

# Game-based Assessments: Design and Validation

Game-based Assessment: An Interdisciplinary Workshop

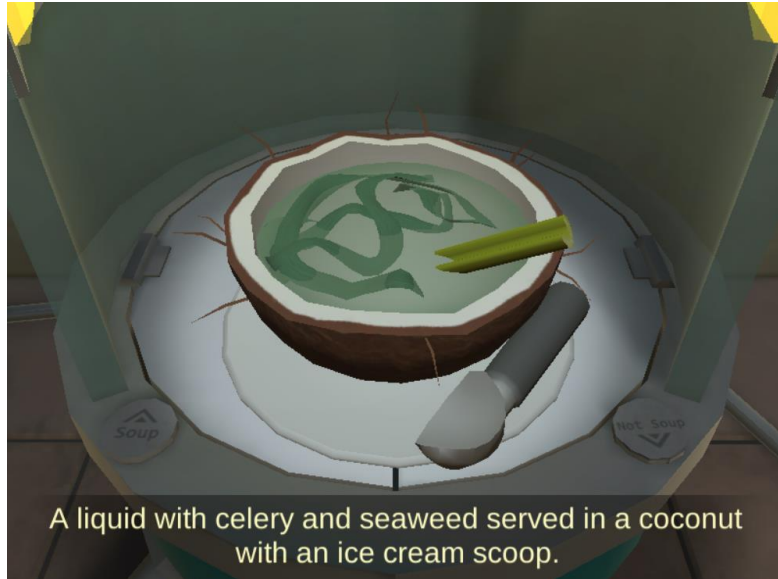
August 22, 2019

# First, some context.

- I am a methodologist in the organizational sciences.
- My goals are identifying high quality measurement approaches to assess job applicants, trainees, and other organizational members.
- You will see this bias emerge quite clearly.



# First, a brief story about GBA and me



A liquid with celery and seaweed served in a coconut with an ice cream scoop.

from: <http://soup.gua-le-ni.com/>



# Inherently Interdisciplinary

- Play vs. Games
  - Play is the unstructured, instinctive way children learn about the world
  - Play with a structured set of rules is a game
  - Children cross the line between play and games freely
  
  - But when is that line crossed?
    - Huizinga's magic circle
    - Easier to compare extremes

Landers, Tondello, Kappen, Collmus, Mekler & Nacke (2019)



# Creating a Fun Game is Already Hard

- Creating a game involves a lot of time and a lot of money
    - Grand Theft Auto V (2013): US\$265M (but earned US\$800M in 24 hours, and at least US\$1.5B in total revenue to date)
    - Most modern AAA titles cost US\$20M-US\$30M; indie can be much less (as little as US\$10K, with typical indies US\$100K-US\$300K)
  - Why is it so complicated and expensive?
  - Because games are extremely complicated
    - Interrelated systems design, intended to create a targeted experience
    - Most common design framework: MDA (Mechanics, Dynamics, Aesthetics)
- Landers, Auer, Collmus & Armstrong (2018)

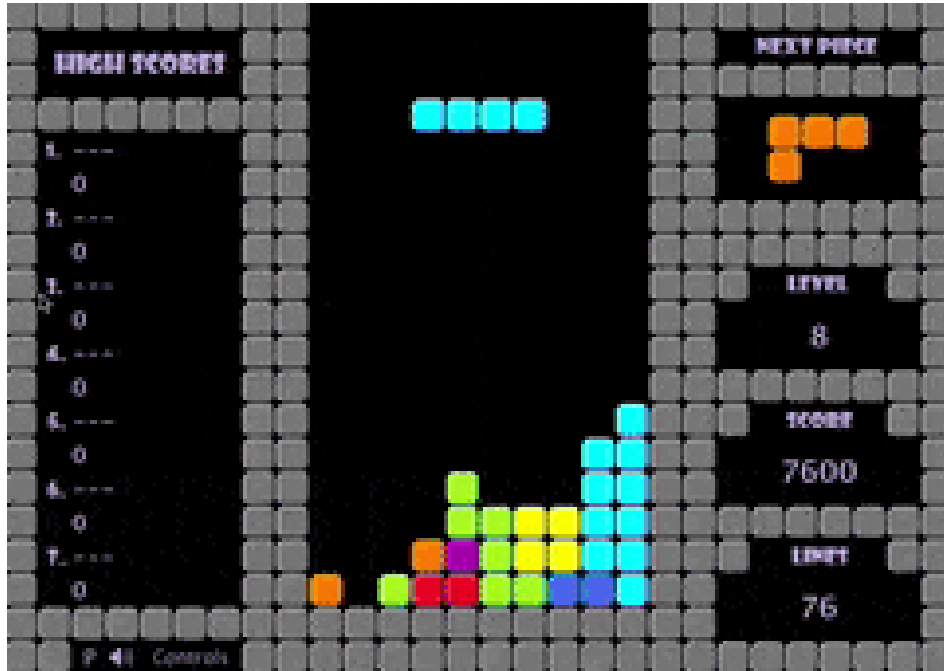


# Example Mechanics

- Scoring  
(such as PBL: points, badges, and leaderboards)
- Turn-taking
- Interfaces (such as dice, game controllers)
- Avatars
- Risk-taking
- Victory conditions (and victory, generally)



# What Are the Mechanics Here?



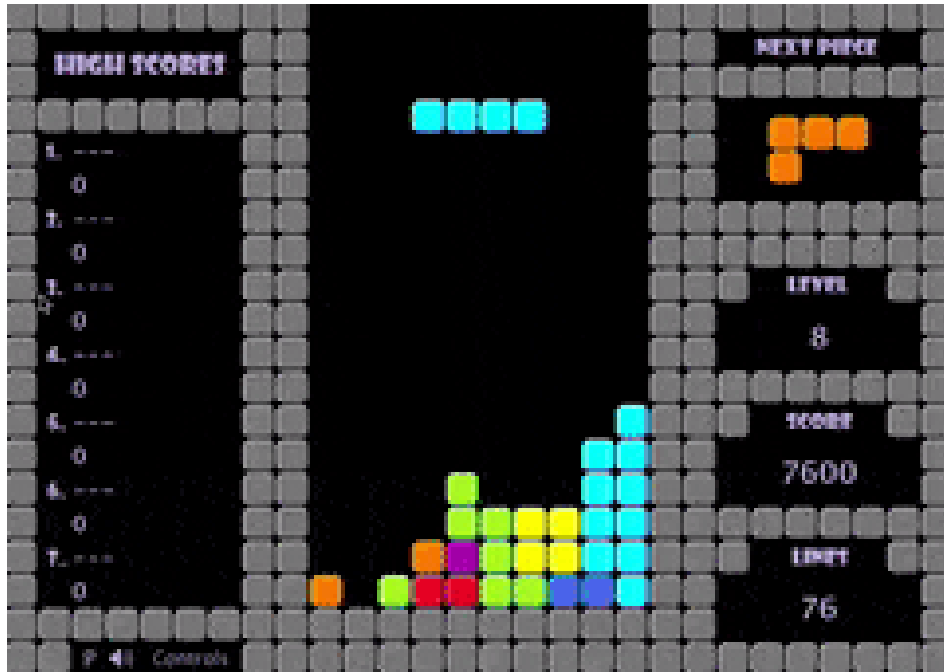
# Basic Mechanics (Game Systems)

- Rotation System
- Color System
- Internal Scoring System
- Piece Selection System
- Piece Preview System
- High Score System
- Piece Movement System
- Line Counting System
- Game Ending System
- Levels System
- Menu System
- Music System
- Sound Effect System
- Control System





# What Dynamics Emerge?



# Example Dynamics

- Emergent interactions created by combining games mechanics with player behaviors over time.
- Piece Movement System + Piece Preview System = *Possible Distraction During Gameplay*
- Piece Movement System + Levels System = *Increasing Time Pressure and Difficulty*
- Piece Movement System + Scoring = *Increased Effort to Score a 4-Line Tetris*

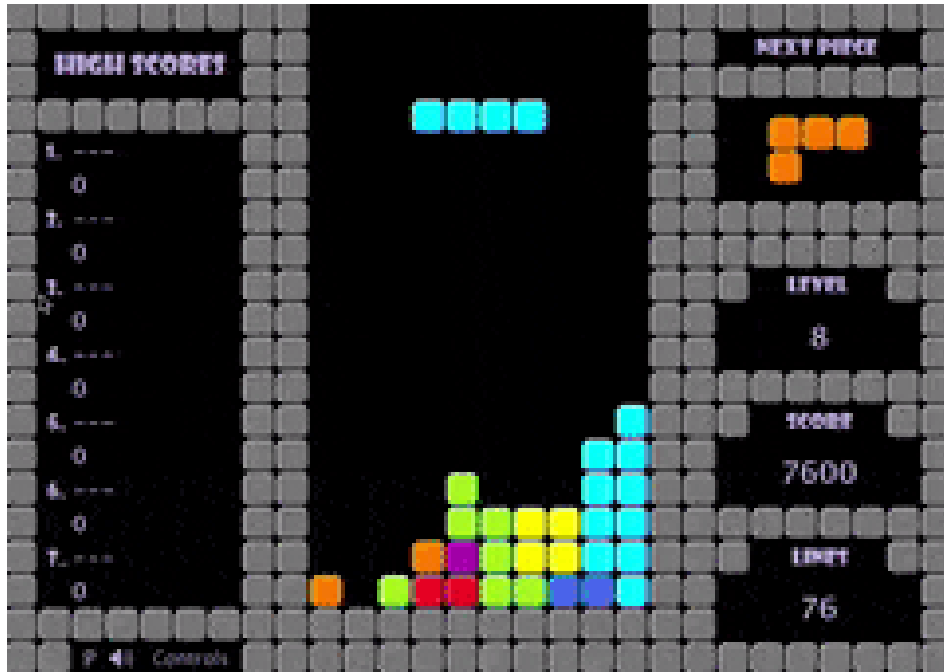


# Types of Aesthetics (from MDA)

- Sensation: provides new experiences
- Narrative: a story that hooks
- Fantasy: a world to immerse oneself
- Fellowship: enabler of social relationships
- Discovery: curiosity about a game environment/world
- Challenge: urge to overcome and master
- Expression: enabling self-discovery
- Submission: immersion into game as a whole



# What Aesthetics Are Created?



# MDA to Deconstruct Any Game



# Assessment Goals Add Complexity

- Psychometric characteristics and gameplay quality are not *necessarily* opposed, but they often are in practice.
  - Reliability
  - Validity
- Aesthetics vs. Assessment Goals
  - Sensation (new experiences) vs. measurement occasions
  - Fantasy vs. serious high-stakes context
  - Fellowship (social relationships) vs. individual assessment
  - Expression (self-discovery) vs. testing time



# Let's Briefly Turn to Gamification

- Businesses saw and liked the money and success of video games but did not like the cost (aside from a few scattered *serious games*)
  - Also led to proliferation of "game" as a sales tactic
- We've defined *gamification* as a **design strategy** in which **game elements** are added to **non-game contexts** (Callan, Bauer & Landers, 2015, building on Deterding)
  - Borrows elements from games and applies them elsewhere (usually PBL)
- Gamification is commonly done rhetorically or just badly (Landers, 2019)



# Gamification Could Create a Game

- But it doesn't necessarily create a game.
- Remember that games are "structured play with imposed rules that a player has agreed to follow."
- Gamification can involve the addition of any game element (e.g., new mechanics, targeted dynamics or aesthetics).
- **Therefore:** Gamification of an existing assessment does not necessarily make it into a GBA.





# Example: Gamifying Personality Assessment (but no game)

- How do we use game elements to take an existing personality assessment and improve its aesthetics?
- We only have control over game mechanics; so which game mechanics are most likely to lead to improvements in targeted aesthetics?



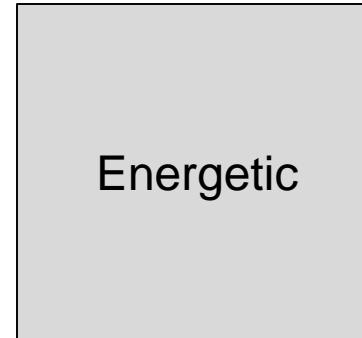
# Inspired by a Gamified Application

- Tinder
  - Makes provision of ratings fun, enjoyable, and motivating



# A Gamification Project

- Project with Nathan Weidner (also here today!)
- Converted a personality inventory into a swipe-based measure based upon Saucier's mini-markers
- Examined reactions to it on MTurk (N=287) versus a traditional Likert-type measure
- Currently under review (R&R!)

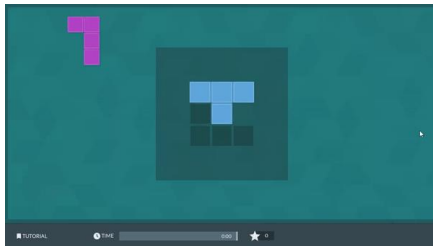
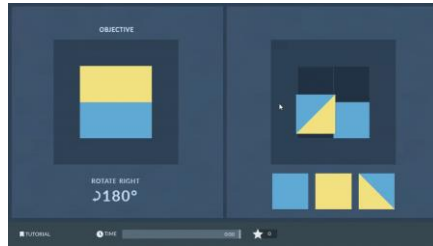


# Gamification != Games

- *Assessment gamification* is a design process that adds game elements to an existing assessment, which may or may not create a game
  - As a *design process*, is like "scale development"
- *Game-based assessments* are assessment methods in the form of a game (i.e., structured play with rules)
  - As a *method*, is like "Likert-type scales"
  - Is more likely created using *game design* than *gamification*



# Validating a GBA: Cognify

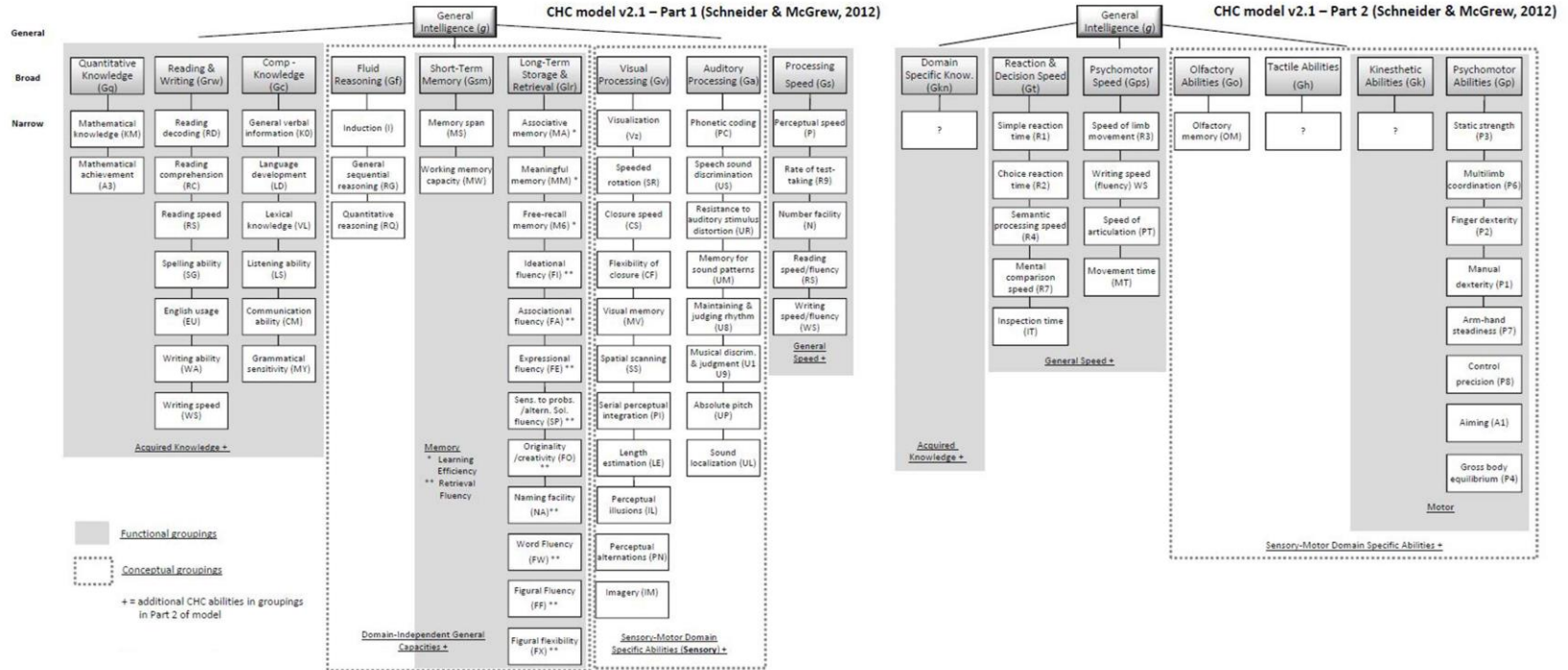


# About Revelian

- Fairly unusual in the current assessment games space because of their complete grounding in I-O psychology (a psychological theory-driven approach)
- Cognify was developed by looking at the CHC model of general cognitive ability and trying to (roughly) target specific abilities



# CHC Theory of Intelligence



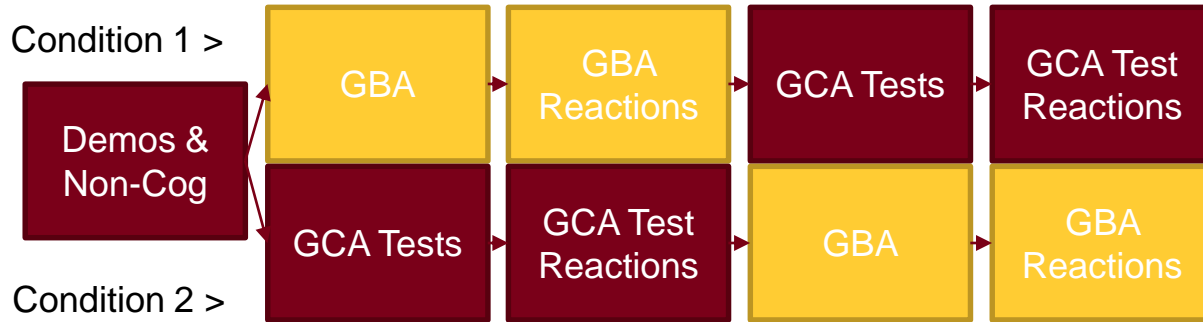
# Study Design

- Two simultaneous recruitment efforts
  - Undergraduates in psychology for extra credit
  - Undergraduates university-wide for \$20
  - \$100 incentive for top 20 participants
- Two-hour study in a semi-controlled environment
  - N=530





# Study Design



- Verbal Ability:
- Processing Speed:
- Fluid Intelligence:
- Quantitative Reasoning:
- Visual Processing:

GRE Verbal Reasoning  
Chicago Non-Verbal Exam  
ETS Kit Nonsense Syllogisms  
GRE Quantitative  
ETS Kit Paper Folding Test



# Game-thinking Cannot Remove AI

- Consider the claim: "This cognitive ability test game-based assessment does not show/shows reduced adverse impact in comparison to traditional cognitive ability tests."
- This is only possible if...
  - A GCA GBA measures different constructs than GCA
  - A GCA GBA measures GCA poorly
- The cause here is (usually) the construct, not the method.
  - Some genres of game are still likely to create AI by gender.



# Theory-based GBA Looks Promising

- Undergrads, at least, liked *this* game-based assessment
  - More intrinsically motivated, believe it's fairer, believe it's more appropriate for job applications
- At least this assessment was designed reasonably effectively
  - Must avoid the Arthur & Villado (2008) trap
  - Likely can be designed and refined to meet psychometric (CTT) assumptions
  - Appears to behave *similarly* to a g measure, has incremental prediction although source is unclear
  - Differential prediction appears similar – if you don't have similar differential prediction in a GCA assessment game, you're not measuring GCA
- Organizational validation: supervisory ratings of job performance at a large multinational consumer goods manufacturer ( $r = .29$  overall,  $.40$  numerical reasoning)



# Lessons Learned and Cautions

- Need to be careful not to consider "this GBA" and "GBAs" as synonymous
  - Design processes are critical, and of the various fields involved in GBA, only game design really studies them
  - Conclusions from one GBA probably do not generalize to GBAs in general
- Need to pursue a rigorous psychometric standard
  - This problem is amplified with many AI-based approaches
- Was likely easier with cognitive ability versus non-cog



# Thank You!

Richard N. Landers, Ph.D.  
[rlanders@umn.edu](mailto:rlanders@umn.edu)

Game-based Assessment: An Interdisciplinary Workshop  
August 22, 2019



UNIVERSITY OF MINNESOTA

**Driven to Discover®**