location corresponds with close

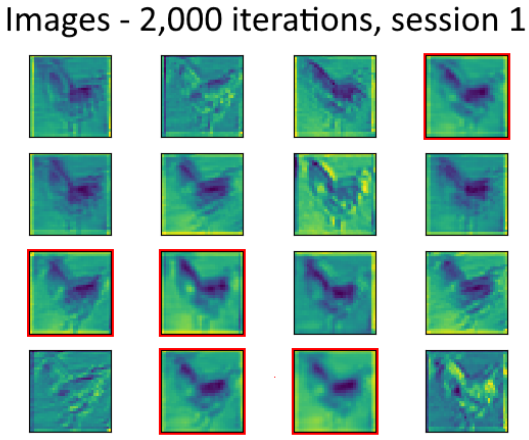

Figure 3. Images generated by applying all 16 filters in layer 1, after 2,000 iterations, to an image in the "bird" class of Cifar-10. Applying filters with little training to sample images generates images with little softness. Each panel was generated using a different filter from layer 1. Very few filters produced images that are soft (soft images highlighted in red).

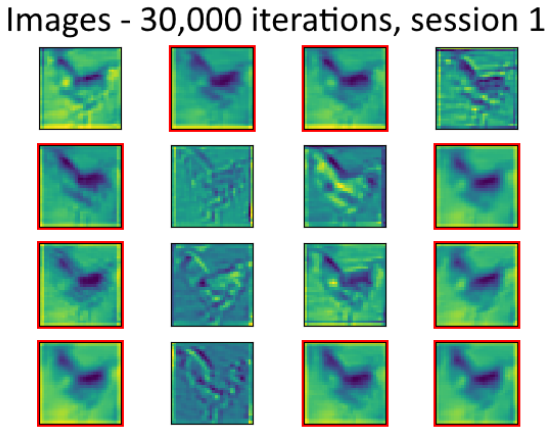

Figure 4. Images generated by applying all 16 filters in layer 1, after 30,000 iterations, to an image in the "bird" class of Cifar-10. Applying filters with more training to sample images generates images with greater softness. Each panel was generated using a different filter from layer 1. Far more filters produced images that are soft compared to the filters that were trained to 30,000 iterations (soft images highlighted in red).

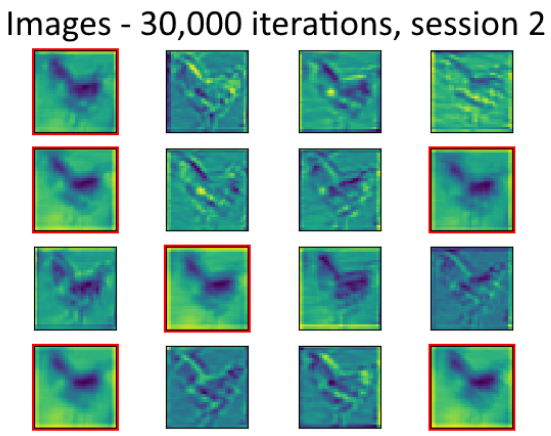

Figure 5. Images generated by applying all 16 filters in layer 1, after 30,000 iterations, to an image in the "bird" class of Cifar-10. Each panel was generated using a different filter from layer 1. The amount of softness is highly variable even with the same amount of training. While there are still more soft images then in Figure 3 with 2,000 iterations, there are noticeably less than in Figure 4, which had the same amount of training and was just a different session.

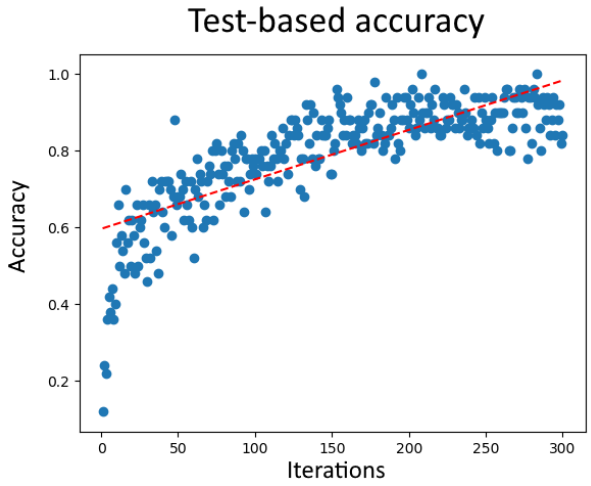

Figure 6. Test-based accuracy during training, reported at intervals of 10 iterations. Test based accuracy suggests a non-linear model. This figure includes the least squares regression line in red as a reference.

values, visually this means there are very few edges and little pixelation. When trained to 2,000 iterations, few of the 16 filters in layer 1 generated images which appeared soft. However, when trained to 30,000 iterations far more of the images were soft. This trend also occurred to some extent with the images generated by applying filters to static noise and when using filters from layer two (data not shown). However, the specific number of soft images varied per session even when the number of iterations remains constant. In two separate sessions filters were trained with 30,000 iterations but the filters of one session produced far more soft images than the other despite the same amount of training (**Figures 4 & 5**).

A graph was generated that showed the accuracy during training (**Figure 6**). This helped show the training trend for the model which gives an idea of what kind of numerical models might be plausible while analyzing the quantitative data. We observed that the rate of increase in accuracy decreased at higher training levels suggesting that a pattern would probably not be linear.

Almost all of the data had very low R and R-squared (R^2) values, meaning that not many of the linear regression equations fit the data very well. The notable exceptions were power functions from the second layer (**Figure 7**). The data generated from the bird image had an R2 value of about 0.9220 and the data generated from the static noise image had an R2 value of about 0.8996. Other than these two, all of the R2 values were below 0.8. Power functions for layer two

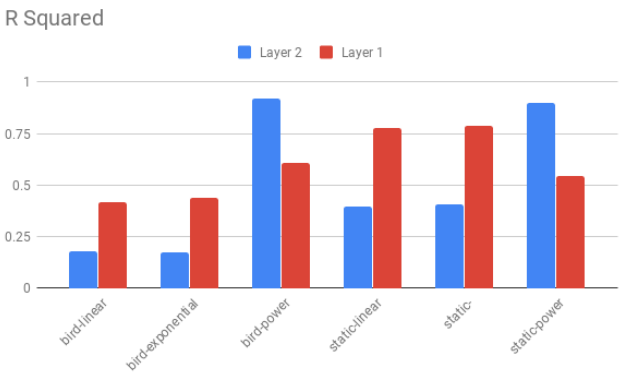

Figure 7. Bar graph showing the values of R2 for a line comparing softness to number of iterations. The graph includes the R2 values for all combinations of variables, that being layer, model, and image. In layer 2, the power model has a very high R2 value for both images. A high R2 value for the power model in layer 2 suggests that this model fits the data well.

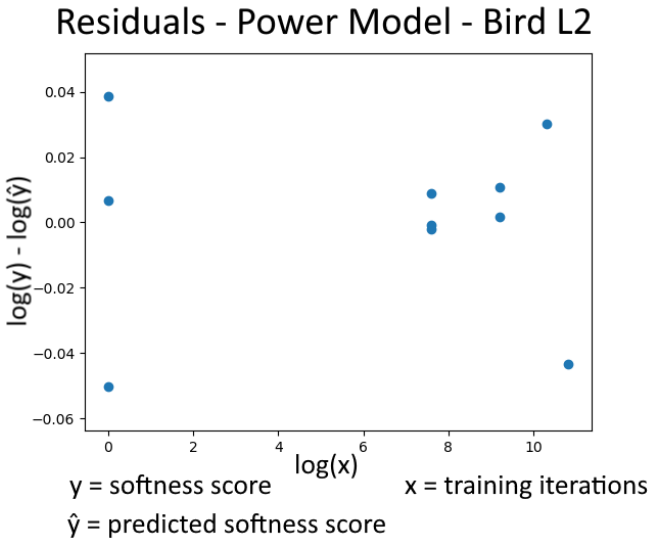

Figure 8. Residual graph showing the residuals for a power model of softness scores. Softness scores were calculated from the bird images generated from layer 2 as described in the methods section. There is not a strong curvature, which indicates that the power model would remain somewhat consistent if the model were extended to higher and lower batch sizes.

of the bird and static noise data, also had the lowest *p*-values for the two-tailed linear regression t-test with 1.05x10-5 and 2.90x10-5 respectively. However, these two were not the only regression equations that were significant at α = 0.05. All of the first layer data had significant *p*-values, but very low R2 values; however, from layer two, only the power function and the exponential model generated from static noise were significant at α = 0.05.

The residual graphs provided additional information as to whether the regression equations held up over time, an example of which can be found below (**Figure 8**). All of the graphs had some degree of curve, but several demonstrated more random scatter than others. The most random graphs were from layer two of the static noise straightened as a power function, both layer one and two of the bird image unstraightened, and layer one of the bird image straightened as a power function. The rest showed much more severe curving, which suggested that the pattern falls off as the data is extended.

DISCUSSION

We observed some consistent patterns in the weights and visualizations which were very localized or inconsistent. There is some evidence to suggest that consistent patterns are present, but evidence also suggests that these patterns may not be consistent across neural network models. The existence of consistent patterns in the weights was supported by qualitative evidence. We observed patterns in the visualizations of the weights through the clustering of strong weights, as well as the increasing number of filters that produce soft images (**Figures 1 & 2 and 3 & 4**). The existence of patterns was further supported by the quantitative evidence, which demonstrates some consistency and rigor to these trends. In layer 2 of the data generated from static noise, there exists an accurate power function that shows a strong relationship between training and softness. A high R2 value suggested that there was a strong correlation between the scores generated by the image and the number of iterations (**Figure 7**). The low *p*-value tells us that there was a significant linear correlation in the data when straightened as a power function. Finally, the random scatter of the residual graph showed that the power function is relatively consistent even as the pattern is extended. (**Figure 8**)

However, evidence suggests that the patterns might not be present in other neural network models or even when testing separate sessions. For example, a power function seemed to be the best fit for the data in layer 2, having higher R2 values and lower *p*-values than any other model for both images. The issue is that this same model was not consistent for layer 1, which is simply a different level of abstraction and doesn't have any fundamental differences from layer 2. Layer 1 has different models for each image which presents the problem that the patterns would be inconsistent as layers vary tremendously across models. The patterns would most likely be inconsistent with different test images and when using different neural networks. Along with this quantitative inconsistency, the qualitative trends were inconsistent across sessions (**Figures 4 & 5**). Therefore, more advanced quantitative research may prove unfruitful.

Overall, there were consistent patterns in Neural Network weights and visualizations within a specific model, but those patterns may not extend much beyond that specific model. In order for the findings of this research to have an impact on applied machine learning, more research would be needed to widen the scope of these findings.

While we conducted this research to the best of our ability, we could take several steps to improve the validity of this research or to further it. One such step would be to use computers that can better handle machine learning. It would have been highly informative to see how the data extended beyond this point, and it might have prompted some new insights. Other steps that could have been taken to generate a more representative sample and outline stronger trends include increasing the sample size, running more sessions, and applying the weights to more images. If future research was to be conducted it would be important to apply more advanced image analysis techniques and statistical analyses so that trends can be better quantified and new trends might be discovered.

METHODS

In order to test for patterns, a Convolutional Neural Network (CNN) was used; as it provides an acceptable level of complexity but is also simple enough that its weights can be easily accessed and analyzed. The CNN was built in Python with the TensorFlow library (5); the CNN had two convolutional layers and two fully connected layers. The batch size is variable, but due to hardware limitations a batch size of 50 was employed to generate the data. Cifar-10 (4) was used to generate the data, as it has a relatively large sample of images, but a manageable number of 10 classes. A function saved the convolutional filters as NumPy array files once the set number of training loop iterations had been completed.

A separate Python program was designed to view the filters in a variety of fashions. This program would display the filters as grids of colors ranging between blue and red depending on the value of the weight for each grid square. The program could also apply the filters to a sample image, in order to see how the filters affect the images they are applied to. The filters and images generated by the filters were grouped by which session and convolutional layer they were from. There were 5 different levels of training: 1, 2,000, 10,000, 30,000, and 50,000 iterations. The sessions with only one iteration were treated as baselines since they were essentially random as they had no time to train. Then, small, moderately small, moderately large, and large numbers of iterations were chosen so that the differences in visualizations at noticeably different training levels could be compared. Due to time and computing restraints, more sessions with fewer iterations were run than those with many iterations. Three sessions with

1 and 2,000 iterations were run, two sessions of 10,000 and 30,000 iterations were run, and only one session of 50,000 iterations was run.

In order to analyze the change in the filters, three kinds of images were qualitatively analyzed for each filter. This qualitative analysis was meant mostly as a heuristic to find possible patterns and support the quantitative analysis. The first of these image types was the red and blue color interpretation of the filters. The next image type was generated when the filters were applied to a 32 x 32 image with randomized black and white pixels. The original black and white image was generated as a 32 x 32 matrix with random ones and zeros, which was then transferred to an image with ones representing white pixels and zeros representing black pixels. The final image type was generated by applying the filter to a sample image from Cifar-10. For this sample image, a simple bird image was chosen to further reveal patterns not identified by applying the black and white image. The images of the same number of iterations and that shared a layer were analyzed together. The images were originally analyzed qualitatively and note was taken of any recurring patterns or features in a layer. The number of "soft" images for each group was also recorded. To quantify image trends, a python program was developed to objectively measure "softness" of an image.

The program first normalizes the images so that the value of each pixel is a floating point between 0 and 1. Once the images are normalized, a 3 x 3 filter is repeatedly applied to each image. This filter computes the standard deviation in that 3 x 3 area and returns it to the location in a matrix where the sample 3 x 3 area was taken. The standard deviations of each image were then averaged for each image, representing a score for the image which should theoretically be lower for softer images. After this score was computed, the scores for each layer were summed to give a total score for the layer. These scores were paired with their respective number of training iterations allowing the scores of each layer to be analyzed as a function of training iterations. Using these scores and their corresponding training iterations, statistics were generated in order to determine whether there was a relationship between training iterations and the softness score. A python program was used to generate a least squares regression line for the relationship between scores and iterations, R, the *p*-value for a two-tailed linear regression t-test with a null hypothesis of zero slope, and graphs of the residuals (**Figure 8**). As the accuracy graphs were clearly curved, the data was straightened to conform to exponential and power function models. Straightening, also called re-expressing, is a process which linearizes non-linear data, which serves many purposes. The same calculations were performed on both of the straightened data sets and the unstraightened data.

Linear regression equations can be used to describe the straightened data and as such different models for a set of data can be compared. Linear data allows for a linear regression t-test to be performed to statistically determine whether the data is linear and thus fits the model which it was straightened for. In general, straightening works by determining which model the data follows and applying the inverse of that to the data. This effectively "undoes" the original function leaving the data as a set of linear points. Curves can often be described by one of two models; exponential, and power models. As such, in order to determine which model described the data curve, the data was straightened as if it was one of these. Several tests were run on the resulting, now linear, data to see which model it fit best.

Received: May 29, 2019
Accepted: November 26, 2019
Published: December --, 2019

REFERENCES

1. Bojarski, Mariusz, *et al*. "End to End Learning for Self-Driving Cars." *arXiv.org ArXiv:1604.07316 [Cs]*, Apr. 2016., arxiv.org/abs/1604.07316.
2. Amato, Filippo, *et al*. "Artificial Neural Networks in Medical Diagnosis." *Journal of Applied Biomedicine*, vol. 11, no. 2, ScienceDirect, Jan. 2013, pp. 47–58., doi:10.2478/v10136-012-0031-x.
3. Krizhevsky, Alex. "Learning Multiple Layers of Features from Tiny Images" cs.toronto.edu, April 8 2009, ch. 3, cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.
4. Wejchert, Jakub and Gerald Tesauro. "Neural Network Visualization." *Advances in Neural Information Processing Systems 2 (NIPS 1989),* Morgan-Kaufmann, 1990, pp. 9. papers.nips.cc/paper/286-neural-network-visualization.pdf.
5. Abadi, Martín, *et al*.. "TensorFlow: Large-scale machine learning on heterogeneous systems", tensorflow.org, Nov 9 2015, static.googleusercontent.com/media/research.google.com/en//pubs/archive/45166.pdf.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Fluorescein or Green Fluorescent Protein: Is it possible to create a sensor for dehydration?

Sinaya Joshi[1], Dr. Louise Pennycook[2], and Dr. Youssef Ismail[2]
[1]Harker High School, San Jose, CA 95129
[2]Schmahl Science Workshops, San Jose, CA 95112

SUMMARY

Dehydration occurs when more fluid leaves the body than enters it. Dehydration can cause symptoms ranging from headaches and dizziness, to more severe symptoms like fever and unconsciousness. Currently there is no early dehydration detection system using temperature and pH as indicators. A sensor could alert the wearer and others of low hydration levels, which would normally be difficult to catch prior to more serious complications resulting from dehydration. The temperature and pH of skin are known to increase and decrease, respectively, with dehydration. These variables are also known to affect the fluorescence of certain fluorophores, which could provide a visible marker of dehydration. In this study, a protein fluorophore, green fluorescent protein (GFP), and a chemical fluorophore, fluorescein, were tested for a change in fluorescence in response to increased temperature or decreased pH. Neither fluorophore was affected by the changes in temperature. However, both lost their fluorescence when the pH of their environment was decreased. We also tested whether the fluorescence returned to its standard brightness when the temperature and pH were normalized. Reversing the pH change did not restore GFP fluorescence, but that of fluorescein was re-established. This finding suggests that fluorescein could be used as a reusable sensor for a dehydration-related pH change.

INTRODUCTION

Dehydration is when more fluid exits the body than enters it. This is a problem because water helps break down nutrients and move waste out of cells, making it essential for health and bodily function. (1, 2). Dehydration can cause symptoms ranging from headaches and lethargy, to fever and unconsciousness (1). Excessive sweating is one of the causes of dehydration, and can be triggered by both exercise and heat (3). For example, football players sweat not only from their many layers of protective equipment, but also as a result of thermoregulation during exercise, a process that maintains the body's core internal temperature. As a consequence, football players sweat 1.5 L/hr, an increase of 2.5% compared to people who do not regularly play sports (4, 5). Increased sweating makes football players more prone to dehydration if they do not replace the water lost through sweat. However, non-athletes can also become dehydrated if they do not consistently drink water throughout the day, resulting in a net loss of fluid. The purpose of this study is to address dehydration in athletes and non-athletes by making a sensor that alerts the dehydrated person to dangerously low fluid levels.

A sensor for dehydration could measure either pH or temperature of the skin to assess whether the person is dehydrated. Normal body temperature is 37 °C and rises to 39.5 °C during exercise (6). The maximum endurable body temperature is 41.1°C (6). An increase in temperature can be used as an indicator for dehydration because as body temperature increases, sweat production increases, which causes dehydration. The pH of skin can also be used as a measure of dehydration. The normal pH of human skin is 6, but sweat causes the pH to become more acidic. During light exercise, skin pH is 5 and further drops to pH 4 during heavy exercise (7). The minimum pH of skin is 3 (7). Acidic skin pH indicates excessive sweating, which can lead to dehydration. In order to monitor hydration levels using temperature and pH, a method must be developed to detect changes in these variables from the physiological ranges mentioned above.

One possible method to detect these changes is by using fluorophores. Fluorophores fluoresce light when their electrons are excited by other light sources. For the fluorescence to be seen, the emitted light must be of a very specific wavelength. One example of a fluorophore is green fluorescent protein (GFP). GFP is a protein that exhibits green fluorescence and comes from the jellyfish Aequorea victoria (8).

Protein fluorophores, like GFP, are known to be highly sensitive to changes in pH and temperature. Proteins rely on electromagnetic bonds between oppositely charged amino acids to help keep their tertiary structure. As the pH of the surrounding environment decreases, the concentration of H+ ions increases (9). Increasing the H+ ions can protonate negatively charged amino acids, thereby disrupting the electromagnetic bonds and causing a conformational change in the protein (9). This conformational change can affect the distribution of electrons, resulting in a loss of fluorescence (9). Temperature changes can also cause a conformational change in GFP. Increased temperature introduces extra energy into the protein. Extra energy can cause the bonds within the protein to vibrate and break, resulting in a change in protein shape and loss of fluorescence (10).

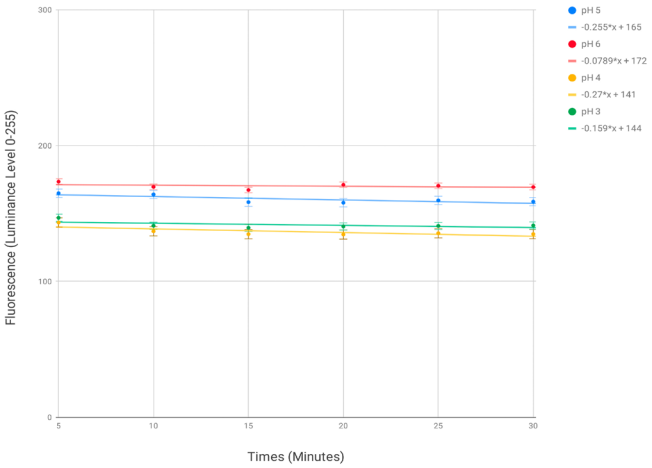

Figure 1: Fluorescence of GFP over time with pH change. This graph shows the change in the fluorescence of GFP when it was held at various pHs measured at intervals of 5 minutes for 30 minutes. GFP fluorescence decreases with decreasing pH, as depicted by the lines with a negative slope for pHs 3 through 5. There were three trials for every pH level, including pH 6, every 5 minutes for 30 minutes, and the averages of the trials were used for the six data points. The error bars represent the standard deviation of the mean of all three trials, showing how much the data varied at each point.

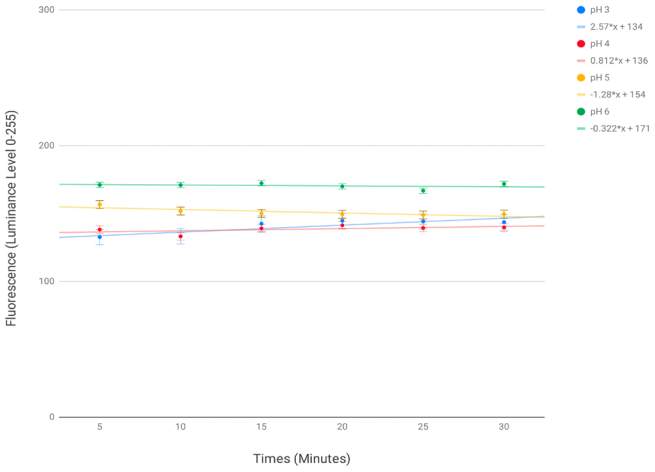

Figure 2: Fluorescence of GFP over time after return to pH 6. GFP fluorescence does not change with increasing pH, as depicted by the horizontal lines for pHs 3 through 5. There were three trials for every pH level, including pH 6, every 5 minutes for 30 minutes, and the averages of the trials were used for the six data points. The error bars represent the standard deviation of the mean of all three trials, showing how much the data varied at each point.

Fluorophores can also be structures other than proteins. Fluorescein is a chemical fluorophore that is used as a synthetic coloring agent and fluoresces yellow-green (11). Fluorescein is less susceptible to a conformational change compared to protein fluorophores, but it still has the potential to be affected by pH and temperature. A decrease in pH may also cause protonation of a chemical similarly to proteins. If fluorescein were to be protonated, the bonds and electron configuration within the molecule would change, thus reducing the fluorescence (12). An increase in temperature could also cause fluorescein to lose its fluorescence. If temperature is increased, the electron configuration of fluorescein could change independently of light input. Changing the electron configuration could affect the chemical's ability to emit light and be interpreted as a loss of fluorescence.

Since both GFP and fluorescein fluorescence are susceptible to changes in pH and temperature within the physiological range, this study aims to measure changes in fluorescence in response to hydration levels. The findings could be used to develop a sensor for dehydration, as change in fluorescent intensity could indicate dehydration. A sensor for dehydration could not only help people live healthier, but could also save lives.

RESULTS

This study used fluorescein and GFP as sensors for dehydration. We tested whether temperature or pH affected the fluorescent intensity of either substance.

After collecting fluorescein and GFP, we put both in tubes and changed their pH and temperature to the experimental groups. The tubes were photographed for 30 minutes. After 30 minutes, the tubes returned to their original temperature and pH and their recovery was photographed for 30 minutes.

The results of this experiment were analyzed using ImageJ to quantify fluorescence.

Fluorescence of GFP Changes with Decrease in pH But Does Not Recover with Return to Original pH

When the pH of GFP was decreased from pH 6 to pH 3 or 4, the fluorescent intensity immediately decreased. When GFP was at pH 5, the brightness decreased, but not as much as at pH 3 or 4. The fluorescent intensity of the control sample at pH 6 did not change. The trendlines for the data for pH between 3 and 5 had a negative slope, showing that the fluorescence decreased over time. The trendline for pH 6 also had a negative slope, but it was closer to 0, so the slight downward trend is likely from the error as shown by the error bars (**Figure 1**).

When the pH was raised to pH 6, the fluorescent intensity of the samples at pH 3 through 5 did not return to the intensity of the control. The fluorescent intensity of the control sample did not change. The trendlines for the data for pH 3 and 4 both had a positive slope, but the final brightness value was smaller than the original fluorescence shown in **Figure 1**. The trendline for pH 5, however, has a negative slope, indicating that the sample decreased in fluorescent intensity rather than returning to its original brightness. Once again, the trendline for pH 6 has a negative slope, but it is a small number and there are a few clear outliers so the negative slope could be due to error (**Figure 2**).

In the ANOVA analysis, the null hypothesis that changing the pH of GFP would not cause the fluorescence to decrease

	GFP	Fluorescein
pH 3	10.014	12.660
pH 4	10.069	13.071
pH 5	2.576	13.879
37 °C	2.004	1.950
39.5 °C	2.000	1.995
41.1 °C	0.960	2.897

Table 1: ANOVA analysis F distribution numbers for the fluorescence of GFP and Fluorescein with changed pHs and temperatures. This table shows the F distribution values for the experimental pHs and temperatures for GFP and fluorescein. If the values are less than 3.11, the null hypothesis is not rejected. If the values are over 3.11, the null hypothesis is rejected.

	GFP	Fluorescein
pH 3	1.039	73.731
pH 4	2.237	36.773
pH 5	1.202	22.739

Table 2: ANOVA analysis F distribution numbers for the fluorescence of GFP and Fluorescein after return to pH 6. This table shows the F distribution values for fluorescein and GFP when the pH was neutralized to pH 6 to determine if they could be reused. If the values are less than 3.11, the null hypothesis is not rejected. If the values are over 3.11, the null hypothesis is rejected.

was rejected, as the F distribution values for pHs 3 through 5 were all above 3.11 (**Table 1**). The results regarding the rejection of the null hypothesis that changing the pH of GFP to 6 would not cause the fluorescence to return to its original brightness were mixed. For pH levels 3 and 4, it was rejected, as the F distribution numbers in these cases were above 3.11. However, for pH level 5, the F distribution number was less than 3.11, indicating that it was not rejected (**Table 2**). This means that the fluorescence of GFP decreased when the pH decreased, but did not return to its original intensity when the pH was neutralized to 6.

Fluorescence of GFP Does Not Change with Temperature Change

The fluorescent intensity of GFP was not affected by any of the experimental temperatures compared to baseline. There were three experimental temperatures (37 °C, 39.5 °C, and 41.1 °C) because they were in the physiological range. GFP was incubated at these three temperatures for 30 minutes. The fluorescent intensity of GFP was expected to drop during these 30 minutes. However, the slopes of all three trendlines from time 0 minutes to time 30 minutes are extremely small and close to zero, signifying that the change in temperature did not affect the fluorescence of GFP (**Figure 3**).

The null hypothesis that changing the temperature of GFP would not cause its fluorescence to decrease was not rejected. For GFP temperature levels of 37, 39 and 41 °C, the F distribution numbers were less than 3.11, clearly indicating the null hypothesis was not rejected and temperature changes did not affect the fluorescence of GFP (**Table 1**).

Fluorescence of Fluorescein Changes with Decrease in pH and Recovers with Return to Original pH

Next we wanted to determine the effect of pH changes on fluorescein. Similar to GFP, when the pH of fluorescein was decreased from pH 6 to pH 3 or 4, the fluorescent intensity immediately decreased. When fluorescein was at pH 5, the brightness decreased, but not as much as at pH 3 or 4. The fluorescent intensity of the control sample at pH 6 did not change. The trendlines for the data for pH between 3 and 5 had a negative slope, showing that the fluorescence decreased over time. The trendline for pH 6 had a positive slope, but there are two outliers for 25 minutes and 30 minutes that are almost 10 units larger than the other four values. We speculated that the lighting in our photo booth may have changed and let in more ambient light, which would have increased the overall light in the picture, appearing as an increase in fluorescence to the image analysis in ImageJ (**Figure 4**).

When the pH was raised to pH 6, the fluorescent intensity of the samples at pH 3 through 5 returned to the intensity of the control. The fluorescent intensity of the control sample did not change. The trendlines for pHs 3-5 in the graph all had a positive slope that resulted in approximately the same fluorescent intensity as baseline at the final timepoint, showing that the fluorescence returned to its original brightness. The trendline for pH 6 is much straighter and, although there is a nonzero slope, it is small enough to assume that any variation

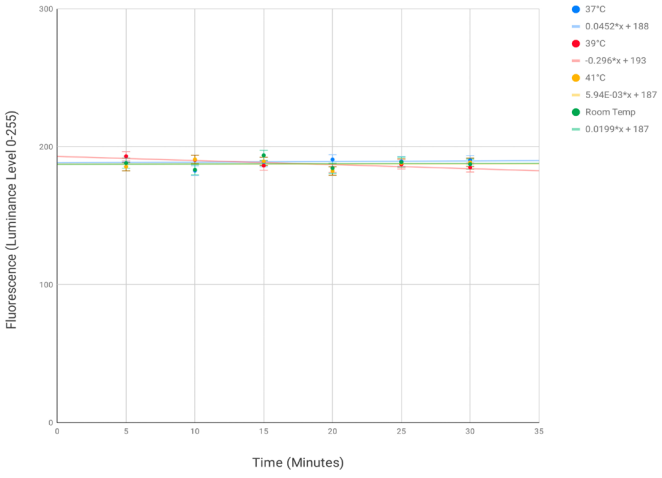

Figure 3: Fluorescence of GFP over time with changed temperatures. GFP fluorescence does not change with increasing temperature, as depicted by the horizontal lines for temperatures 37 °C, 39.5 °C, and 41.1 °C. There were three trials for every temperature, including room temperature (23 °C), every 5 minutes for 30 minutes, and the averages of the trials were used for the six data points. The error bars represent the standard deviation of the mean of all three trials, showing how much the data varied at each point.

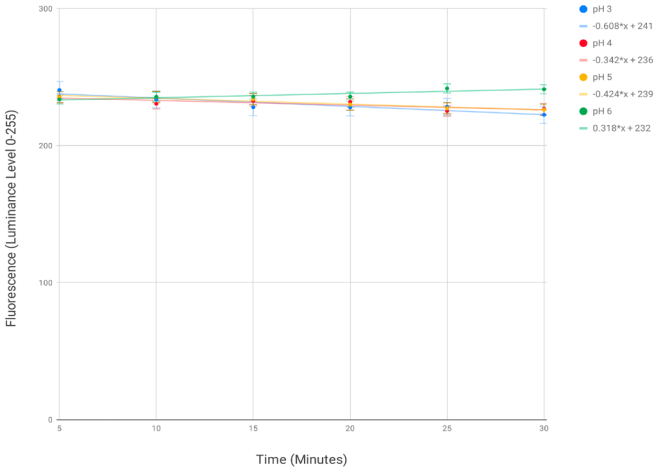

Figure 4: Fluorescence of Fluorescein over time when pH was changed. Fluorescein fluorescence decreases with decreasing pH, as depicted by the lines with a negative slope for pHs 3 through 5. There were three trials for every pH level, including pH 6, every 5 minutes for 30 minutes, and the averages of the trials were used for the six data points. The error bars represent the standard deviation of the mean of all three trials, showing how much the data varied at each point.

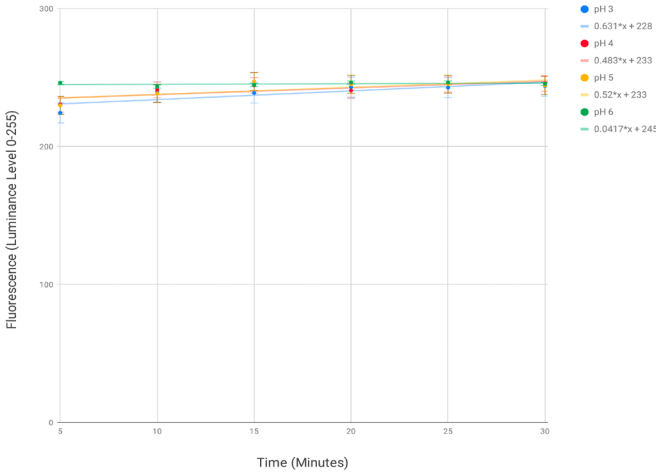

Figure 5: Fluorescence of Fluorescein over time after return to pH 6. Fluorescein fluorescence increases with increasing pH, as depicted by the lines with a positive slope for pHs 3 through 5. There were three trials for every pH level, including pH 6, every 5 minutes for 30 minutes, and the averages of the trials were used for the six data points. The error bars represent the standard deviation of the mean of all three trials, showing how much the data varied at each point.

was a result of error (**Figure 5**).

In the ANOVA analysis, both the null hypothesis that changing the pH of fluorescein would not cause the fluorescence to decrease and the null hypothesis that returning the pH of fluorescein to 6 would not cause the fluorescence to return to its original brightness were rejected. The F distribution values for the fluorescence of fluorescein with the changed pHs in **Table 1** and for the fluorescence of fluorescein after the pH was returned to pH 6 in **Table 2** were above 3.11, rejecting the null hypothesis. This indicates that the fluorescence of fluorescein decreased when the pH decreased and increased when the ph was returned to 6.

Fluorescence of Fluorescein Does Not Change with a Change in Temperature

The fluorescent intensity of fluorescein was not affected by any of the experimental temperatures compared to baseline. Like GFP, there were three experimental temperatures (37 °C, 39.5 °C, and 41.1 °C) because they were in the physiological range. Fluorescein was incubated at these three temperatures for 30 minutes. The fluorescent intensity of fluorescein was expected to drop during these 30 minutes. Instead, the slopes of all the trendlines are extremely small and close to zero, signifying that the change in temperature did not affect the fluorescence of fluorescein (**Figure 6**).

The null hypothesis that changing the temperature of fluorescein would not cause its fluorescence to decrease was not rejected. For fluorescein temperature levels of 37, 39 and 41 °C, the F distribution numbers were lower than 3.11, so the null hypothesis is not rejected, indicating that it could be true and decreasing the temperature does not have an effect on the fluorescence of fluorescein (**Table 1**).

DISCUSSION

In this study, we created a sensor for dehydration that uses changes in fluorescent intensity of a fluorophore to detect changes in pH and temperature. Since dehydration changes temperature and pH of the skin, the fluorescent intensity of the fluorophores GFP and fluorescein were observed at varying temperature and pH. The fluorophores were also tested to see if their fluorescent intensity returned to baseline brightness when their environment returned to the physiologic pH and temperature. This would mean that the sensor could be reusable, making it more cost-effective and therefore a better sensor.

The results for GFP indicated that increasing the temperature of the environment had no effect on its fluorescent intensity. The null hypothesis in the ANOVA analysis was not rejected, indicating that it could be true and decreasing the temperature did not affect the fluorescence of GFP. However, decreasing the pH of the GFP solution caused its fluorescence to decrease. When the pH of the environment was returned to pH 6, the fluorescent intensity of GFP did not return to its baseline brightness. The null hypothesis for when the pH was changed was rejected, meaning that it is not true and decreasing the pH of GFP decreases its fluorescence. However, the null hypothesis for when the pH was neutralized was not rejected, suggesting that it could be true and neutralizing the pH does not affect the fluorescence of GFP. This meant that GFP could be used in a dehydration test, but it would not be reusable.

One plausible explanation for why temperature changes did not affect the fluorescent intensity of GFP was that the range of temperatures tested was small (37 °C to 41.1 °C). However, the range of temperatures must be relevant to the

human body. When the pH of the GFP solution returned to 6, the fluorescent intensity of GFP might not have returned to its baseline brightness because the conformational change caused by low pH was too large.

Similar to the results for GFP, the results for fluorescein indicated that increasing the temperature had no effect on its fluorescent intensity. Likewise, the null hypothesis was not rejected, indicating that decreasing the temperature did not affect the fluorescence of fluorescein. The temperature might not have affected the fluorescent intensity of fluorescein for the same reason as GFP: the range that was tested was not very large because the range of temperatures possible in the human body is not very large. On the other hand, the fluorescence of fluorescein also decreased when the pH decreased. The null hypothesis was, similarly, rejected, meaning that decreasing the pH caused the fluorescence of fluorescein to decrease. Unlike GFP, when the pH of the environment returned to pH 6, the fluorescent intensity of fluorescein returned to its original brightness. The null hypothesis from the ANOVA analysis was rejected, indicating that it is false and the returning the pH to 6 caused the fluorescence to return to its original intensity. Thus, a reusable sensor containing fluorescein may be possible. In dehydrated patients, the pH of their sweat will change, causing the fluorescein brightness inside of the wearable sensor to decrease. When the dehydrated person drinks water after seeing the sensor, their skin pH would return to pH 6 and cause the fluorescent intensity to return to its original brightness, signaling that the person is now rehydrated.

Unlike pH, changing the temperature of both fluorescein and GFP does not change the fluorescence, meaning that they can't be used in a sensor where the factor affecting fluorescence is temperature. However, a sensor based on skin temperature would not be reliable anyway, because body temperature and skin temperature are different, so skin temperature cannot be a reliable indicator of dehydration. Although body temperature while exercising may be 39.5 °C, skin temperature may not be the same. The skin temperature of an athlete is constantly changing depending on activity level and perspiration. Furthermore, one of the symptoms of heat exhaustion is cold and clammy skin, which was not tested. Therefore, skin temperature is an unreliable measure of dehydration. Since our results indicate that skin temperature does not affect the fluorescence of either fluorophore, skin temperature is not an effective indicator of dehydration.

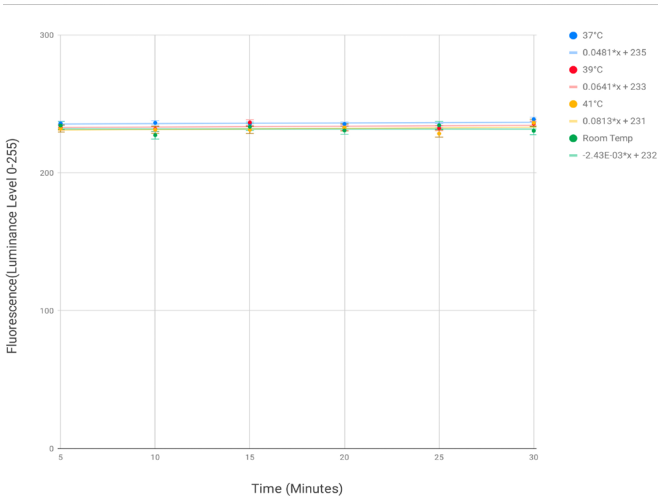

Figure 6: Fluorescence of Fluorescein over time with changed temperatures. Fluorescein fluorescence does not change with increasing temperature, as depicted by the horizontal lines for temperatures 37 °C, 39.5 °C, and 41.1 °C. There were three trials for every temperature, including room temperature (23 °C), every 5 minutes for 30 minutes, and the averages of the trials were used for the six data points. The error bars represent the standard deviation of the mean of all three trials, showing how much the data varied at each point.

Further work on this project would include testing whether changes in pH could have the same effect on fluorescein in a gel. A gel is a mass of liquid where the particles are spread throughout the system evenly. Gels are often used to encase proteins and chemicals. For example, doctors use protein-based hydrogels for tissue engineering and repair because of its structural properties (13). Putting fluorescein in a gel would be the next step in producing a useful sensor, as this would allow the fluorescein to be attached to someone's skin. This experiment would determine whether the required pH changes occur in human subjects. As the environment of the hydrogels is different than water-based buffers, the first step would be to test whether fluorescein can retain its function in this environment. However, one consideration when making this sensor is that some people have a naturally low skin pH, which would make the sensor react even when the person is not dehydrated. In order to solve this problem, before testing a sensor on a person, the pH of their skin must be checked to ensure that the baseline pH is 6.

After testing different temperatures and pHs on both fluorescein and GFP, fluorescein appears to be the most suitable candidate for a dehydration sensor. At the time of writing, the cost of pure fluorescein is almost 150 dollars less than the cost of the materials needed to collect GFP, making it much cheaper. Since GFP can only be used within a short time frame after purification, while fluorescein is more stable, it is even more cost-effective. A sensor containing fluorescein would also be reusable as its fluorescent intensity returns to its original brightness when the pH returns to pH 6, which is not true of GFP. Additionally, the baseline fluorescence of fluorescein is brighter than GFP, making loss of signal during dehydration more obvious. Not only is the baseline fluorescence brighter than GFP, but fluorescein loses its color as well when the pH decreases, making the sensor visible without special equipment. This research suggests a possible cost-effective, reusable solution to prevent severe dehydration, especially in high-risk populations such as the elderly, athletes, and young children.

MATERIALS AND METHODS

Fluorescein Sample Collection

Fluorescein was extracted from a neon yellow highlighter by diluting the ink in 5 mL of water. Fluorescein is the primary ingredient in the ink of these highlighters, and the other ingredient is a soluble binder that prevents the ink from bleeding through pages. The top of the highlighter was removed with pliers in order to get to the tube with the ink in it. This tube was placed into the water until all of the ink had come out and the solution was a mix between fluorescein, the binder, and the water. Neither the water nor the binder, however, is fluorescent, so they would not have an effect on the fluorescence of the mixture. Thus, only the fluorescence of fluorescein was measured in our data.

GFP Sample Collection

A 6 mL liquid culture of HB101 E. Coli with pGLO (a plasmid containing the gene for GFP), ampicillin, and arabinose was grown for 48 hours. After centrifuging the bacteria for 3 min, the pellet was resuspended in 1 mL of Tris-EDTA (TE) Buffer with 0.1 mM EDTA. Then 40 µL of lysozyme was added, and the mixtures were left at room temperature for 30 minutes to lyse the bacteria. To separate the protein from the rest of the bacteria, the lysate was centrifuged for 10 minutes at 16,000 xg. The supernatant was removed and 250 µL of ammonium sulfate (4 M) was added.

The hydrophobic interaction chromatography (HIC) column was equilibrated using 10 column lengths (5 cm) of 2 M ammonium sulfate. The mixture of the supernatant and 4 M ammonium sulfate was added to the HIC column. The column was washed using 1 mL of wash buffer (1.3 M ammonium sulfate) and the GFP was removed from the column by adding 1 mL of elution buffer (TE Buffer).

Temperature Data Collection

Twelve clear tubes, each containing 70 µL of fluorescein, were prepared to determine fluorescence at four temperatures (n=3): room temperature (23 °C), 37 °C, 39.5 °C, and 41.1 °C. One tube was maintained at room temperature as a control.

The tubes were photographed at intervals of 5 minutes for 30 minutes. After 30 minutes, the four temperature groups were moved to room temperature. Their recovery was photographed every 5 minutes for 30 minutes.

The process above was repeated with the GFP samples.

pH Data Collection

Twelve clear tubes, each containing 70 µL of fluorescein, were prepared to determine fluorescence at four different pHs (n=3): 3, 4, 5, and 6. The pH was changed to the listed pHs using HCl, an acid, and NaOH, a base. One tube was maintained at pH 6 as a control. The pH was measured using pH papers.

The tubes were photographed at intervals of 5 minutes for 30 minutes. After 30 minutes, the four pH groups were neutralized to pH 6 using HCl and NaOH. Their recovery was photographed every 5 minutes for 30 minutes.

The process above was repeated with the GFP samples.

Image Analysis

All photos were taken while the tubes were being exposed to UV light. Our photo booth consisted of a black blanket covering the photographer and the samples to minimize the ambient light. The camera used the same exposure settings for all samples.

The change in fluorescence was measured using ImageJ. ImageJ records the brightness of the light from an image. The results were put into a line graph using Google sheets.

The fluorescein samples were too bright to be measured directly by ImageJ, so each image was identically processed to lower the brightness and collect quantifiable data.

Using Photoshop, the image was duplicated as a new layer. The copy was screened, then blurred using a Gaussian Blur filter set to 5.0 pixels. The interaction mode between the copy and the original image was changed from Normal to Multiply. The image was flattened by combining both the original and duplicate layers into one layer. This process reduced the brightness level to within the luminosity range of 0 to 255 for ImageJ analysis.

ANOVA Analysis

An ANOVA analysis was conducted on the fluorescence of both fluorophores at each temperature and pH.

For our data set in the ANOVA analysis, the degrees of freedom in our numerator was 5 and the denominator was 12, giving us a Critical F distribution value of 3.11, $F(5, 12) = 3.11$. If our analysis showed that the F distribution of the data set was less than 3.11 the null hypothesis would not be rejected. If the F distribution were greater than or equal to 3.11, the null hypothesis would be rejected, indicating that our hypothesis was correct.

ACKNOWLEDGEMENTS

I would like to thank Mrs. Belinda Lowe-Schmahl and Schmahl Science Workshops for providing equipment and their lab to work in, and Dr. Louise Pennycook for mentoring my work in the lab.

REFERENCES

1. M.A., Peter Crosta. “Dehydration: Symptoms, Causes, and Treatments.” Medical News Today, MediLexicon International, 20 Dec. 2017, www.medicalnewstoday.com/articles/153363.php.
2. Ghose, Tia. “Why Is Water So Essential for Life?” LiveScience, Purch, 29 Sept. 2015, www.livescience.com/52332-why-is-water-needed-for-life.html.
3. Baker, Lindsay B. “Sweating Rate and Sweat Sodium Concentration in Athletes: A Review of Methodology and Intra/Interindividual Variability.” Sports Medicine, vol. 47, no. S1, 2017, pp. 111–128., doi:10.1007/s40279-017-0691-5.

4. Shirreffs, S. M. “Hydration: Special Issues for Playing Football in Warm and Hot Environments.” Scandinavian Journal of Medicine & Science in Sports, vol. 20, 2010, pp. 90–94., doi:10.1111/j.1600-0838.2010.01213.x.
5. Godek, S F., et al. “Sweat Rate and Fluid Turnover in American Football Players Compared with Runners in a Hot and Humid Environment * Commentary.” British Journal of Sports Medicine, vol. 39, no. 4, 2005, pp. 205–211., doi:10.1136/bjsm.2004.011767.
6. Trautner, Barbara W, et al. “Prospective Evaluation of the Risk of Serious Bacterial Infection in Children Who Present to the Emergency Department with Hyperpyrexia (Temperature of 106 Degrees F or Higher).” Pediatrics, U.S. National Library of Medicine, July 2006, www.ncbi.nlm.nih.gov/pmc/articles/PMC2077849/.
7. Herrmann, Franz, and Leona Mandol. “Studies of Ph of Sweat Produced by Different Forms of Stimulation.” Journal of Investigative Dermatology, Elsevier, 7 Jan. 2016, www.sciencedirect.com/science/article/pii/S0022202X15486798.
8. Findlay, Kim. “Green Fluorescent Protein(GFP).” What Is Electron Microscopy?, John Innes Center, www.jisc.ac.uk.
9. Campbell TN, Choy FYM, “The Effect of pH on Green Fluorescent Protein: a Brief Review”, Center for Environmental Health, Department of Biology, University of Victoria, Box 3020 STN CSC Victoria, British Columbia V8W 3N5 Canada. Molecular Biology Today (2001) 2(1): 1-4.
10. Zhang C, Liu MS, Xing XH. “Temperature influence on fluorescence intensity and enzyme activity of the fusion protein of GFP and hyperthermophilic xylanase.” Appl Microbiol Biotechnol. 2009 Sep;84(3):511-7.
11. PubChem. “Fluorescein.” National Center for Biotechnology Information. PubChem Compound Database, U.S. National Library of Medicine, 26 Mar. 2005, pubchem.ncbi.nlm.nih.gov.
12. Zhu, H., Derksen, R., Krause, C., Fox, R., Brazee, R., and Ozkan, H., "Fluorescent Intensity of Dye Solutions under Different pH Conditions," Journal of ASTM International, June 2005, Vol. 2, No. 6.
13. Schloss, Ashley C et al. “Protein-Based Hydrogels for Tissue Engineering” Advances in experimental medicine and biology, vol. 940 (2016): 167-177.

Article submitted: April 11, 2019
Article accepted: May 6, 2019
Article published: December 10, 2019

JOURNAL OF
EMERGING INVESTIGATORS

Article

Investigation of everyday locations for antibiotic-resistant bacteria in Cambridge, Massachusetts

Ethan Maggio[1,2], Oliver Price[2], Suzanne Bastien[2], and Caroline Palavicino-Maggio[3]

[1] Boston College High School, 150 Morrissey Blvd, Boston, Massachusetts 02125

[2] St. Peter's School, 96 Concord Ave, Cambridge, Massachusetts 02138

[3] Harvard Medical School, 220 Longwood Ave, Boston, Massachusetts 02115

SUMMARY

As a society we have become dependent on the use of antibiotics for the treatment of bacterial infections. However, bacteria can evolve in ways of reducing the effectiveness of antibiotics and become antibiotic-resistant. The main objective of our project was to determine if any of the bacteria collected from various locations in Cambridge, Massachusetts grew in the presence of an antibiotic. To test our hypothesis, we collected bacterial samples from five different everyday trafficked locations in Cambridge, MA. These locations included the Harvard MBTA subway station (T-station), Fresh Pond Park water fountain, the button of a traffic light, a Cambridge resident's sneaker bottom and cell phone. Then, we asked if any of these bacteria would grow in the presence or absence of ampicillin. We observed an increase in the growth of bacterial colonies in samples obtained from the Harvard T-station, water fountain, traffic light, bottom of the shoe, and cell phone screen. However, no colonies were present in the antibiotic dish except for the bacterial sample obtained from the Harvard T-Station sample, it grew bacterial colonies in the presence of ampicillin.

INTRODUCTION

Bacteria are found almost everywhere, from thermal vents in the ocean to the human digestive tract (1). Today, there are more and more microbes that are becoming resistant to ampicillin (2). We need to be cautious what children and young adults touch in order to avoid diarrheal diseases (3). Many studies have shown that the risk of contracting an infectious illness is reduced when people are consistent with washing their hands in public settings. There is then less of a risk of contracting an infectious illness (4).

Our society is dependent on the use of antibiotics for the treatment of bacterial infections, such as ear infections, strep throat and meningitis (5). Recently, there has been a lot of media attention around antibiotic-resistant bacteria. Antibiotic resistance occurs when bacteria genetically change in a way that reduces the effectiveness of antibiotics designed to cure or prevent infections (5). Once resistance has been acquired, bacteria survive and continue to multiply, even when the antibiotic treatment is administered, causing more harm (6). For our study, we wanted to understand what this meant and how to test for it, since certain bacteria can cause sepsis and other horrid diseases (5). Therefore, our main objective was to determine if any of the bacteria collected from various locations in Cambridge, Massachusetts, grew in the presence of an antibiotic. We asked, "If we collect bacteria from five different commonly traffic locations in Cambridge, will any of them have an antibiotic resistant response when exposed to the common antibiotic ampicillin?" We decided to use ampicillin because it is a penicillin used to treat a wide variety of diseases such as ear infections, stomach infections, bladder infections and even meningitis, which can be deadly (7). It is also the most commonly used antibiotic treatment, usually from a prescription by a doctor. We hypothesized that there will not be any antibiotic-resistant responses.

The five locations in Cambridge that we chose were: a push-button for a street cross-walk, a water fountain in Fresh Pond Park, a cell phone screen, the bottom of a sneaker that had walked around Cambridge, and an electric subway escalator railing from The Massachusetts Bay Transportation Authority, also commonly known as T-station, at Harvard Square (Harvard T-Station). After collecting samples from these five locations, we tested the ability of the bacteria to grow in the presence of ampicillin. Our results showed that most of the bacteria samples were not resistant to the ampicillin. However, one sample did contain some antibiotic-resistant bacteria. With our study, we hope to make the public more aware of the potential risk of contracting a disease from bacteria lurking right in our own neighborhood.

RESULTS

We employed multiple controls in this experiment to ensure that there was no bacterial contamination of the plates, water, or antibiotic used for the bacterial growth assay. Our first control consisted of a petri dish with just agar and no bacterial sample. This control informed us that our agar plates were not contaminated with any previous bacteria and that what we saw in our dishes were the actual bacteria from the original source. Our second control consisted of a petri dish with agar and the sterile water we boiled without any bacterial sample. Our third control was a petri dish containing agar and the ampicillin antibiotic only. This control indicated to us that the antibiotic alone did not induce any type of bacteria growth. After 24 hours of incubation, we observed that no bacterial colonies were present in any of our three controls (**Figure 1**).

After a 24-hour incubation period, we observed more than 250 colonies grew in the absence of ampicillin from the

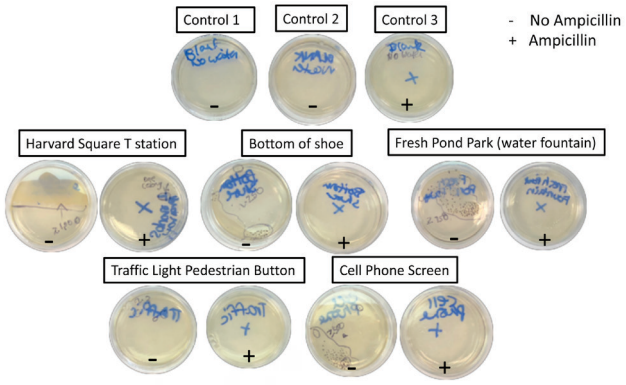

Figure 1. Representative images of samples after 24-hour incubation. Samples were incubated for 24 hours at 37ºC in petri dishes with LB Agar made with (+) or without (-) ampicillin.

sample we scraped off of the bottom of the shoe (**Figures 1 and 2A**). These colonies were white and opaque, with a circular form and a convex elevation; its margin was entire with a smooth surface (**Table 1**). No bacterial colonies grew in the presence of ampicillin (**Figures 1 and 2B**).

The sample we obtained from the Harvard T-Station's escalator rail grew more than 1,000 bacterial colonies after 24 hours of incubation (**Figures 1 and 2A**). These colonies were irregular in shape but smooth in their elevation form; although their color was white-yellowish, they were still translucent with a smooth elevation and glistening surface (**Table 1**). We also found a large bacterial colony that grew in the presence of the ampicillin antibiotic, suggesting that these bacteria were resistant to ampicillin (**Figure 2B**). This colony's characteristic was very similar to the bacterial colonies that grew in the petri dish treated with no antibiotic. We repeated the same experiments at a later time, and similar results were obtained (**Figure 3**).

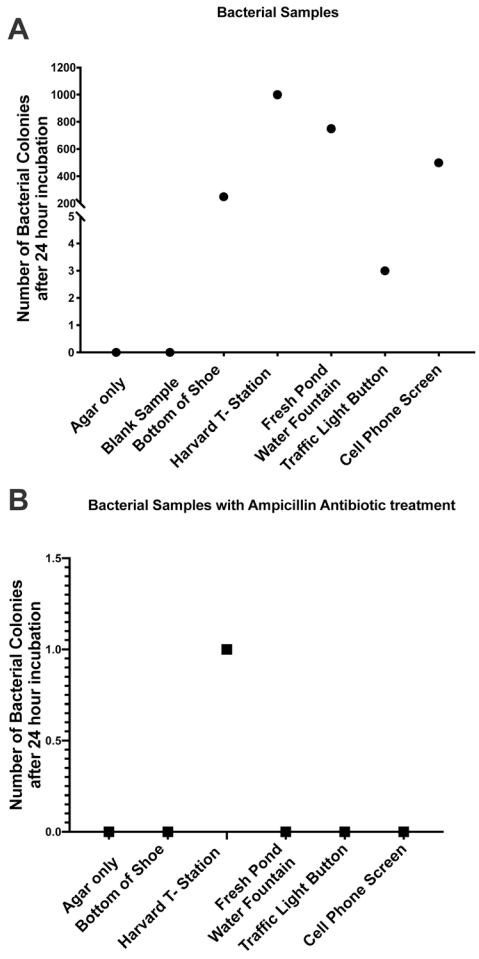

Figure 2: After 24-hour incubation (A) number of bacterial colonies in petri dishes with LB Agar without (-) ampicillin and (B) the number of bacterial colonies in the presence of ampicillin.

After 24 hours of incubation, we observed over 700 bacterial colonies grown from the sample taken from the water fountain at Fresh Pond Park (**Figures 1 and 2A**), but no colonies were observed in the presence of ampicillin (**Figure**

Table 1. Bacterial Colonies and their Characteristics (see ref. 8 for criteria)

	Ampicillin	
Bacteria Collection Sites	-	+
Bottom of Shoe	Characteristics: Form - circular Elevation - convex Margin - entire Surface - smooth Opacity - opaque Chromogenesis - white	None observed
Harvard Square – T-Station's Escalator	Characteristics: Form - irregular Elevation - flat Margin - undulate Surface - glistening Opacity - translucent Chromogenesis - white yellowish	Characteristics: Form - irregular Elevation - flat Margin - undulate Surface - glistening Opacity - translucent Chromogenesis - white yellowish
Fresh Pond Water Fountain	Characteristics: Form - irregular Elevation - raised Margin - lobate Surface - glistening Opacity - iridescent Chromogenesis-white yellowish	None observed
Traffic Light Button	Characteristics: Form - circular Elevation - flat Margin - entire Surface - dull Opacity - opaque Chromogenesis- white	None observed
Cell Phone Screen	Characteristics: Form - filamentous Elevation - umbonate Margin - filiform Surface - rugose Opacity - opaque Chromogenesis - white	None observed

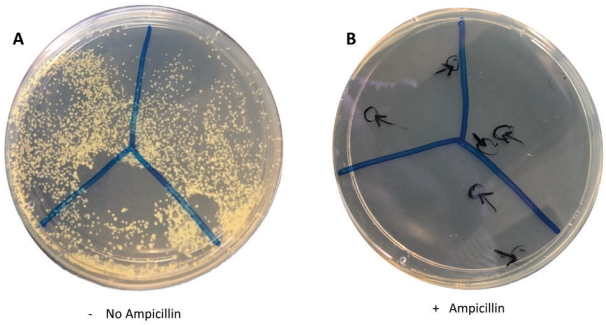

Figure 3: Representative images of samples obtained from Harvard T-Station in August 2019 after a 24-hour incubation at 37°C in petri dishes with LB Agar made with (+) (B) or without (-) (A) ampicillin.

2B). These colonies were white-yellowish and had an irregular, raised form with a glistening and iridescent clarity (**Table 1**).

In the sample obtained from the traffic light, only three colonies grew overnight; no colonies grew in the presence of ampicillin (**Figures 1 and 2**). These colonies were white, circular in form, flat in elevation and had a dull-looking surface (**Table 1**).

We observed that, after an overnight incubation, more than 500 bacterial colonies were present from a sample that we took off of the cell phone screen; no colonies were observed in the presence of ampicillin (**Figures 1 and 2**). These colonies were white and opaque. They were also filamentous in form, and their elevation was umbonate, meaning that although the colonies looked convex, they were also bumpy. Its margin was also different from the bacterial colonies in that it had a filiform shape (**Table 1**).

DISCUSSION

After collecting samples in five commonly trafficked locations, we tested whether or not the bacteria present in these samples would grow in the presence of ampicillin. Without the presence of ampicillin, we observed growth of bacteria colonies obtained from all the samples in our experiment, except for our controls.

The number of bacteria colonies that grew from the sample obtained from the bottom of the shoe was not surprising, since we walk around everywhere and likely pick up a large number of bacteria. However, the samples taken from the electric banister at the metro T-station grew five times the number of bacterial colonies, with an average of more than over 1,000 colonies grown from each sample. The electric handrail is touched every day by many people, which may explain the number of colonies. We also observed that over 700 colonies grew from the sample obtained from the water fountain at Fresh Pond Park. This was not unsurprising, as there are a large number of people in the park every day. Interestingly, the least number of bacterial colonies we observed came from the sample scraped off of the button at the traffic light on Concord Avenue opposite Wheeler Street. This may be because a lower number of people use the cross street or use the traffic assistance button. We observed that more than 500 bacterial colonies grew overnight from a sample we took off of the cell phone screen. This was not surprising because, although we wash our hands often, our phones are touched no matter if we did wash them or not. Our phones are also placed down on surfaces around everyday places.

Our results showed that most of the bacteria samples were not resistant to ampicillin. However, one of our samples did contain antibiotic-resistant bacteria. We observed the growth of large colonies in the ampicillin plate. This result indicated that the sample taken from the electric subway escalator handrail was resistant to ampicillin. We repeated these experiments at a later time and obtained similar results of bacteria growth in the presence of ampicillin.

We concluded that the subway rail samples would likely have the most amount of hand contact on a daily basis. Therefore, it had the highest chance of containing a bacterium that was resistant to the common ampicillin antibiotic. Another reason, for bacterial growth in the presence of an antibiotic, may be that the colonies that grew in the petri dish were not a strain that can be affected by ampicillin.

Furthermore, testing of various sites in Cambridge and other cities could prove useful in knowing which types of antibiotics can be most valuable in treating everyday bacterial infections, or even more deadly infections. This information would make the public more aware of how to avoid touching public surfaces and reinforce the importance of handwashing.

METHODS

Sample Collection

Sample collections took place in Cambridge, MA in May 2018 and August 2019. Samples were obtained from the Harvard Square-T-station's escalator rail, a water fountain at Fresh Pond Park, a cell phone, the bottom of a shoe, and a traffic light button (on Concord Ave., opposite of Wheeler St.). Sites were evenly swabbed using a sterile 10 µL innoculating loop and carefully submerged into a 1.5 mL Eppendorf tube containing 1 mL of sterile Molecular Grade Water (G Biosciences Cat. #786-293).

Petri Dish Preparation

Ten grams of LB Broth, Miller (BD Bioscience Cat. #244620) and six grams of agar (ICN Cat. #100262) were added to a final volume of 400 mL of deionized water and then autoclaved. Once cooled to 55˚C at room temperature, 10-11 mL of LB solution was poured evenly into each sterile 100 x 20 mm (Corning Cat. #353003) petri dish. Petri dishes were then stored at 4oC until use. For the agar-ampicillin plates, the same procedure was repeated with the exception that we added 5 mL of 10 mg/mL ampicillin (Sigma-Aldrich Cat. #69-53-4).

Plating the Samples

Five hundred µL of each sample was pipetted onto the

LB Agar petri dish, and another 500 µL of the same sample was pipetted onto an LB Agar petri dish plate containing ampicillin. Plates were then incubated at 37˚C overnight. After 24 hours, we removed the samples from the incubator and counted bacterial colonies.

Received: April 14, 2019
Accepted: December 5, 2019
Published: December 12, 2019

REFERENCES

1. Täubel, M. et al., "The occupant as a source of house dust bacteria". *The Journal of allergy and clinical immunology*, vol. 124, no. 4, 2009, pp.834–40.e47.
2. Heinz, E., "The return of Pfeiffer's bacillus: Rising incidence of ampicillin resistance in Haemophilus influenzae". *Microbial genomics*, vol .4, no. 9, 2018
3. Watson, J. et al., "Child's play: Harnessing play and curiosity motives to improve child handwashing in a humanitarian setting". *International Journal of Hygiene and Environmental Health*, 2018, pp.1–6.
4. Aiello, A.E. et al., "Effect of hand hygiene on infectious disease risk in the community setting: a meta-analysis". *American journal of public health*, vol. 98, no. 8, 2008, pp.1372–1381.
5. Laxminarayan R, et al., "Antibiotic resistance—the need for global solutions". *The Lancet. Infectious diseases*, vol. 13, 2013, pp. 1057–98.
6. Koch, A.L., "Control of the bacterial cell cycle by cytoplasmic growth". *Critical reviews in microbiology*, vol. 28, no. 1, 2002, pp.61–77.
7. Centers for Disease Control and Prevention, Jan. 2019, https://www.cdc.gov
8. Madigan, M. T., et al. 2015. *Brock biology of microorganisms* (15th ed.). Boston: Pearson.

JOURNAL OF
EMERGING INVESTIGATORS

Article

An Analysis of the Mathematical Accuracy of Perspective in Paintings

Sidak Singh Grewal[1], Kanwarpreet Grewal[2]
[1] Lotus Valley International School, Noida, India
[2] Express Greens, Sector 44, Noida, India

SUMMARY

Art is a great human endeavor and artists, since time immemorial, have tried to capture the beauty of what lies around us in their canvases. But does art accurately depict real-world objects in terms of their sizes and proportions? A particular area where artists have been challenged is in the depiction of the three-dimensional world on a two-dimensional canvas. This is known as perspective. Mathematicians have done a lot of work to understand perspective and this field of mathematics is known as projective geometry.

Our hypothesis was that there are mathematical inaccuracies of perspective in artists' paintings that we are unable to detect with our naked eyes. Errors of perspective may be more tolerated as the distance from the eye to the object in the painting increases. We wanted to understand the degree of mathematical inaccuracy in several famous paintings and then draw conclusions about the limits of our perception of depth.

We have used a mathematical method from the world of projective geometry, which is called cross-ratio, to analyze paintings for accuracy of perspective. Cross-ratio is measured by using four fixed points on a straight line as reference; this value is always the same irrespective of the viewing angle or the viewing distance. We took structures/buildings in paintings as our four points and compared the painting's cross-ratio to that of a photograph of the same building. For our research we have taken three famous paintings made by different artists and measured their accuracies. From our analysis we concluded that there are significant errors in perspective in these paintings and the errors increase as the structures in the paintings recede from the viewer.

INTRODUCTION

Perspective is the technique to create the illusion of three-dimensional (3D) objects on a two-dimensional (2D) plane in art. The basic idea of perspective is that objects of a certain size decrease in size as they move farther away from the point of sight. Another technique employed by artists is called foreshortening. Foreshortening is the illusion that objects appear distorted if they are placed at an angle to the observer. For example, a circle may appear as an ellipse when seen from the side. Additionally, professional artists use the technique of vanishing points in their paintings. There may be one or more vanishing points on a plane. We treat these points as infinity, and all parallel lines meet at these point(s). The placement of objects on the painting with respect to these points and their relative sizes determines the visual accuracy of perspective (1–3). Foreshortening and the technique of vanishing points are complementary to each other and are important tools for artists (4).

Brunelleschi was the first to fully understand the importance of perspective and was able to incorporate the mathematical theory of perspective in his paintings in the fifteenth century. He claimed to discover a method to find out an object's dimensions in the picture by finding out the object's distance from the vanishing points. He applied his technique and created both paintings and architectural works. He is famous for creating the Duomo, a famous church in Florence. But many of his paintings that illustrate his techniques of perspective using mathematics have been lost (5,6). So, we asked if there is a way to measure the accuracy of perspective in paintings?

We were able to apply a simple but powerful method that could answer this question. From a branch of mathematics known as projective geometry we found a technique called cross-ratio (7). Cross-ratio is the ratio of four points lying on a line. The distances between these points are inserted into a formula to obtain the cross-ratio of those points.

If you draw lines starting from one of the vanishing points intersecting all four of our special points individually, extend them, and make a different line at any angle, you will get a set of four new points whose cross-ratio will be exactly equal to the previous one (6). As an example, the cross-ratio of the points A, B, C, and D will be exactly same to the cross-ratio of points A', B', C', and D' **(Figure 1A)**.
Cross-ratio is measured by evaluating the following formula:

$$\text{Cross-ratio of points (A, B, C, D)} = \frac{AC \times BD}{BC \times AD}$$

We infer from the above example that we get the same cross-ratio even if we scale the distances between points or change the angle of the line on which our points lie. Therefore, the angle of sight or the scale of the picture has no effect on the cross-ratio of any four points on a line (8).

The applications of the technique of cross-ratio are wide ranging. Scientists have proposed that accidents can

be reconstructed by measuring the cross-ratio using a single camera placed at an angle to the collision (9). Since cross-ratio is invariant with respect to the angle of view, the speed of colliding vehicles can be accurately determined. This technique may be very useful in the field of forensic science. The method of cross-ratio can also be very important in the study of architecture, and it has been used to study the visual perception of old Roman buildings (10).

Figure 1: A) An illustration of the method of cross-ratio. B) The same four points in a room seen from different angles and distances. The cross-ratio of the four points is invariant.

We have used the method of cross-ratio very effectively in our study to analyze how accurate artists have been in their paintings using perspective. But the question remains, which four points in the painting should be analyzed?

Many artists make pictures of buildings or structures of great importance. We used the fact that several of those buildings/structures have survived to this date. Therefore, we have assigned our four points to lie on the buildings. Pillars have been the choice of placement for our points, as they are cylindrical and are typically of constant width to maintain the symmetry of the building. By comparing the cross-ratios of any four points on a building depicted in a painting to the cross-ratio of the same points on the picture of the real building, we could determine the accuracy of the painting to a great degree of detail.

Our hypothesis was that there are mathematical inaccuracies of perspective in the artists' paintings that we are unable to detect with our naked eyes. Errors of perspective are likely to be overlooked as the distance from the eye to the object in the painting increases. We wanted to understand the degree of mathematical inaccuracy in several famous paintings and then draw conclusions about the limits of human perception of depth.

The paintings we analyzed were *The Doge's Palace* and *Riva degli Schiavoni* by Giovanni Canal, *St. Peter's Basilica* by Giovanni Paolo Panini and *The Wedding at Cana* by Paolo Veronese.

RESULTS

Paintings by Giovanni Canal, Giovanni Paolo Panini, and Paolo Veronese were analyzed in order to find out if their depictions of buildings were mathematically accurate.

Before we analyzed the paintings of the masters, we studied the usefulness of the method of cross-ratio by taking a practical example. For this illustration we took photos of the interior of a building at different angles and distances **(Figure 1B).** We then chose four points on a straight line and measured their cross-ratio.

We assigned the end points of the door frames as our special points and named them A, B, C, and D.

We made a Python program to calculate the cross-ratio of the pictures by inserting the distances between the special points. We used a software that enabled us to zoom in the picture and measured the distances accurately and all the pictures were of very high quality, so the details were clear even when zoomed in.

We found out that the cross-ratio of the four points in the three parts measured were 1.32, 1.34, and 1.33 **(Figure 1B).** Thus, the cross-ratios varied just by 1-2% of each other, which was within the uncertainty of measurement in calculating distances.

Therefore, we concluded that the method of cross-ratio is a very reliable and precise method to measure how the relative distances between points change as the image is rotated or resized. We demonstrated that the cross-ratio of any four points on an image is invariant even if we change the angle of observation or distance from the object to the point of observation. Using this method, we could now proceed to measure the precision of several paintings.

Giovanni Antonio Canal

Giovanni Antonio Canal is considered one of the most famous and influential Italian painters in history. He is known for his exceptional interest in painting famous buildings and bridges in Venice and other cities of Italy (11). He painted over 600 paintings including the famous T*he Doge's Palace and Riva degli Schiavoni, The Grand Canal from the Palazzo Vendramin-Calergi towards S. Geremia, View of the Grand Canal, Piazza San Marco* and *The Molo from the Basin of San Marco* (12). He was known for the accuracy and precision in replicating the buildings and all their details in his paintings.

We looked at the accuracy of *The Doge's Palace* and *Riva degli Schiavoni* by comparing its cross-ratio to that of

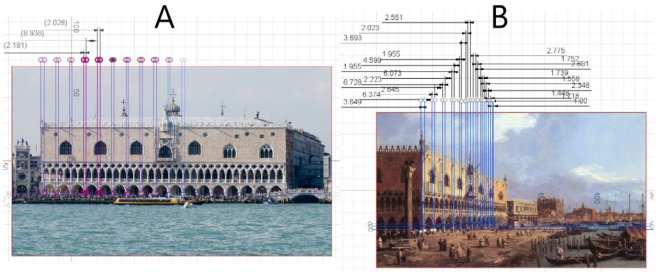

Figure 2: A) A photograph of the *Doge's Palace* in Venice. The pillars are marked to calculate the cross-ratios which will be used to compare with the painting. B) Painting of *Doge's Palace* by Canal on which we have marked the distances between the pillars. The distances are used to calculate the cross-ratio of the points taken four at a time.

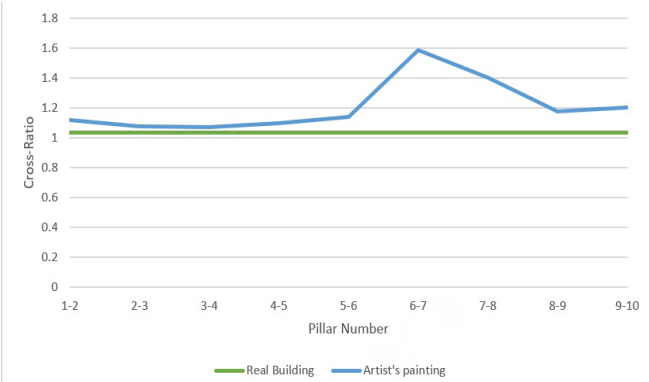

Figure 3: Comparison of cross-ratios of the real building (Doge's Palace) to that of the painting by Canal.

the real building using a photograph **(Figure 2).** By using the software, Autodesk Fusion 360 (13), we marked points on each of the pillars in the Canal painting and measured the distances between them **(Figure 2B)**. This software enabled us to zoom into the picture to measure the distances accurately. In the photograph of Doge's Palace we marked the pillars that we marked in the painting as reference **(Figure 2A)**. We assigned the same special points to each of the pillars of the palace and measured the distances between them. Since we needed four points to measure the cross-ratio, we combined two points of one pillar with two of the subsequent one.

We showed that there is a significant variation of cross-ratios in the artist's painting **(Figure 3)**. The artist was very accurate in the first few pillars and the cross-ratio of the dimensions of his pillars is quite close to that of the real building, but after the fifth pillar the difference in the cross-ratios becomes considerable. The artist was most inaccurate in column six, where he had an error of 47 %. In pillar three the artist was most accurate, where the error was only 3%. On average the error in the painting compared to the actual building was 16.5% **(Figure 3)**.

Giovanni Paolo Panini

Giovanni Paolo Panini was a famous Roman painter and architect in the eighteenth century. He painted many different vistas of Rome, from its architecture to its landscape. He painted many paintings of different Roman monuments including *The Pantheon, A Capriccio of the Roman Forum, The Colosseum, St. Peter's Basilica*, and many more (14). His legacy remains in his precision of replicating Roman buildings on his canvases. We examined the accuracy of his painting, *St. Peter's Basilica*.

We assigned special points to the pillars on the right in the Panini painting and calculated their cross-ratio **(Figure 4)**, as well as those in the real picture taken in *St. Peter's Basilica* **(Figure 5)** (15). We have shown the dimensions of the points on the pillars in the picture which helped us find the cross-ratios. These points correspond to the ones we used to calculate the cross-ratios of Panini's painting.

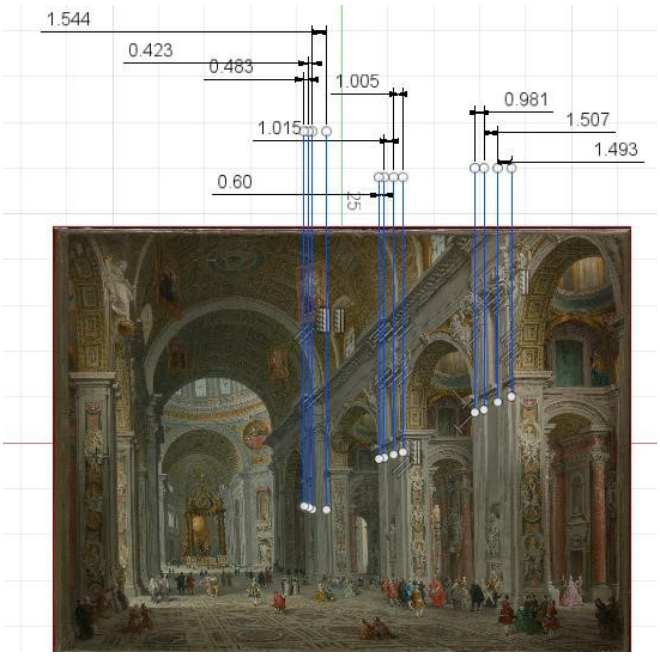

Figure 4: *St Peter's Basilica* painted by Panini. We added the lines to measure distances to calculate cross-ratios.

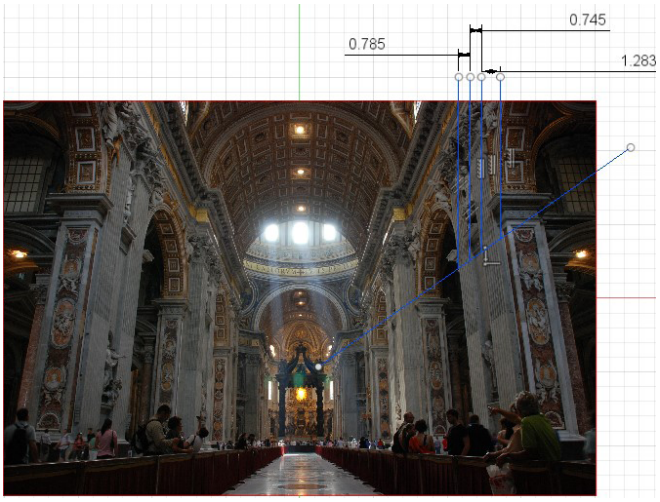

Figure 5: A photograph of *St Peter's Basilica*. We added the lines to measure distances to calculate cross-ratios.

Thus, we saw that the artist was most inaccurate in the dimensions of pillar three, where his error was 16.2% **(Figure 6)**. The artist was slightly more accurate in pillar 1 in which the error was 15.6%. When compared to the real building, the painting has on average 16% error in perspective. But we could see that the cross-ratio jumped significantly from pillar two to pillar three. This variation was about 33% between pillar 2 and pillar 3 **(Figure 6)**. We showed from this analysis that Panini was fairly inconsistent and inaccurate in his painting.

Paolo Veronese

Paolo Veronese was one of the most famous painters in the sixteenth century. He was famous for making large-scale and grand paintings which had a lot of detail (16). Some of his

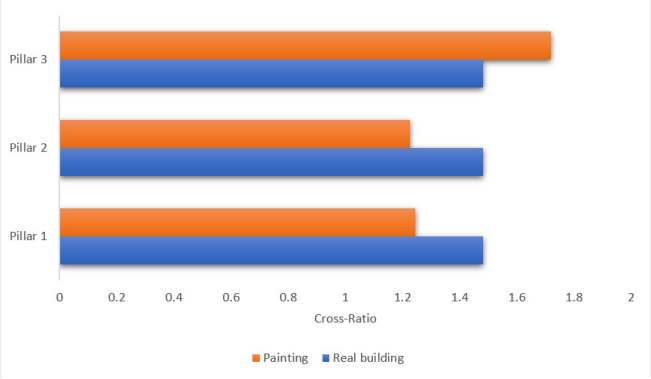

Figure 6: Cross-ratios of Panini's painting compared to a photograph of *St Peter's Basilica*.

famous works are *Venus and Adonis*, *The Wedding at Cana, The Vision of St. Helena* and *The Resurrection of Christ*. We wanted to analyze the accuracy of *The Wedding at Cana* (17). This painting is the largest painting in the Musée du Louvre in Paris. We used the technique of cross-ratio described earlier to determine the accuracy of the artist's painting. We have depicted the painting and the points that we used to calculate the cross-ratios and the distances between them **(Figure 7)**.

The building depicted in the painting is not real, but is based on the imagination of the artist. We assumed that the artist completely mirrored the marble designs on both sides, and therefore the cross-ratio should remain same. We calculated the cross-ratio of the repeated design on top of the columns of the building at the left and the cross-ratio of the repeated design on top of the columns of the building at the right. We concluded from this data that the maximum variation between the two similar designs on both sides was 26%, and on average the cross-ratios of both sides agreed with each other by 95% **(Figure 8)**. Therefore, in this painting we could also see inconsistency in the cross-ratio on both sides.

DISCUSSION

With the analyses of the cross-ratio in three paintings we showed that all of these artists became less accurate in perspective as their work receded to one of the vanishing points. All the above-mentioned paintings show inaccuracies of approximately 3% to 16% in the foreground, and it becomes worse as the objects in the painting recede into the horizon. All the paintings were painted by popular artists and have been visited by millions of people. One finds it strange that most viewers haven't noticed inaccuracies in these paintings. We conclude that the human eye ignores inaccuracies to a certain limit, and in this study we have shown that even inaccuracies of more than 15% in some cases are ignored. However, our study could not find the limit at which the human eye can detect this inaccuracy. We suggest that a more detailed study on this is needed to establish this limit accurately. However, it is clear from our analysis that as the structures in the building recede into the vanishing point the inaccuracies might increase and remain undetected by the artist and the viewers. As the distance from the eye increases, the details of a structure's features become so minute and blurred that the human eye completely ignores the errors in perspective.

The mathematical idea of cross-ratio is very useful as it is invariant with respect to the angle of sight. However, we believe that human perception will be impacted as the angle of the viewer changes and will also depend on other factors like size of structures and distance. Therefore, we think that this could be an idea for a future work in which surveys can be conducted to see how people view inaccuracies as buildings are rotated by various angles and are zoomed in/out.

Another idea of a potential future research project could be to rank some selected popular paintings by the accuracy of perspective using the methods presented in this paper. We could then conduct a survey by a set of people and ask them to rank the same paintings by how accurate they appear visually. We could then compare both rankings to find out the correlation between human visual perception and the mathematical accuracy of paintings.

Mathematics is a very useful tool to help us make sense

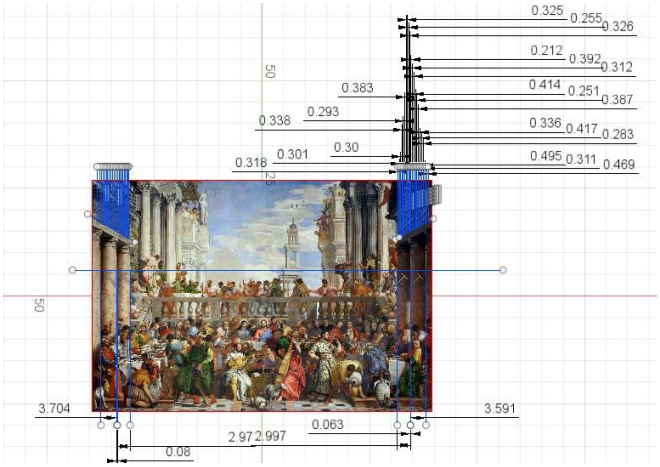

Figure 7: *Wedding Feast at Cana*. We have marked the points we used in calculating the cross-ratios.

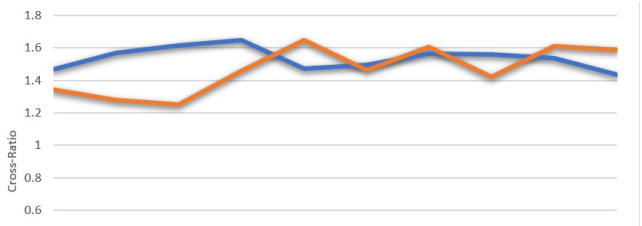

Figure 8: Comparing Cross-ratios of similar patterns in the painting, "*The Wedding at Cana*".

of the world around us. Art is an expression of human emotion and a desire to capture the beauty of the outside world. But artists can make mistakes in capturing the external world because of the limits of what can be seen and perceived by the naked eye. We can use mathematics to find out these mistakes and correct them. Artists can use the techniques of projective geometry when making rough sketches of their paintings so that when the final painting comes out, the results are more accurate and more visually appealing and appear more realistic. In this paper we have successfully used mathematics to find out flaws of perspective that artists tend to make in their work.

MATERIALS AND METHODS

For this research, famous paintings which depicted monuments were selected. The reference image, which could be an actual picture of the same monument, was also searched online. High pixel density images (of at least 300 DPI) were chosen as these could be zoomed in without getting blurred. Once the paintings were selected, they were loaded into Autodesk Fusion 360 software. These images were zoomed in on and prominent structures of the buildings in the images were chosen. Pillars or other symmetric objects which were repeated in the structure were selected. We knew that the method of cross-ratio required four points, but a pillar gives us only two (the left edge and the right edge). So, to get the four points we combined two adjacent pillars in our calculations for the first painting (Canal). In the second painting (Panini), each pillar has four points as each pillar had substructures. Therefore, we did not need to combine pillars. Similarly, in the third painting (Veronese), the patterns on top of the pillars gave us many choices from which we could choose our four points from. Once the images were loaded into the software and the structures to be analyzed chosen, the line tool was used to draw lines on the edges of these structures. These lines were extended parallel to the y axis. The sketch dimension tool was used to find distances between the vertical lines. Once we had four distances, a simple Python program was used to find the cross-ratio of the points. These calculations were repeated for the row of pillars or other symmetric structures. Similar calculations were done for the real image and then the two sets of cross-ratios were compared to determine the accuracy of the painting. We defined the error rate in our calculations as follows:

$$\text{Error \%} = \left| \frac{\text{Cross ratio of four points in the real building} - \text{cross ratio of same points in painting}}{\text{cross ratio of the four points in the real building}} \right| \times 100$$

With respect to the choice of paintings there are some limitations using the technique of the cross-ratio. We required high quality images to be able to determine the cross-ratio accurately. The minimum resolution required for our analysis was 300 DPI. This enabled us to zoom in and accurately measure distances between objects without the details becoming blurred. Another requirement of this technique was that structures in the painting had to be repetitive while retaining their original size and be in a straight line. This enabled us to mark points in a straight line and measure their cross-ratio accurately and then compare this with the cross-ratio of the same structure in a photograph. Thus, this method will only apply to paintings that try to depict the details of buildings and similar structures accurately. This technique will not apply to paintings which are a part of art movements like impressionism, cubism, surrealism and so on, where the artist is not trying to capture the details of structures accurately, but takes artistic liberties with the dimensions of the structures.

Received: June 7, 2019
Accepted: November 21, 2019
Published: December 13, 2019

REFERENCES

1. Weisstein, Eric W. "Perspective." From *MathWorld-A Wolfram Web Resource*, mathworld.wolfram.com/Perspective.html
2. "The Beginning Artist's Guide to Perspective Drawing." *Artists Network*, 28 Sept. 2018 artistsnetwork.com/art-mediums/drawing/learn-to-draw-perspective/
3. Blumberg, Naomi. "Linear Perspective." *Encyclopædia Britannica,* Encyclopædia Britannica, Inc., 17 Mar. 2016, www.britannica.com/art/linear-perspective
4. Blumberg, Naomi. "Foreshortening." *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., 25 Jan. 2019, www.britannica.com/art/foreshortening
5. O'Connor, J.J, and E.F Robertson. "Mathematics and Art-Perspective." *MacTutor History of Mathematics*, University of St Andrews, Scotland, Jan. 2003, www-groups.dcs.st-and.ac.uk/history/HistTopics/Art.html
6. Hyman, Isabelle. "Filippo Brunelleschi." *Encyclopædia Britannica,* Encyclopædia Britannica, Inc., 11 Apr. 2019, www.britannica.com/biography/Filippo-Brunelleschi
7. "Cross-Ratio." *Wikipedia*, Wikimedia Foundation, 11 May 2019, en.wikipedia.org/wiki/Cross-ratio
8. Ardila, Federico. "The Cross-ratio - Numberphile" *Youtube*, Brady Haran, 6th Jul. 2018, www.youtube.com/watch?v=ffvojZONF_A
9. Wong T.W, et al. "Application of cross-ratio in traffic accident reconstruction.", *Forensic Science International*, vol. 235, Feb. 2014, pp. 19-23. doi.org/10.1016/j.forsciint.2013.11.012
10. Migliari Riccardo, and Leonardo Baglioni. "Application of the cross-ratio to the analysis of architecture", *16th International conference on geometry and graphics*, Aug. 4-8, 2014, Innsbruck, Austria, www.academia.edu/32225067/APPLICATION_OF_THE_CROSS-RATIO_TO_THE_ANALYSIS_OF_ARCHITECTURE
11. "Canaletto." *Wikipedia*, Wikimedia Foundation, 20 May 2019, en.wikipedia.org/wiki/Canaletto
12. "The Complete Works." Giovanni Antonio Canal (Canaletto), www.canalettogallery.org/ the-complete-works.html
13. "Fusion 360." *Autodesk,* www.autodesk.com/products/

fusion-360/overview
14. "Giovanni Paolo Panini." *Wikipedia,* Wikimedia Foundation, 25 Feb. 2019, en.wikipedia.org/wiki/Giovanni_Paolo_Panini
15. "St. Peter's Basilica." *Wikipedia*, Wikimedia Foundation, 17 May 2019, en.wikipedia.org/wiki/St._Peter%27s_Basilica
16. "Paolo Veronese." Wikipedia, Wikimedia Foundation, 18 Apr. 2019, en.wikipedia.org/wiki/Paolo_Veronese
17. "The Wedding at Cana." *Wikipedia*, Wikimedia Foundation, 20 May 2019, en.wikipedia.org/wiki/The_Wedding_at_Cana

JOURNAL OF
EMERGING INVESTIGATORS

Article

The effects of confinement on the associative learning of *Gallus gallus domesticus*

Amanda Jaworsky, Jocelyn Reid, Evey Peplowski
Williamston High School Math and Science Academy, Williamston, Michigan

SUMMARY

In this study, we aimed to determine if confinement affects associative learning in chickens. We assessed associative learning ability by training chickens to recognize two plates of cottage cheese: on one plate, the cottage cheese was stained green with no other additives, and on the other plate, the cottage cheese was stained pink and contained methyl anthranilate, a common chemical in bird repellent. After a training period of 7 days, we again presented each chicken with the two plates of cottage cheese and measured the amount of time before the chickens began to consume cottage cheese and the number of pecks the chickens made to each cottage cheese plate during a two minute interval. Two trials were conducted; the first was conducted before the chickens were subjected to confinement and the second was conducted after the chickens were subjected to confinement for 15 hours. We found that the difference in the time lapsed before the chickens began to consume the cottage cheese before and after confinement was significant, though the difference in the number of pecks was not. These results suggest that confinement distresses chickens, as it impairs associative learning without inducing confusion.

INTRODUCTION

For most people in the United States, though chickens are considered profitable to farm, these animals are not considered particularly intelligent (1). However, chickens are very intelligent animals (2). They engage in complicated social hierarchies and have high levels of self-awareness, numerical understanding, and visual capabilities (3, 4). Despite these high levels of intelligence, many chickens are subjected to controversial conditions in factory farms. Factory farms subject chickens to crowded, noxious living conditions; each chicken is given an average living area smaller than a sheet of paper, in which the ambient noise of surrounding chickens is deafening and the buildup of feces can lead to ammonia levels that are harmful to chickens' eyes, throats, and skin (5). Since chickens are very profitable animals to farm, are not considered particularly intelligent, and are not traditional pets in the United States, this abuse is generally unregulated. We hypothesized that chickens subjected to confinement similar to that experienced by chickens held in factory farms would demonstrate signs of distress and struggle to complete associative learning tasks.

As a result, we designed a study to test this hypothesis, in which the independent variable was confinement and the dependent variables were the chickens' tendency to peck the more desirable, green-colored cottage cheese, the amount of time before pecking the cottage cheese, and the number of pecks of each color. The goal of this study was to determine if confinement affects chickens' abilities to learn by association. Our hypothesis was that confinement would cause either an increase in the amount of time it takes the chickens to begin the test or an inability of the chickens to accurately complete the test. After completing our experiment, we found that the amount of time it took for the chickens to begin the test increased after confinement. However, we also found that chickens were equally able to accurately participate in the test after confinement as before.

RESULTS

We hypothesized that chickens would show behavioral differences after being subjected to 15 hours of confinement by either a variation in the average number of pecks or the amount of times that they pecked at the cottage cheese. To investigate this, we compared two groups: the chickens before experiencing confinement, and the chickens after having been confined for 15 hours. Tests for both groups were conducted after the chickens had been fed cottage cheese for approximately 7 days. The median amount of time to peck in the chickens before confinement was 8.5 seconds, while the chickens after 15 hours of confinement showed a median of 35 seconds to peck. This indicated an increase in the amount of time for chickens to peck the cottage cheese after being confined (**Figure 1**). We found this increase in time to peck was significant using the Wilcoxon signed-rank test; at $p = 0.05$ $T^* = 13$ and our data obtained a measure of $T^* = 11$, since $11<13$ we rejected the null hypothesis. Furthermore, we anecdotally observed that, after confinement, chickens were more hesitant and cautious when presented with the plates of cottage cheese.

The control group demonstrated a median of 31.5 pecks directed towards the green cottage cheese and the treatment group demonstrated a median of 30 pecks (**Figure 2**). We found that the apparent decrease in the number of pecks by the treatment group was not significant by Wilcoxon signed-ranks test at significance level $p = 0.05$. We also hypothesized that the number of times the chickens pecked the tainted

Figure 1. Boxplot of seconds elapsed until cottage cheese was pecked.

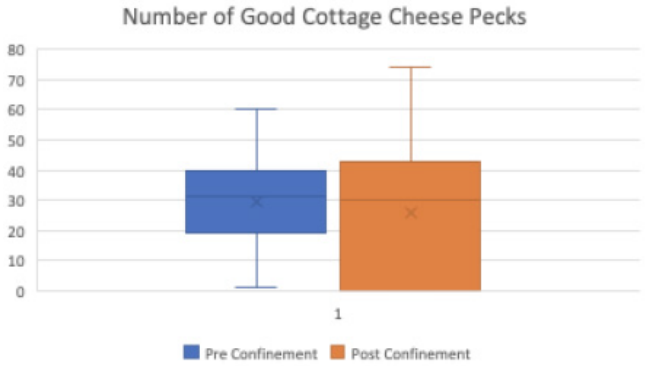

Figure 2. Boxplot of the number of times each chicken pecked the green cottage cheese.

cottage cheese (pink) would increase after the chickens were placed in confinement. However, the chickens showed absolutely no difference in the amount of times they pecked the pink cottage cheese, so we could not run a test on this data.

DISCUSSION

The goal of our research was to measure chickens' color recognition and association capabilities and determine if factory farm conditions, such as confinement, affect these capabilities. We attempted to measure these attributes using various approaches throughout the course of our research. Originally, we planned to condition each chicken to recognize the color red as "good" and the color green as "bad" using a reward system. In this planned experiment, chickens would demonstrate this association by pecking one of two small, colored squares of paper: one green and one red. In order to train them, they would receive a treat (a dried mealworm, or a click from a dog training clicker) when they pecked the red piece of paper, and they would not receive a reward when they pecked the green piece of paper. This set-up was inspired by an experiment we read about in which chicks were tested on their color preference through the use of colored circles laid on the floor (6). This study was able to effectively condition chickens to show color preference, so we were confident that a similar set up would work for our chickens as well.

We would then conduct our test by offering each chicken the two pieces of paper, and measuring how many times they pecked the red colored paper (in hopes of getting a treat) and the amount of times they pecked the green colored paper (incorrectly assuming they would get a treat). The pecks directed towards the red paper would be labeled as "correct" and pecks directed toward the green paper would be labeled as "incorrect." The chickens' scores would be calculated as the number of correct pecks over the total number of pecks. After confining the chickens, we would repeat the test and compare the changes in scores.

This approach was attempted repeatedly with six different chickens throughout the course of six weeks, yet after this time, it became apparent that our approach was not appropriate. We struggled to entice the skittish chickens to peck the paper even once, because they did not find the task appealing. Some of the more intelligent chickens appeared to pick up on the pattern, but even those chickens were unreliable. Ultimately, we could not encourage almost any of the chickens to peck either of the paper squares or even show interest in the test. After running one test run where a chicken did not peck either color for the entirety of the test, we decided to amend our approach. After reaching out to a researcher who worked with our class and has done a lot of biology research with live animals, we decided that using colored foods would be a more promising way for us to study chickens' abilities to learn by association.

We first attempted to use green- and pink-dyed yogurt, and supplemented the pink yogurt with methyl anthranilate, which is commonly used as bird repellent (7). However, on the first day we attempted our new method, the chickens would not even touch the yogurt. After a second failed attempt the following day, we decided we needed to use a different food product. We decided to use cottage cheese, because we thought that cottage cheese would be easier for the chickens to eat due to the curds. Fortunately, the first day the chickens tried the cottage cheese, they ate it without hesitation.

With our new experiment, we began to rethink how our hypothesis would translate to our results. We suspected that placing the chickens into confinement would induce distress and make them less willing to taste the cottage cheese. We also thought that it might lead to increased confusion, causing them to peck both types of cottage cheese even though they had already been trained that the pink cottage cheese would taste bad. We hypothesized that, after confinement, the amount of time before the chickens pecked the cottage cheese would increase, the amount of times the chickens pecked the green cottage cheese would decrease, and the amount of time the chickens pecked the pink cottage cheese would increase.

Our first hypothesis, that the time before pecking the cottage cheese would increase, was supported by our experiment. The previously confined chickens were wholly unwilling to try the cottage cheese until they had investigated

it long enough, unlike in the control test, where they ate the cottage cheese very quickly and without hesitation. The next two hypotheses, that the amount of times the chickens pecked the green would change and that the amount of times they pecked the red would increase, were not supported by our experiments. The first reason that we believe the number of green cottage cheese pecks did not change is due to the fact that several times, the chickens spilled the contents of their provided food and water while they were in confinement. Because the food got spilled, it is likely that when they were tested immediately after being brought out of confinement, they were quite hungry. Although they were more hesitant to eat, once they tasted the untainted green cottage cheese, they were hungry enough to let their hunger override their potential skittishness. We also suspect that we may have flavored the pink cottage cheese with methyl anthranilate too strongly. This would have made it harder for the chickens to show signs of confusion; if it were very strongly ingrained in their memory that the pink cottage cheese tasted terrible, they would be less likely to make any mistakes in the future than they might have been if we had flavored it slightly more moderately, perhaps skewing our data by making it harder for them to demonstrate any confusion.

Our data suggests that the confinement experienced by chickens in most factory farms has the potential to be distressing. The effects of stress on the chickens influenced them to regard the cottage cheese hesitantly, when they had previously considered it harmless. These results led us to recognize the importance of improving conditions in factory farms for chickens. However, this data is only applicable to female egg-laying strains of chickens. Since we were only using egg-laying chickens in our experiment, the results could not be used to represent the response of any broiler type chicken or male chickens.

If we were able to replicate this experiment with more time and resources, we would have introduced controls for noise and light. Factory farms are extremely loud and have bright lights on all year round, which are not natural living conditions for chickens. Chickens lay eggs based on the season, which is signaled to their body by light changes (i.e., lengthening days in spring and shortening days in fall). Therefore, the presence of constant light can put a stress on a chicken's body by constantly signaling to it to lay; we are curious as to how this variable might affect chicken performance on our experimental task. We also noticed that loud noises seriously distressed the chickens, which made us wonder how constant noise might affect them. We also would have introduced more colors to our test to see if the chickens had potential color preferences that could have altered the results.

MATERIALS AND METHODS

In this experiment, we tested chickens' ability to learn by association before and after confinement. Their level of intelligence was determined by using a simple, yet effective, associative learning test: color recognition. For this test, chickens were presented with two different plates of cottage cheese. We decided to use cottage cheese, because it is easy for us to mix with other substances and because it is safe and easy for chickens to eat. One of the plates had unflavored cottage cheese that had been dyed a bright green, while the other plate had a portion of cottage cheese that had been dyed bright pink and flavored with methyl anthranilate (half a teaspoon of methyl anthranilate per cup of cottage cheese). The chemical we mixed in, methyl anthranilate, is a harmless chemical that is often used as bird-repellent due to its taste. We decided to use the colors green and pink because both are vivid colors that have been shown to be of equal preference to chicks (5) In this way, choosing the "good" green cottage cheese granted the chickens a snack, while choosing the "bad" pink cottage cheese gave the chickens a foul taste in their mouth.

We split the chickens into two randomly selected groups of five and one group of four. To accurately simulate the confinement used in factory farms, we wanted the chickens to be confined with multiple other chickens. We only had access to two cages, so we could not do them all together. We decided on three groups because each group was an appropriate size for confinement and because we could monitor each chicken to measure their understanding of the color association. For approximately 7 days before we conducted the control test, we fed the chickens the green and pink cottage cheeses to condition them with the mindset that the green would taste good and pink would taste bad, effectively training them to always choose to eat the green and avoid the pink. We gave them seven days, because we observed that after this time period the chickens consistently peck the green cottage cheese and avoid the pink.

After training, we then tested each chicken individually, wherein each chicken was given a cup each of the green and pink cottage cheese. We then recorded their behavior for two minutes, which started immediately after the chicken was placed on the table. We recorded the time at which the first peck occurred, as well as how many times the chicken pecked each color. Following the initial test, we placed each group in confinement, giving each chicken the spatial limitations of approximately 8.5 inches by 11 inches. These limitations were based off the amount of space that chickens are generally given in factory farms, which is around 66 square inches. We did not choose to use the exact size limitations because we did not want to the chickens so crowded that they would overheat. The increment (15 hours) was chosen due to our schedule. We confined the chickens at night, when all of the chickens would usually be checked on, and returned the next day shortly after 2 pm, which resulted in a 15-hour period of confinement. After 15 hours in these conditions, we repeated the testing again on each chicken individually.

After collecting the data, we used a Wilcoxon signed-rank test to determine if confinement had a significant effect on the amount of time it took the chickens to first peck the cottage cheese and the number of times they pecked each

type of cottage cheese. We used this test because we wanted to compare the before and after confinement data and were unable to determine that our data was normally distributed.

Received: May 31, 2019
Accepted: December 4, 2019
Published: December --, 2019

REFERENCES

1. Rose, Nick. "Chickens Are Way Smarter Than We Give Them Credit For." *Vice*, 5 Jan. 2017, www.vice.com/en_us/article/8qeg4b/chickens-are-way-smarter-than-you-think.
2. Smith, Carolynn L, and Jane Johnson. "The Chicken Challenge – What Contemporary Studies of Fowl Mean For Science And Ethics." *Between the Species*, Macquarie University, 2012, digitalcommons.calpoly.edu/bts/vol15/iss1/6/.
3. Marino, L. "Thinking chickens: a review of cognition, emotion, and behavior in the domestic chicken." *Animal Cognition*, vol. 20, no. 2, 2017, pp. 127-147, *SpringerLink*, doi:10.1007/s1007.
4. "10 Interesting Facts about Chicken Vision - VAL-CO." *VAL-CO*, 22 Dec. 2017, www.val-co.com/10-interesting-facts-chicken-vision.
5. "Animals on Factory Farms." *ASPCA*, American Society for the Prevention of Cruelty to Animals, www.aspca.org/animal-cruelty/farm-animal-welfare/animals-factory-farms.
6. Ham, AD and D Osorio. "Colour preferences and colour vision in poultry chicks." *Proceedings of the Royal Society* B, Vol 274, No. 1621, 2007, pp. 1941-1948, doi:10.1098/rspb.2007.0538.
7. Crowe, Simon. "A Comparison of Protocols for Passive and Discriminative Avoidance Learning Tasks In the Domestic Chick." *Brain Research Bulletin*, 2008, www.academia.edu/11640266/A_Comparison_of_Protocols_for_Passive_and_Discriminative_Avoidance_Learning_Tasks_In_the_Domestic_Chick.

Sponsorship

Editor's Circle $10,000+

ALLOY THERAPEUTICS

Patron $5,000+

NEW ENGLAND BioLabs® Inc.
enabling technologies in the life sciences

PORTFOLIOS WITH PURPOSE®

Institutional Supporters

HARVARD UNIVERSITY

HARVARD MEDICAL SCHOOL

Tufts UNIVERSITY

Charitable Contributions

We need your help to provide mentorship to young scientists everywhere.

JEI is supported by an entirely volunteer staff, and over 90% of our funds go towards providing educational experiences for students. Our costs include manuscript management fees, web hosting, creation of STEM education resources for teachers, and local outreach programs at our affiliate universities. We provide these services to students and teachers entirely free of any cost, and rely on generous benefactors to support our programs.

A donation of $30 will sponsor one student's scientific mentorship, peer review and publication, a six month scientific experience that in one student's words, 're-energized my curiosity towards science', and 'gave me confidence that I could take an idea I had and turn it into something that I could put out into the world'. **If you would like to donate to JEI, please visit https://emerginginvestigators.org/support, or contact us at questions@emerginginvestigators.org.** Thank you for supporting the next generation of scientists!

'Journal of Emerging Investigators, Inc. is a Section 501(c)(3) public charity organization (EIN: 45-2206379). Your donation to JEI is tax-deductible. '

emerginginvestigators.org