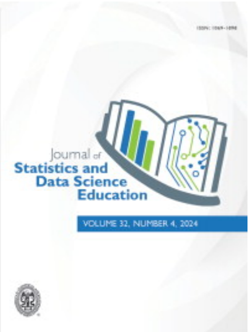


Journal of Statistics and Data Science Education

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ujse21

Age Guessing: A Game to Introduce Fundamental Statistical Concepts

Amir Rastpour & Abraham Amini

To cite this article: Amir Rastpour & Abraham Amini (18 Mar 2024): Age Guessing: A Game to Introduce Fundamental Statistical Concepts, Journal of Statistics and Data Science Education, DOI: 10.1080/26939169.2024.2319152

To link to this article: https://doi.org/10.1080/26939169.2024.2319152

JOURNAL OF STATISTICS AND DATA SCIENCE EDUCATION
2024, VOL. 00, NO. 0, 1–6
https://doi.org/10.1080/26939169.2024.2319152

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Age Guessing: A Game to Introduce Fundamental Statistical Concepts

Amir Rastpour[a] and Abraham Amini[b]

[a]Faculty of Business and Information Technology, Ontario Tech University, Oshawa, ON, Canada; [b]Faculty of Science, Ontario Tech University, Oshawa, ON, Canada

ABSTRACT
We develop a spreadsheet-based game to illustrate fundamental statistical concepts in the first class of an undergraduate Statistics course to motivate students about the topics that they will learn in upcoming classes. This game has been implemented by Google Forms and Google Sheets and can be played in both online and in-person classes of small and large sizes. Statistics is one of the most anxiety-inducing courses for undergraduate students, especially if mathematics is not the focus of their program. Negative anecdotes about the course, mathematics anxiety, and not knowing what the course is exactly about and how practical it can be are among the reasons that contribute to statistics anxiety. The first class provides a good opportunity for an instructor to mitigate these negative impressions and to set a positive attitude toward the course. A pre- and post-game group discussion that we have conducted systematically for six years suggests that the game addresses the students' negative impression about the course and helps them gain a clearer understanding of the tools and skills they will learn in Statistics. Supplementary materials for this article are available online.

ARTICLE HISTORY
Received May 2022
Accepted January 2024

KEYWORDS
Statistics anxiety; Undergraduate statistics

1. Introduction

Motivated by the active learning model (Bonwell and Eison 1991), we propose a spreadsheet-based game to illustrate basic concepts of descriptive and inferential statistics by using student-generated data in a fun and interactive way. This game links abstract statistical concepts to a very natural skill that everyone has some experience with, namely estimating the age of a stranger based on their appearance. The game is designed to be used in the first class of an undergraduate introductory Statistics course, where students do not necessarily have background knowledge about statistics and do not know what to expect from the course. The goal of the game is to motivate concepts that the students will learn in the upcoming classes, and to help them overcome their learning anxiety by providing a clear understanding of what they should expect from the course, what type of statistical tools they will learn and how useful and practical these tools can be in daily life. The data collection and analyses of the data are conducted using Google Forms and Google Sheets. The game can be played in small and large classes and can be implemented in-person and online.

Statistics is fun, but one would not enjoy it unless one understands it and appreciates its power. Unfortunately, many students experience difficulties with learning statistics and do not understand the concepts to a degree that they could enjoy the subject; instead, they often resort to memorizing the concepts (Smith and Martinez-Moyano 2012). Students' learning challenges may not be a result of their intellectual incompetence, but rather may be attributed to their misconception, wrong attitudes, and anxiety (Baloğlu 2004). As documented by Berk and Nanda (1998) and observed by the first author, many students join introductory statistics classes with negative attitudes toward the subject, a presumption of the uselessness of the course to their field and career goals, and anxiety.

Statistics anxiety, which is defined as adverse mental and physiological reaction to statistics content (Primi and Chiesi 2018), plays a key role in students' negative perspective toward the subject. Chew and Dillon (2014) provide a comprehensive survey of empirical and theoretical findings related to implications of statistics anxiety on students' learning experience. Lesser et al. (2013) suggest that incorporating "fun" into statistics classes can decrease students' anxiety and improve their learning motivation. Playing educational games is considered as one of several tools that adds fun into the class environment (Lesser and Pearl 2008) and is widely used by statistics (Lesser et al. 2013) and mathematics (Kärki, McMullen, and Lehtinen 2022) instructors at different levels.

The use of educational games can be even more important in an online setting because online classes tend to be more stressful for students, in comparison to in-person classes. It is suggested that increasing student interaction during online learning improves the students' learning experience through increased motivation to learn and decreased stress (Conrad et al. 2022). Connecting and interacting with peers and instructors during an online class can be a major challenge for students. An interactive online communication decreases "transactional distance," the perceived space for potential misunderstanding

CONTACT Amir Rastpour amir.rastpour@ontariotechu.ca Faculty of Business and Information Technology, Ontario Tech University, Oshawa, ON L1H 7K4, Canada.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/ujse.

between the instructor and learner, and improves the overall online learning experience (Yeung and Yau 2021). Playing games provides a good environment for such interactions (Zulfiqar et al. 2021).

Using games in classrooms fits well within the active learning framework, which requires the participation of students in the learning process, from collecting data to discussing and analyzing the data, and to finding the intended outcomes (Bonwell and Eison 1991). It has been shown that an active learning framework can significantly reduce statistics anxiety and improve students' attitude toward the subject (Carlson and Winquist 2011). To enhance the learning experience, statistics instructors use an active learning framework to develop educational games that not only engages students in learning the concept, but also in question formulation, data collection, and steps to interpret results (e.g., Eadie et al. 2019).

The first class is an ideal opportunity to address students' negative attitudes toward statistics and to incorporate strategies that mitigate their statistics anxiety (Hulsizer and Woolf 2009). Components of our age guessing game are designed to create a collaborative environment during the first class that provides the opportunity for students to get together in groups, and to evaluate their impressions toward statistics. Although none of the authors have conducted a formal survey, the first author has raised questions such as "what is your impression toward the course?" and "what do you think the course is about?" before and after playing the game and has consistently observed that students have a clearer idea of what they should expect from the course and how their new skills could help them in their daily life and future career. It is expected (and confirmed through the in-class questions and discussions) that the game sets a healthy foundation for the upcoming classes by reducing statistics anxiety among students and motivating upcoming topics.

There are other activities, such as introductions, icebreakers, course outline discussions, and expectation setting, that instructors usually conduct during the first class to create a positive classroom climate and to assure that everyone understands the expectations. Our age guessing game complements these activities by providing an opportunity to students to get to know each other through their smaller groups, or through discussions that follow the game. It is also designed to help students set their expectations regarding what they will learn during the semester.

The idea of using an age guessing game in Statistics classes has been used before (e.g., Gelman and Nolan 2017). However, it appears that previous applications were limited to smaller classes, did not motivate a wide range of statistical concepts, and did not provide a link between the game and useful scientific insights. Our proposed game has the following unique features:

1. By linking the game to useful scientific insights provided in a psychology paper (Voelkle et al. 2012), students appreciate the value of statistical tools that they will learn.
2. It can be played in small and large classes.
3. The game is implemented on Google Forms and Google Sheets, which allow it to be played in both physical and online classrooms.
4. A wide range of statistical concepts from average and variance, and visualization tools, to more complex concepts, such as sampling and hypothesis testing are motivated and discussed using the game.
5. The results are analyzed instantaneously.

The rest of the article is organized as follows: Observed impacts of the game is outlined in Section 2; details of the game and how it is played are described in Section 3; and concluding remarks are provided in Section 4.

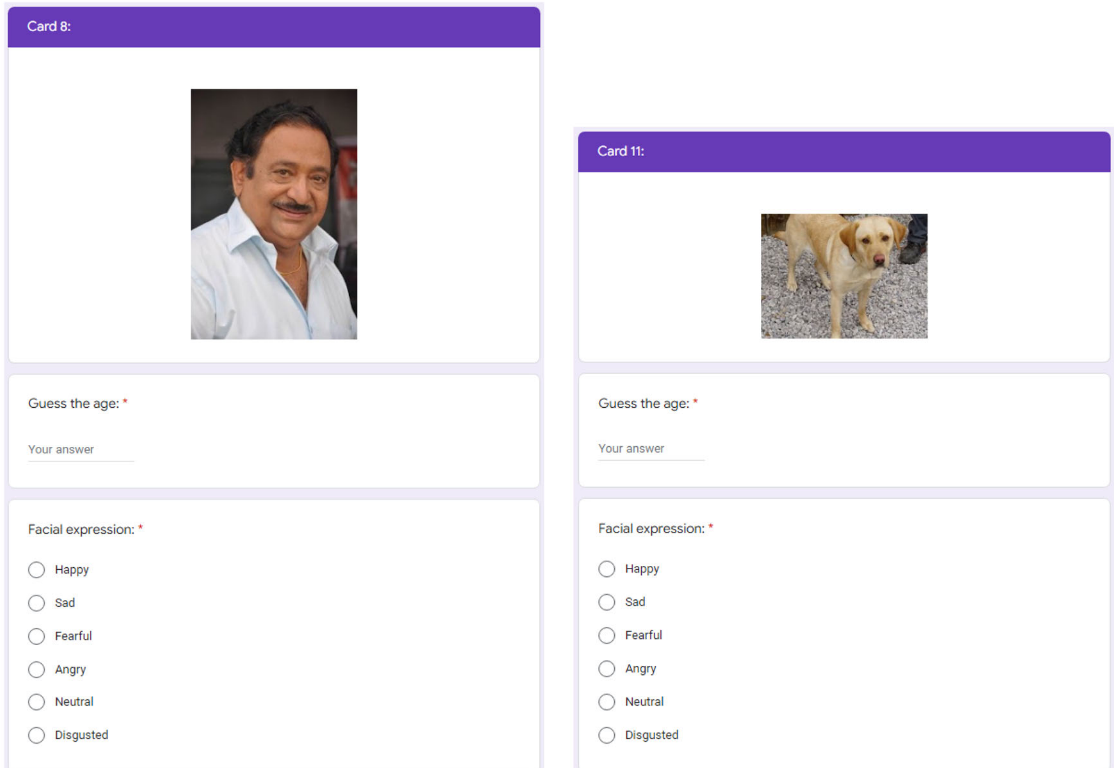
2. The Game

This game is played under the supervision of an instructor. The data collection procedure is automated using Google Forms, which organizes all student inputs into a spreadsheet on Google Sheets. The combination of Google Forms and Google Sheets allows instructors to make changes in both the data collection and the data analyses/interpretation steps as needed.

To initiate the game, the instructor invites students to share their personal experience about estimating the age of strangers, and elicit factors that they think influence the accuracy of these estimates. The instructor actively moderates the discussion and reveals that determinants of accuracy in age estimation and factors that impact the perceived age of a person have been studied by psychologists, and that statistical tools can be used to extract knowledge in such contexts. The instructor introduces a research paper on the topic (Voelkle et al. 2012), without providing any details on findings of the paper (these details are released after playing the game).

2.1. Data Collection

The instructor provides the class with a link to the game's Google Forms page and asks that each group selects a member as a representative: the representative signs into the Google Forms page on behalf of the group, chooses a nickname for the group, and answers questions about the group size and the number of male and female members on the group. Then, Google Forms shows 10 faces whose ages are known only to the instructor (and not to the students) and asks each group to guess the age of each face. In addition, the forms ask the students to guess the mood (happy, sad, fearful, angry, neutral, and disgusted) of the photos based on their facial expression. (Although the focus of the game is on estimating the age, we collect the mood data just to have discussions about the categorical variables as well.) These faces belong to people from different age, gender, and ethnic groups, and are publicly accessible through the UTKFace dataset (Zhang, Song, and Qi 2017) (the students are not aware of the source of the photos). After going through these 10 faces, Google Forms shows five photos from dogs of different breeds and asks the students to guess their age and mood as well—usually, students find these dog photos so cute and it softens up the class environment even more, which helps facilitate better conversations. Figure 1 shows two of the photos used in class. Usually, it takes around 20 min until all groups enter all their answers. While the students work in their groups, the instructor joins each group (or the breakout room in an online setting) and facilitates discussions, if needed.



(a) (b)

Figure 1. Two sample Google Sheets cards used in the Age Guessing game. Pictures are from Zhang, Song, and Qi (2017).

2.2. Data Analyses

Google Forms automatically saves all data into a Google Sheets document. The instructor introduces the concept of error, defined as estimation minus real, and uses a macro to automatically calculate the error of estimation of each group for each card (i.e., each photo). The macro also arranges the data in a format that the instructor can readily use it to draw different plots and to conduct different preliminary analyses. Here is a list of discussions that the first author initiates in class after collecting the data:

- *Data Visualization:* First, the instructor shows the raw data to students and asks them to discuss what they can learn from this data. After a few minutes, usually a consensus is reached that it is difficult (if not impossible) to learn useful insights just by looking at the data. Then, the instructor asks the students to discuss what they know about "data visualization" and how it can help them to extract useful information from the data. After having some discussions, the instructor poses three specific questions, one at a time: (a) "how can you visualize the group size distribution"? (b) "How can you visualize the distribution of the number of male and female members in groups?" (c) "how can you visualize the distribution of estimation errors for a specific photo?", and (d) "how can you visualize whether having more people in a group will necessarily lead to better, or worse, estimations?" The instructor leads the discussions by talking about different data types (i.e., quantitative vs. qualitative) and helps the students to figure out that they can use a pie chart to answer the first question, a bar chart for the second one, a point chart for the third one, and a scatterplot for the last one. The instructor creates these plots and initiates further discussions. Figure 2(a)–(d) demonstrate the plots created in one of classes (plots from other classes are very similar).
- *Finding patterns:* Figure 2(a)–(b) demonstrates the student groups and their configuration based on gender. Figure 2(c)–(d) demonstrates students' estimations and their errors from different perspectives. The instructor encourages students to look at Figure 2(c)–(d) critically and observe patterns. If needed, the instructor guides the conversations by asking questions. For example, while discussing Figure 2(c), the instructor asks whether the groups tended to underestimate, or to overestimate the ages, and whether this pattern was different for older and younger faces. Another interesting question that triggers discussions is whether the "size of zigzags" is the same across different curves in Figure 2(c), or not, and what does a bigger zigzag mean? While looking at Figure 2(d), the instructor asks whether larger groups provided more accurate estimations, and whether this trend changed from one face to another.
- *Sampling issues:* The instructor asks whether the observations the students made in the previous stage were generalizable or not. For example, "Figure 2(c) demonstrates that almost all groups underestimated the age of the 13 years old girl shown on Card 7; can we claim that *everyone* would

underestimate this girl's age?" This question initiates the conversation on the fact that the students in class are not a good representative of the population (i.e., everyone), and then the instructor encourages students to think about different factors (i.e., age, education level, gender, etc.) that should be considered while defining a good sample.

- *Inferential statistics:* After discussing issues around sampling, the students get a better understanding about the concepts of sample versus population, and they are ready to discuss more complex concepts like Inferential Statistics. The instructor leads a discussion in class by asking whether one can use Figure 2(d) and claim that the age estimation of *every* older man will get better if we increase the group size from two to three and to four, and will the estimation get worse if we increase it to five (the pattern that the blue line illustrates). Then, the instructor leads the conversation by asking "what would have possibly happened to the pattern if you [the students] were divided into different groups than your current groups? would you necessarily observe the same blue curve?" The instructor guides the conversations toward the concept of randomness and that in order to generalize an observation, the students should keep in mind that a pattern might be created just out of randomness, and there might not be any underlying reasons for that pattern. Then, the instructor discusses that the topic of inferential statistics will help students to take randomness into account when they want to generalize patterns that they observe in a sample.
- *Linking observations from the collected data to findings of Voelkle et al. (2012):* The game is wrapped up by having a conversation about the fact that statistical tools can be used to scientifically analyze all the items discussed in this game. For example, the instructor reminds that, as shown in Figure 2(c), the estimation error associated with Card 8 (a 57 year old man), has a bigger zigzag (i.e., higher variation), in comparison to the size of the zigzag of Card 7. This is what Voelkle et al. (2012) predicted by stating that "generally, the age of older faces [are] more difficult to estimate than the age of younger faces." The instructor also points out another finding of the article that people are more accurate in estimating the age of other people from their own age group, which could be a reason why the students tend to estimate the age of young faces more accurately (Figure 2(c)–(d) do not include curves associated with young faces to avoid clutter.)

3. The Observed Game Impacts

The first author uses this game in the first lecture of an introductory statistics course, which is offered to second-year Bachelor of Commerce students and includes 100–150 students. First, the instructor assigns students to groups of size 2–5 (small groups were allowed to demonstrate the impacts of having small groups on estimation accuracy), and then poses open-ended questions to students to evaluate their feelings toward the course and to gauge their understanding of what the course is about. Although the purpose of this game has not been to scientifically study the impacts of a game on learning experience of students, the first author has consistently conducted oral surveys over past six years to assess the impact of the game on students. The context of these surveys has been constructed around the four pillars introduced by Hannula (2002) to measure students' attitude toward mathematics. These pillars are: (a) *the emotions the student experiences during mathematics related activities*; (b) *the emotions that the student automatically associates with the concept 'mathematics'*; (c) *evaluations of situations that the student expects to follow as a consequence of doing mathematics*; and (d) *the value of mathematics-related goals in the student's global goal structure*. Before playing the game, the instructor asks the students to discuss the following questions in their groups for 5–10 min:

- Generally speaking, how do you feel about Statistics?
- What are your expectations from this course?
- What have you heard from others about this course?
- Do you think if it is a useful course?"

This conversation is an effective icebreaker and often reveals that most students are either stressed about the course and think it is impossible for them to get a good grade, or they are too naïve and see it as a brief extension for some high school courses that is not necessarily useful for a commerce student who intends to major in "less-quantitative" fields. Some common student responses[1] include: "I took Data Management at high school and I was good at it. I think I will be fine at this course.", "I think this course is very difficult because it involves lots of math," and "My friends told me that I should work really hard to get a good mark."

After playing the game, the instructor paraphrases the same questions to observe changes in students' perception about the course and their understanding of what they will learn. The paraphrased questions are:

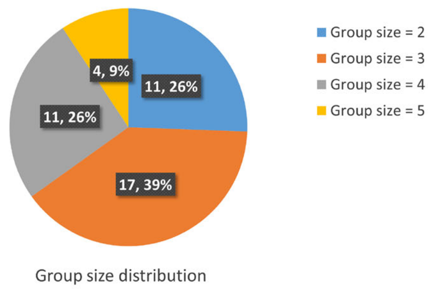
- Are you scared of Statistics, or are you at peace with it?
- How do you compare the practicality of this course to your other courses?
- Do you think this course can help you land better jobs?
- Who can list three topics that we will discuss in this course?

Over the past six years, it has consistently been observed that students feel better about the course and become more realistic about course outcomes after playing the game. Some common student responses include: "It seems the course is not pure mathematics, which make me feel better about it," "It is very cool that we can use statistics to study something like age guessing and it can be applied to many business problems," and "I enjoyed finding patterns in those plots and I think it involves lots of creativity which is fun."

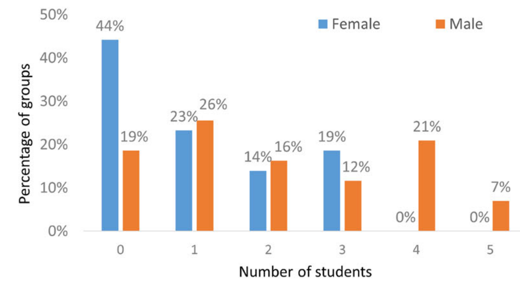
Using the concept of age guessing, which is very easy to grasp to everyone, and generating some real data from students, help students to relate to the game. In turn, the game helps them see what they will learn and how they can apply their statistical skills to extract knowledge from data in their daily life.

During the pre- and post-game group discussions, and while the game is played, the instructor adheres to the Believing and Doubting conceptual framework (Harkness 2009). According to this framework, in order to have rich conversations on topics

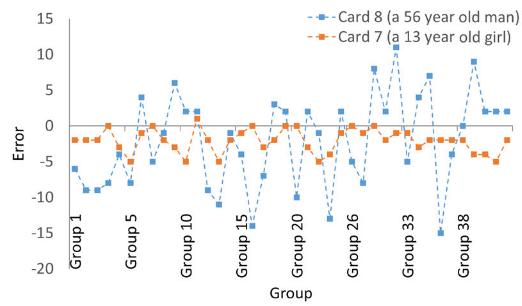
[1] All student quotes in this article are indirect quotes and do not reflect the exact wordings of the students



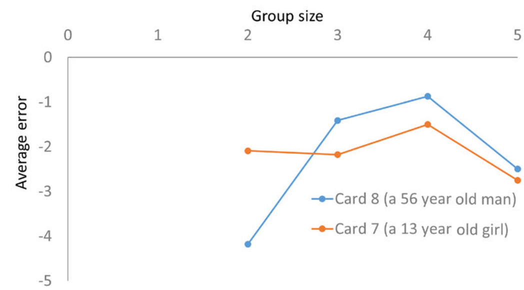
(a) Group size distribution by count and percentage



(b) The distribution of group configuration



(c) The estimation error of groups for two cards



(d) The average error vs. the group size

Figure 2. Using visualization to extract information from data.

related to mathematics, the instructor should "try to believe" students' opinions rather than "doubting" their arguments and thoughts. That is, even when students provide "wrong answers," the instructor should believe that there are sparks of rightness in those answers and try to "unpack" the assumptions and thoughts behind those answers, rather than doubting the students' capabilities and moving on to perceived right answers. According to the first author's experience, this approach helps students feel more comfortable in class and become more active in discussions.

4. Conclusion

We developed a game based on estimating the age of strangers to introduce fundamental statistical concepts, from descriptive to inferential statistics. This game is implemented using Google Forms and Google Sheets, which allow the instructor to tailor it toward the needs of their students. This game is designed for small and large undergraduate classes and can be played in both online and in-person settings. This game includes data collection and data analysis steps. It is designed to trigger fruitful discussions since students often relate to the action of estimating the age of strangers. A systematically conducted oral survey over the past six years confirms that the game helps students to be more realistic about the course and the topics that they will learn, and decreases their negative impressions about the course.

Data Availability Statement

- Google Forms page for playing the game: *https://forms.gle/Y2yVSe7fNEoPZg1t6*
- Sample processed data and Figure 2 data: *https://osf.io/eq2c6/?view_only=7a5677e964184c01a3d95627c44caddf*

Disclosure Statement

The authors have no relevant financial or nonfinancial interests to disclose.

Funding

We acknowledge that this work has been partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

Baloğlu, M. (2004), "Statistics Anxiety and Mathematics Anxiety: Some Interesting Differences I," *Educational Research Quarterly*, 27, 38–48.

Berk, R. A., and Nanda, J. P. (1998), "Effects of Jocular Instructional Methods on Attitudes, Anxiety, and Achievement in Statistics Courses," *Humor - International Journal of Humor Research*, 11, 383–409. DOI:10.1515/humr.1998.11.4.383.

Bonwell, C. C., and Eison, J. A. (1991), *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports* (ERIC).

Carlson, K. A., and Winquist, J. R. (2011), "Evaluating an Active Learning Approach to Teaching Introductory Statistics: A Classroom Workbook Approach," *Journal of Statistics Education*, 19. DOI:10.1080/10691898.2011.11889596.

Chew, P. K., and Dillon, D. B. (2014), "Statistics Anxiety Update: Refining the Construct and Recommendations for a New Research Agenda," *Perspectives on Psychological Science*, 9, 196–208. DOI:10.1177/1745691613518077.

Conrad, C., Deng, Q., Caron, I., Shkurska, O., Skerrett, P., and Sundararajan, B. (2022), "How Student Perceptions About Online Learning Difficulty Influenced Their Satisfaction During Canada's Covid-19

Response. *British Journal of Educational Technology*, 53, 534–557. DOI:10.1111/bjet.13206.

Eadie, G., Huppenkothen, D., Springford, A., and McCormick, T. (2019), "Introducing Bayesian Analysis With M&M'S®: An Active-Learning Exercise for Undergraduates," *Journal of Statistics Education*, 27, 60–67. DOI:10.1080/10691898.2019.1604106.

Gelman, A., and Nolan, D. (2017), *Teaching Statistics: A Bag of Tricks*, Oxford: Oxford University Press.

Hannula, M. S. (2002), "Attitude Towards Mathematics: Emotions, Expectations and Values," *Educational studies in Mathematics* 49, 25–46. *https://www.jstor.org/stable/3483259*.

Harkness, S. S. (2009), "Social Constructivism and the Believing Game: A Mathematics Teacher's Practice and Its Implications," *Educational Studies in Mathematics*, 70, 243–258. DOI:10.1007/s10649-008-9151-3.

Hulsizer, M. R., and Woolf, L. M. (2009), *A Guide to Teaching Statistics: Innovations and Best Practices* (Vol. 10), Chichester: Wiley. DOI:10.1002/9781444305234.

Kärki, T., McMullen, J., and Lehtinen, E. (2022), "Improving Rational Number Knowledge Using the NanoRoboMath Digital Game," *Educational Studies in Mathematics*, 110, 101–123. DOI:10.1007/s10649-021-10120-6.

Lesser, L. M., and Pearl, D. K. (2008), "Functional Fun in Statistics Teaching: Resources, Research and Recommendations," *Journal of Statistics Education*, 16, 1–10. DOI:10.1080/10691898.2008.11889572.

Lesser, L. M., Wall, A. A., Carver, R. H., Pearl, D. K., Martin, N., Kuiper, S., Posner, M. A., Erickson, P., Liao, S. M., Albert, J., et al. (2013), "Using Fun in the Statistics Classroom: An Exploratory Study of College Instructors' Hesitations and Motivations," *Journal of Statistics Education*, 21, 1–33. DOI:10.1080/10691898.2013.11889659.

Primi, C., and Chiesi, F. (2018), "The Role of Mathematics Anxiety and Statistics Anxiety in Learning Statistics," in *10th International Conference on Teaching Statistics, Kyoto, Japan.*

Smith, A. E., and Martinez-Moyano, I. J. (2012), "Techniques in Teaching Statistics: Linking Research Production and Research Use," *Journal of Public Affairs Education*, 18, 107–136. DOI:10.1080/15236803.2012.12001674.

Voelkle, M. C., Ebner, N. C., Lindenberger, U., and Riediger, M. (2012), "Let Me Guess How Old You Are: Effects of Age, Gender, and Facial Expression on Perceptions of Age," *Psychology and Aging*, 27, 265–277. DOI:10.1037/a0025065.

Yeung, M. W., and Yau, A. H. (2021), "A Thematic Analysis of Higher Education Students' Perceptions of Online Learning in Hong Kong Under Covid-19: Challenges, Strategies and Support," *Education and Information Technologies*, 27, 181–208. DOI:10.1007/s10639-021-10656-3.

Zhang, Z., Song, Y., and Qi, H. (2017), "Age Progression/Regression by Conditional Adversarial Autoencoder," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Accessed 27 May 2022. Available at *https://susanqq.github.io/UTKFace/*.

Zulfiqar, S., Al-reshidi, H. A., Al Moteri, M. A., Feroz, H. M. B., Yahya, N., and Al-Rahmi, W. M. (2021), "Understanding and Predicting Students' Entrepreneurial Intention Through Business Simulation Games: A Perspective of Covid-19," *Sustainability*, 13, 1838. DOI:10.3390/su13041838.