



# Using Animations in Assessment

Anastasios Karakolidis  
anastasios.karakolidis@gmail.com



# The reality

The vast majority of tests in use today rely heavily on static text



# The problem with text-based assessments

Limited in terms of the complexity of what can be presented as stimuli and/or responses [1]

Test-takers are required to have skills (e.g., reading comprehension) that may be irrelevant to the construct that the assessment purports to capture [1, 2]

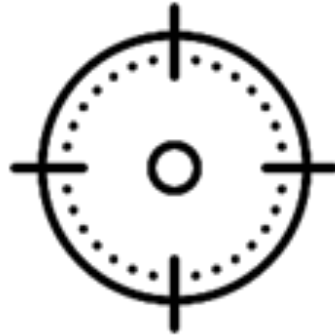
Such equity and fairness issues might affect:

- test-takers with poor language and reading skills
- non-native speakers
- people with dyslexia

***Construct-irrelevant variance → validity [3]***

# This innovation...

explores the extent to which animated videos are a useful alternative to written text in testing



# Why animated videos?

## Advantages:

- replace much of written text
- enable the visualisation of complex information that cannot be easily communicated via text
- more effective than other video-based media

## Challenges:

- development complexities
- cost

# Development of an animated test

Selection of a situational judgement test (SJT) designed by Steven Stemler, Robert Sternberg and their colleagues at Yale to measure teachers' practical knowledge of how to deal with challenging social situations at work [4, 5]

Collaboration with an animation company to create 15 animated videos based on the original text-based scenarios

# A sample text-based scenario

6. Ms. Green sometimes groups her students when teaching them. Using different colours, she divides the students in her class into three groups: the strongest students comprise the yellow group, the average-achieving students make up the grey group, and finally, the weakest students form the blue group. Groups of different colours receive different tasks and so the method helps to individualize teaching. Ms. Green also forms different groups for different subjects. She makes sure that her students understand why she groups them. George was initially in the blue group for reading because his reading was below average. However, he was often one of the best in the blue group. A week ago George asked Ms. Green if he could work with the grey group. Ms. Green thought about it and decided that it was worth trying. On the very next day she noticed that George was struggling in the grey group. Reading at that level was really difficult for him. Moreover, there was a difference in how George felt about being in the grey group versus the blue group. In the blue group he was one of the best, while in the grey group he was the weakest student. George is very sensitive to his classmates' and teacher's opinions, and Ms. Green is worried about him. Yesterday on George's 4th day in the grey group, Ms. Green noticed that he was very upset following the reading lesson.

**What should Ms. Green do? Rate the extent to which you agree or disagree with each of the following statements.**

Ms. Green should:

1. make it clear that once a student moves up to a new level, he or she has to stay in that group for a week at least.

Strongly Disagree

Disagree

Neither Agree Nor Disagree

Agree

Strongly Agree

2. talk to George on his own and encourage him to stay in the grey group, saying that eventually things will get better.

3. offer George after class help so that he knows that his teacher is trying to help him stay in the grey group.

4. phone George's parents and talk to them about the situation.

5. continue to observe the situation before making her decision.

6. tell George that he needs to return to the blue group because he is not performing well enough in the grey group.

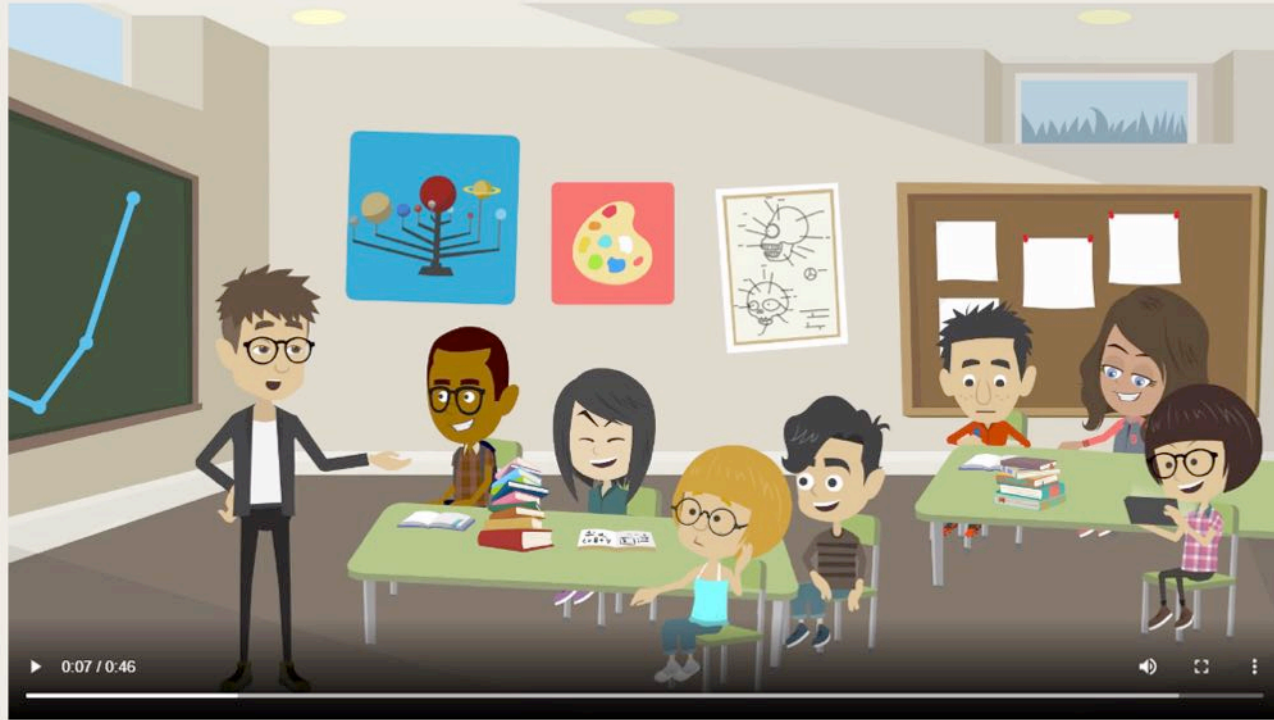
7. get a student from the grey group to help George during class.



## The animated SJT

# The animated test

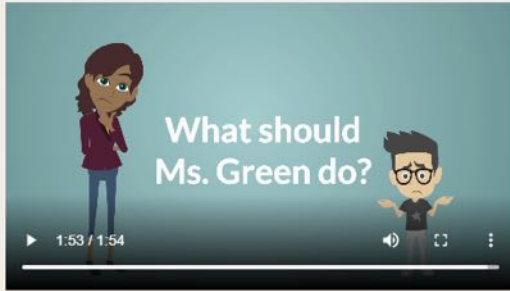
1. Watch the video and rate the extent to which you agree or disagree with each of the following statements.



1 / 15

# The animated test

6. Watch the video and rate the extent to which you agree or disagree with each of the following statements.



Ms. Green should:

1. make it clear that once a student moves up to a new level, he or she has to stay in that group for a week at least. (Agree)

2. talk to George on his own and encourage him to stay in the grey group, saying that eventually things will get better.

Strongly Disagree

Disagree

Neither Agree Nor Disagree

Agree

Strongly Agree

3. offer George after class help so that he knows that his teacher is trying to help him stay in the grey group.

4. phone George's parents and talk to them about the situation.

5. continue to observe the situation before making her decision.

6. tell George that he needs to return to the blue group because he is not performing well enough in the grey group.

7. get a student from the grey group to help George during class.

Next Scenario

# Process of animating

1st draft



2nd draft



Final draft



# The research study

Experimental comparison of the animated vs. the text-based version of the test

Sample: native (n=51) and non-native (n=78) English-speaking pre-service teachers - random assignment

Measures:

- SJT performance
- English proficiency
- reading comprehension ability in English

# Research question #1: What impact does the use of animated videos have on construct-irrelevant variance attributed to language and reading skills?

- The variance attributed to construct-irrelevant factors was reduced by 9.5%
- In other words, the use of animated videos reduced the dependency of test performance on irrelevant factors, such as participants' proficiency and reading comprehension in English



# Research question #2: What impact does the use of animated videos have on test-takers' reactions to the test?

The animated version of the test



was perceived as more valid and fair ( $d = 0.38$ )





was perceived as more enjoyable ( $d = 0.50$ )



was found to use language that was less difficult to understand ( $\eta^2 = 0.043$ )

# So what?

- Did animations make a substantial difference? 
- We should care about what test-takers think of the tests they take!
- Test-takers perceptions of a test can have an impact on:
  - their attitudes towards the body organising the assessment <sup>[1]</sup>
  - their engagement with the assessment <sup>[6, 7]</sup>
  - their performance <sup>[6, 7]</sup>
- Is it worth the effort and money? 



# The potential

- Improving equity and fairness
- Improving standardisation in assessment
- Testing across different populations
- Testing hard-to-measure constructs
- Professional development



# References

- [1] E. C. Popp, K. Tuzinski, and M. Fetzter, “Actor or avatar? Considerations in selecting appropriate formats for assessment content,” in *Technology and Testing: Improving Educational and Psychological Measurement*, F. Drasgow, Ed. New York, NY: Routledge, 2016, pp. 79–103.
- [2] J. Abedi, “Linguistic factors in the assessment of English language learners,” in *The SAGE Handbook of Measurement*, G. Walford, E. Tucker, and M. Viswanathan, Eds. California, Ca: SAGE, 2010, pp. 129–105.
- [3] American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME), *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association, 2014.
- [4] S. E. Stemler, J. G. Elliott, E. L. Grigorenko, and R. J. Sternberg, “There’s more to teaching than instruction: Seven strategies for dealing with the practical side of teaching,” *Educ. Stud.*, vol. 32, no. 1, pp. 101–118, 2006.
- [5] J. G. Elliott, S. E. Stemler, R. J. Sternberg, E. L. Grigorenko, and N. Hoffman, “The socially skilled teacher and the development of tacit knowledge,” *Br. Educ. Res. J.*, vol. 37, no. 1, pp. 83–103, Feb. 2011.
- [6] Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology*, 82(2), 300–310.
- [7] Scott, J. C., & Mead, A. D. (2011). *Foundations for measurement*. In N. Tippins & S. Adler (Eds.), *Technology-Enhanced Assessment of Talent* (pp. 21–65). San Francisco, CA: Jossey-Bass.

Email address: [anastasios.karakolidis@gmail.com](mailto:anastasios.karakolidis@gmail.com)