Human perception, action, and motor control 101

Game Analysis 2024

Aalto University

Perttu Hämäläinen

Disclaimer: This content is perpetually work-in-progress, updated every year.

Fourth Edition

Motor Learning and Performance

A Situation-Based Learning Approach

Includes access to
- online study guide
- 20 experiential labs

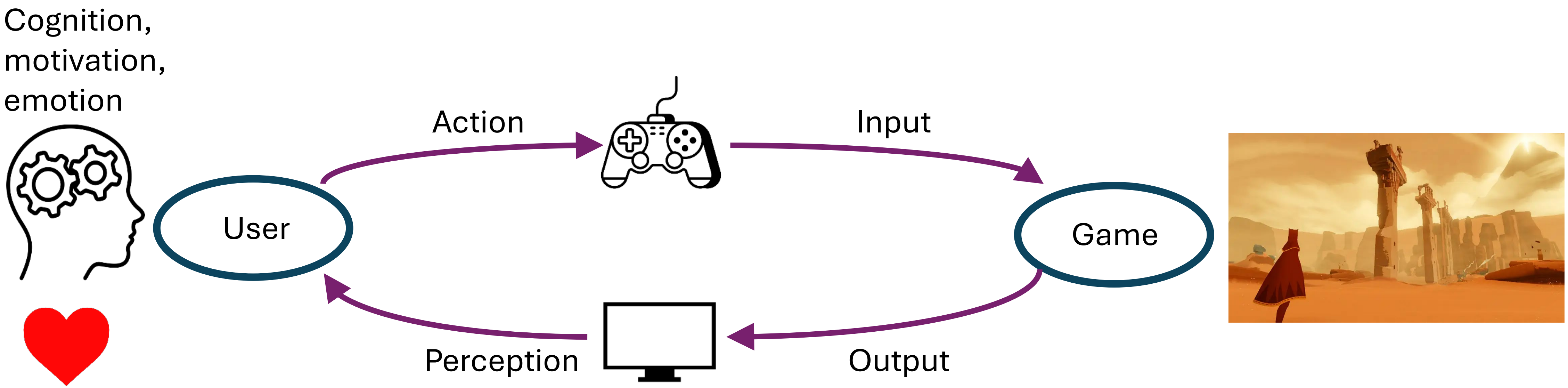
Richard A. Schmidt
Craig A. Wrisberg

Topics

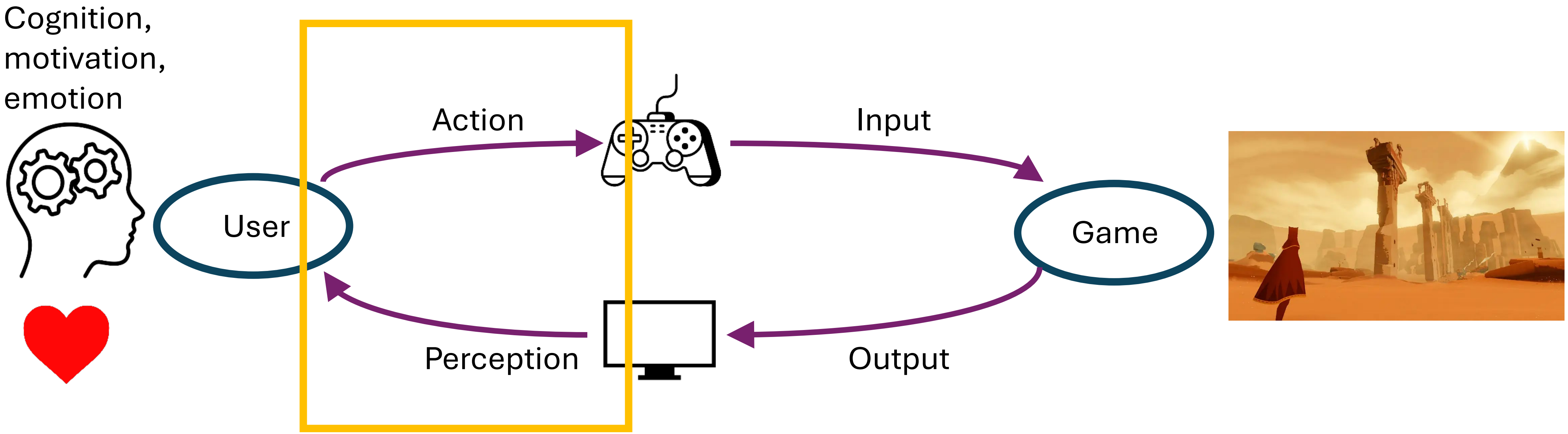
- Intro
- Reaction time
- Aiming, hitting, catching
- Importance of automation
- Limits of attention

Intro

The interaction loop



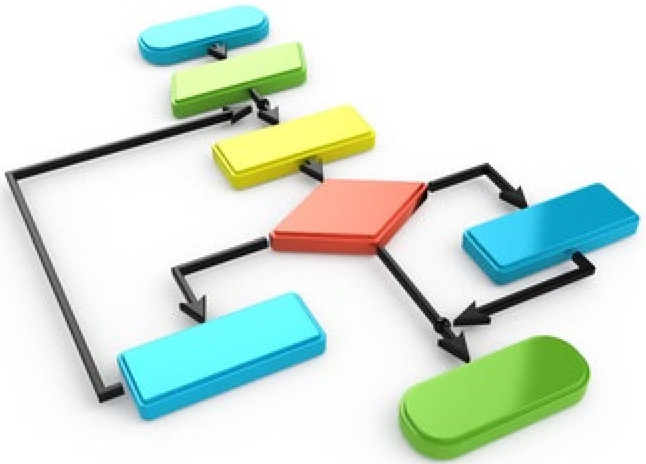
The interaction loop



Action games & exergames: perceptual and motor challenges are central part of the game

Two types of thinking

SYSTEM 2: EXPLICIT/CONSCIOUS:
SLOW, ALGORITHMIC, EFFORTFUL



ATTENTION,
PRIMING,
INHIBITION

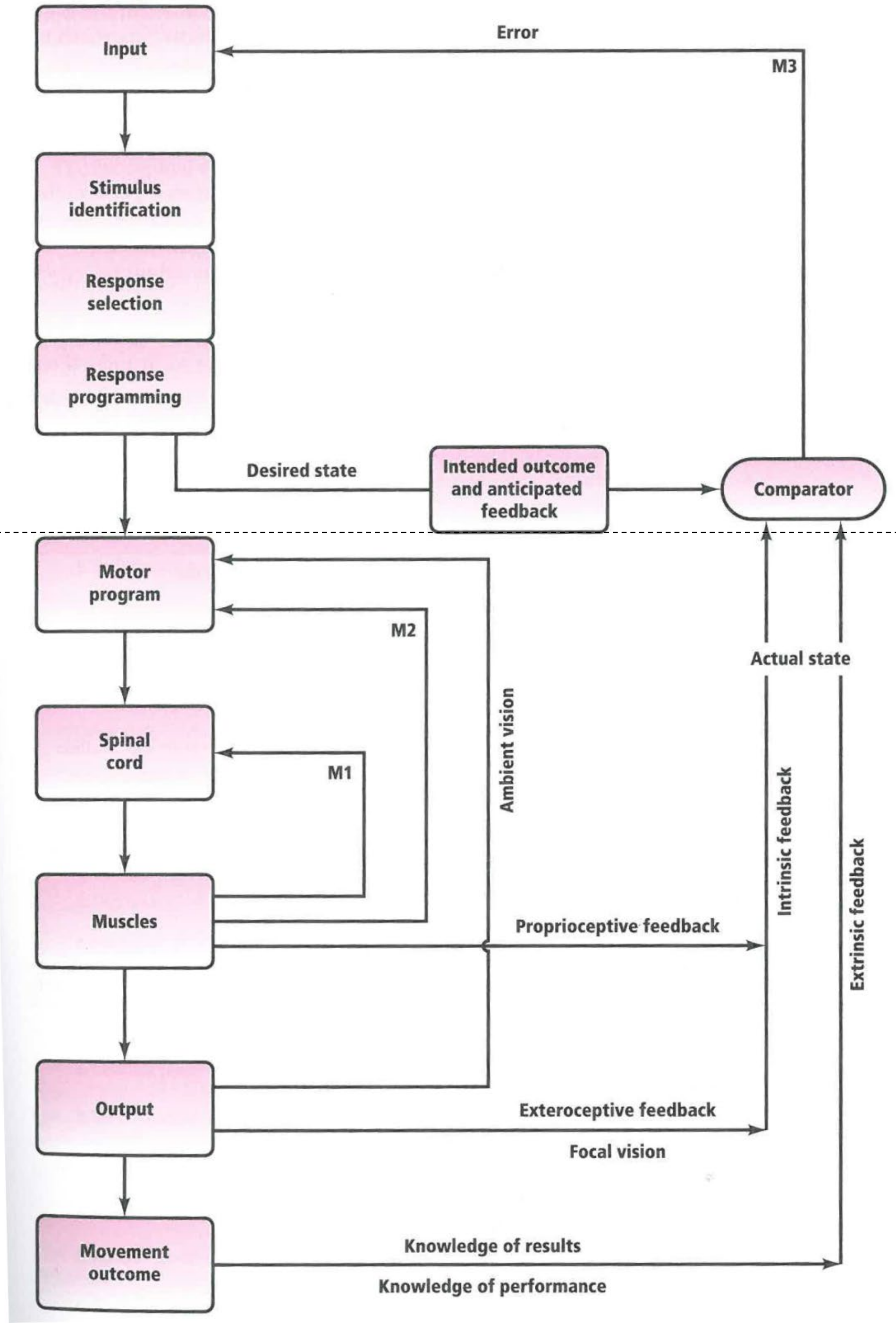
IDEAS,
ASSOCIATIONS,
IMPULSES,
INTUITIONS

SYSTEM 1: IMPLICIT/UNCONSCIOUS:
FAST, ASSOCIATIVE, EFFORTLESS,
INTUITIVE

ASSOCIATION

Slow,
(mostly) conscious

Fast,
non-conscious



Reaction time

Reaction time

- Our reactions can be controlled or automatic
- Controlled reactions are slower
- In general, our reaction time is around 150-200ms
- Even faster reactions possible through automatic processing, acting without thinking
- Stimulus intensity, size and location matter

Slater-Hammel's Experiment

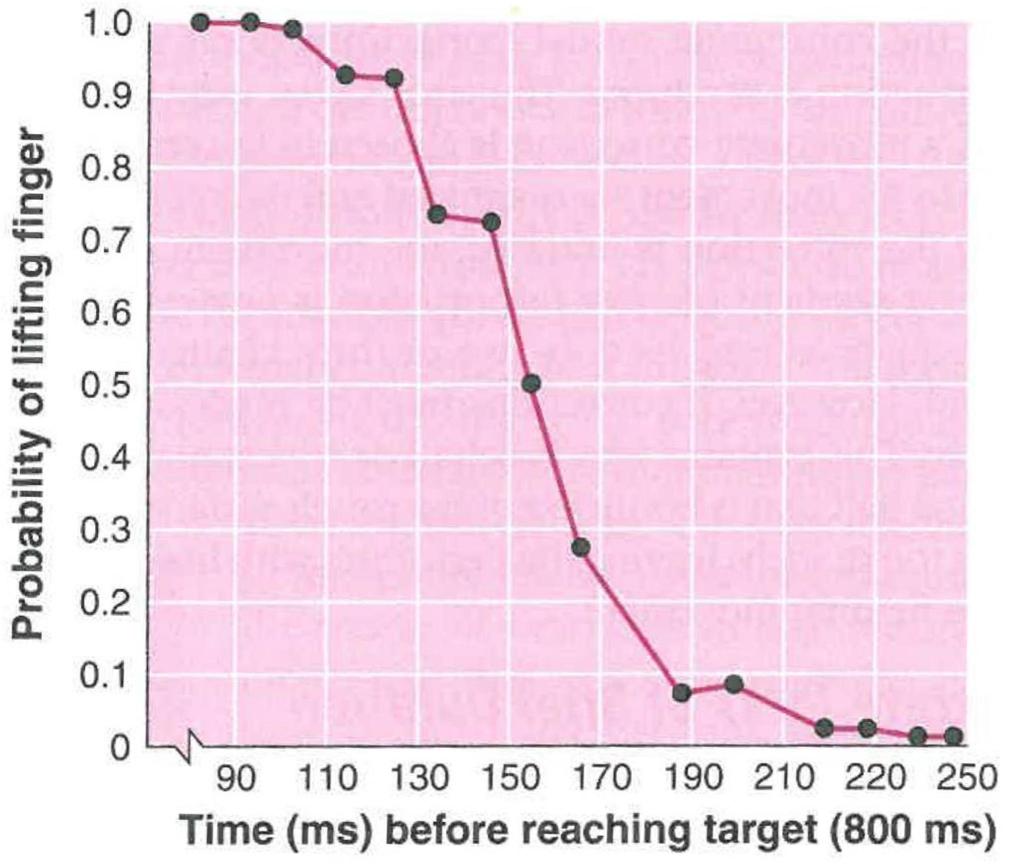


Figure 3.7 The probability of lifting the finger though the clock hand had stopped, plotted as a function of the interval of time before the 800 ms position. (Adapted from Slater-Hammel, 1960.)

Movements

are initiated 150-170ms ahead and can't be inhibited once initiated.

Things affecting reaction time

- **Hick's Law (number of alternative reactions)**
- Intuitiveness (stimulus-response compatibility)
- Anticipation (animations matter)

Hick's law

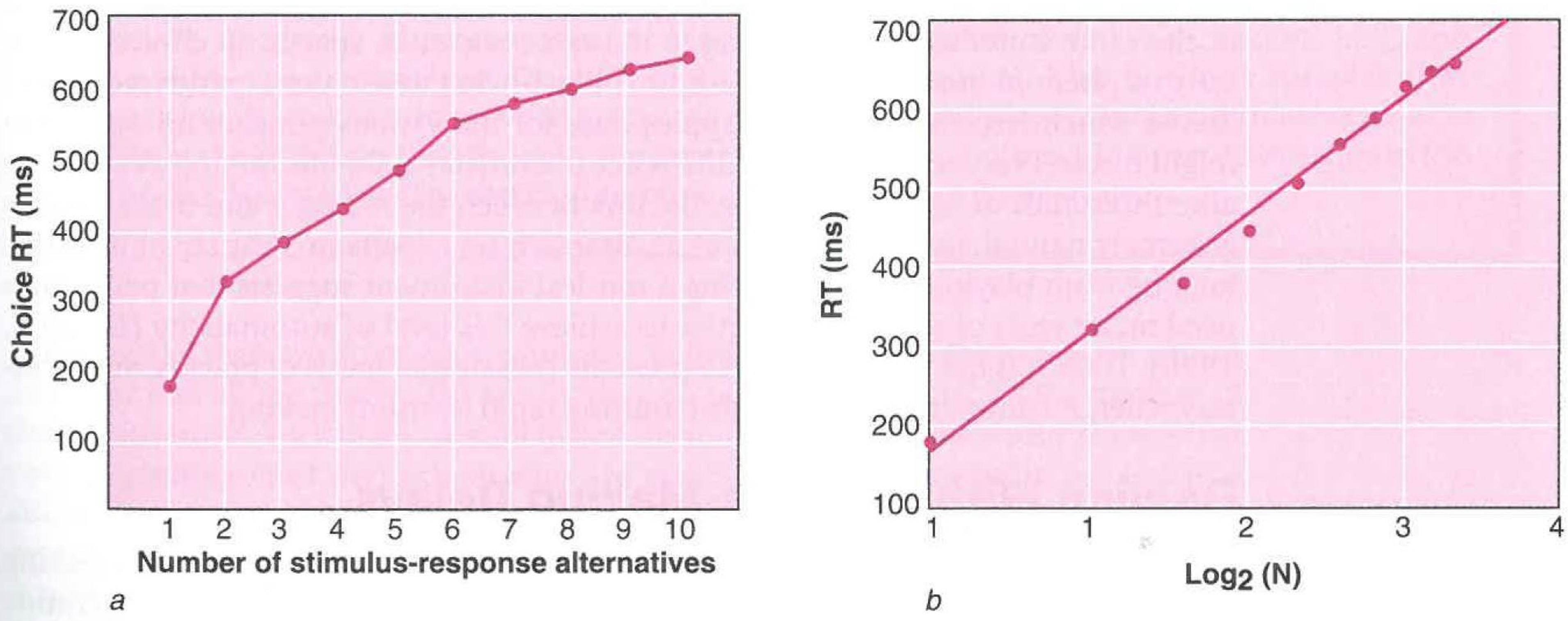


Figure 2.3 *(a)* The relationship between choice reaction time and the number of stimulus-response alternatives and *(b)* the linear relationship between choice reaction time and the $\log_2$ of the number of stimulus-response alternatives, which is the relationship known as Hick's law. (Adapted from Woodworth, 1938; data obtained by Merkel in 1885.)

Things affecting reaction time

- Hick's Law (number of alternative reactions)
- **Intuitiveness (stimulus-response compatibility)**
- Anticipation (animations matter)

"Intuitive use of products involves utilising knowledge gained through other experience(s). Therefore, products that people use intuitively are those with features they have encountered before. **Intuitive interaction is fast and generally non-conscious**, so people may be unable to explain how they made decisions during intuitive interaction (Blackler, 2008; Blackler, Popovic, and Mahar, 2002; Blackler, Popovic, and Mahar, 2003)"

Intuitiveness

- Stimulus-response compatibility
- Mouse look inverting
- Even more important with full-body control where common design patterns yet to emerge

Stimulus-response compatibility

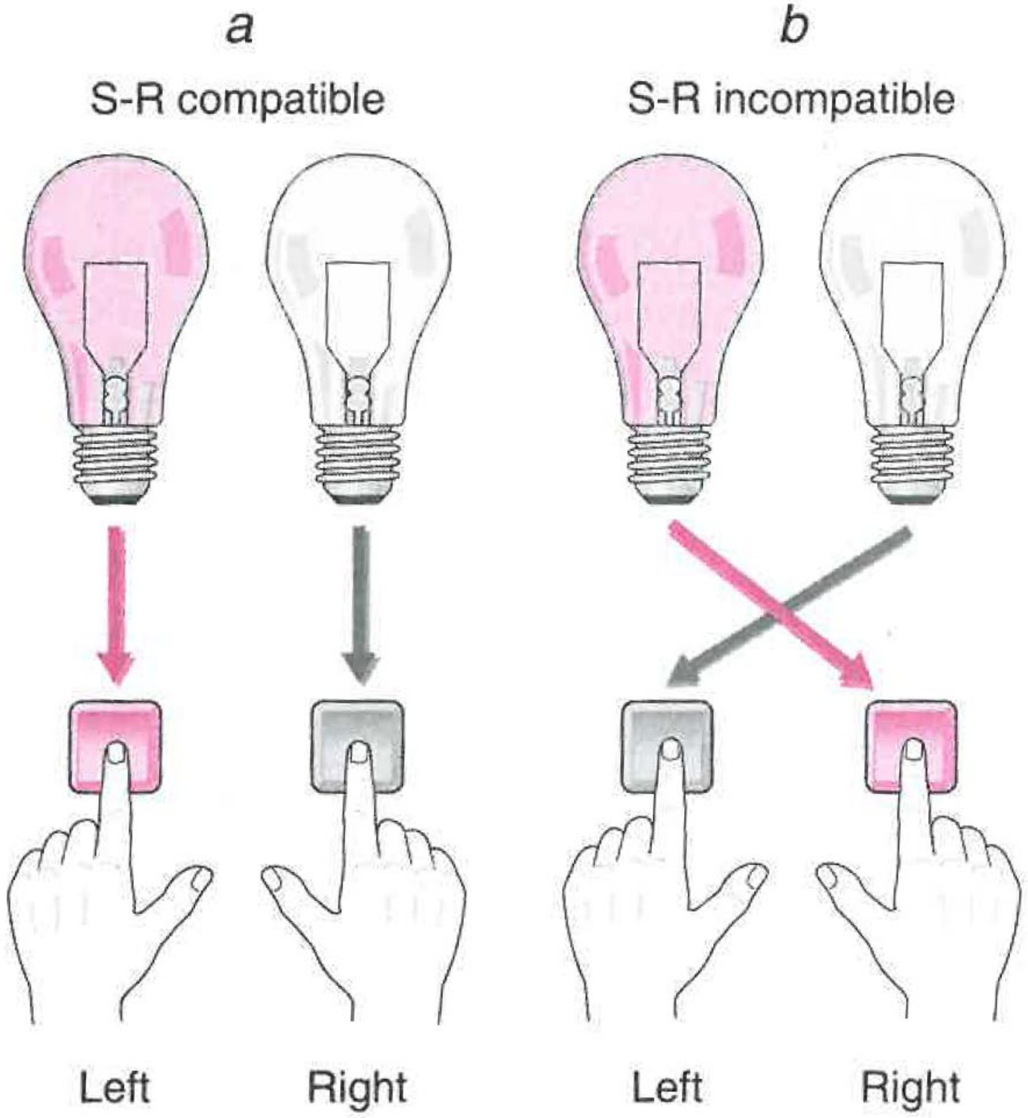


Figure 2.4 Stimulus-response compatibility. The relationship between stimulus and response is more natural, or compatible, in the situation on the left *(a)* than on the right *(b)*.

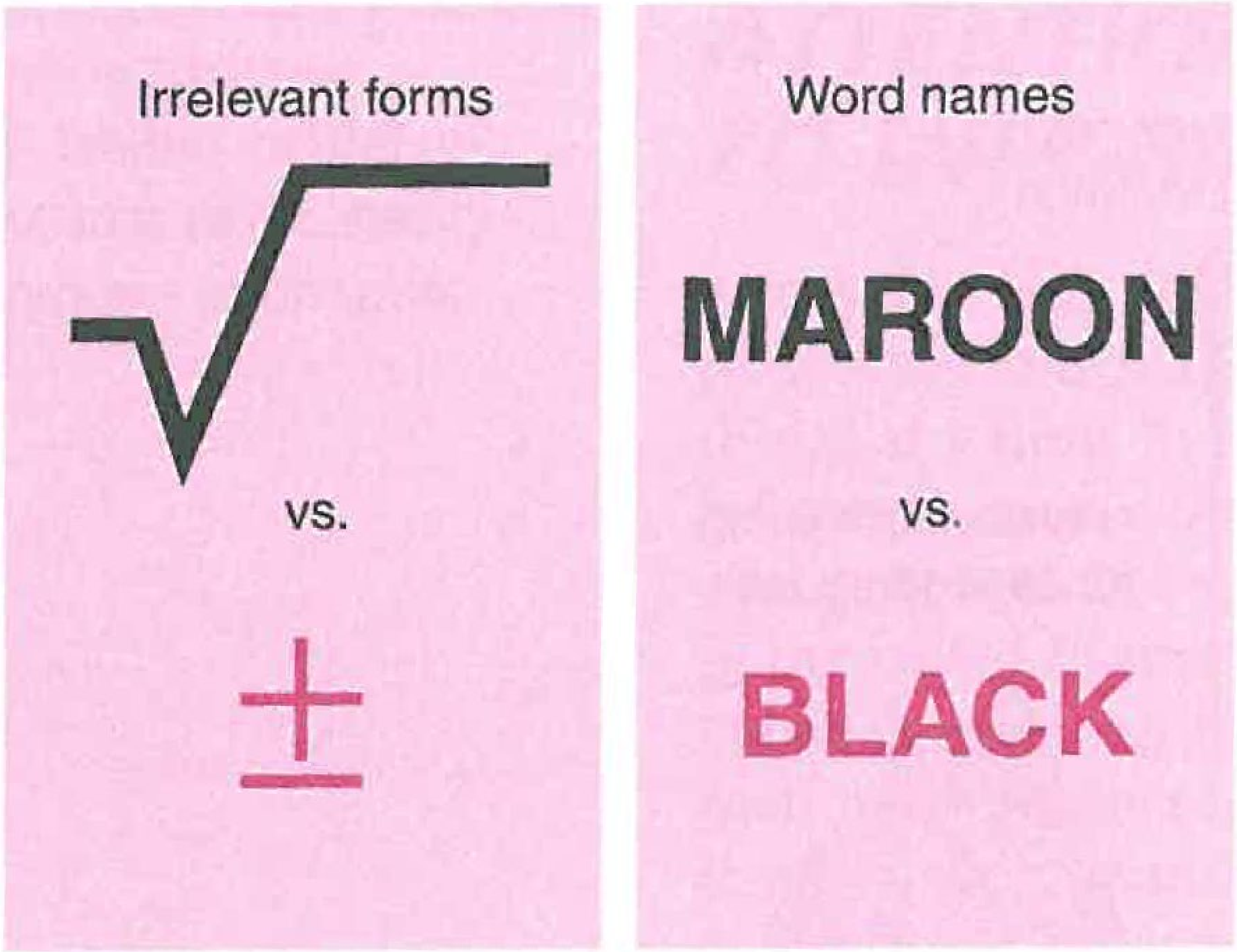
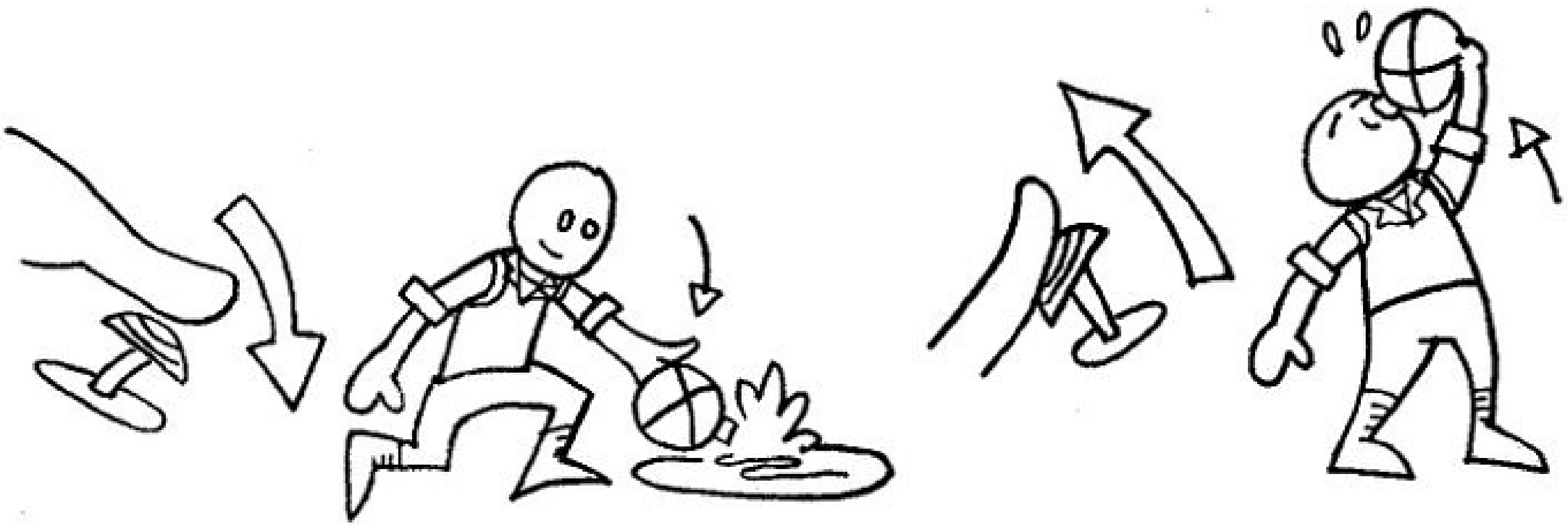
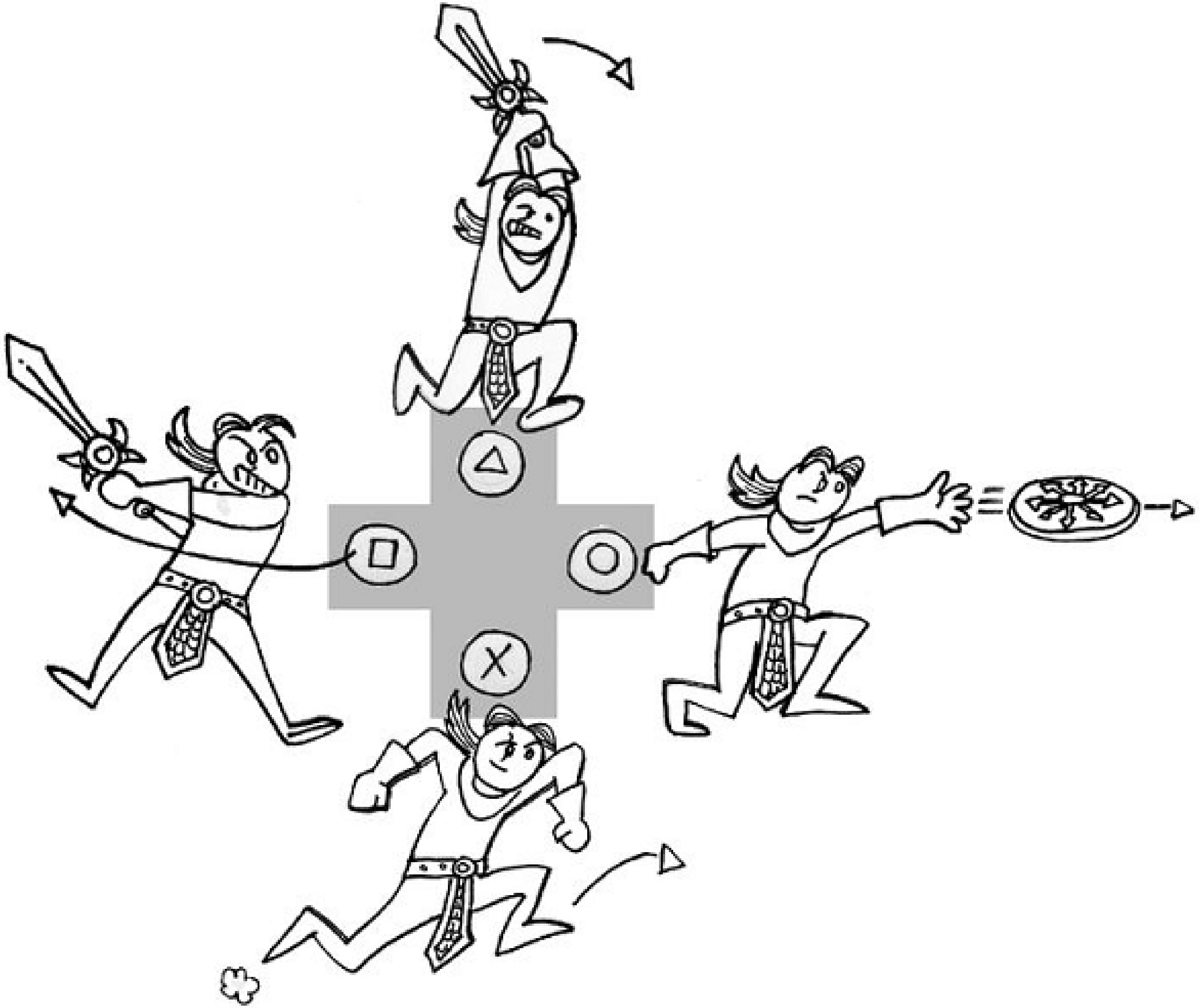


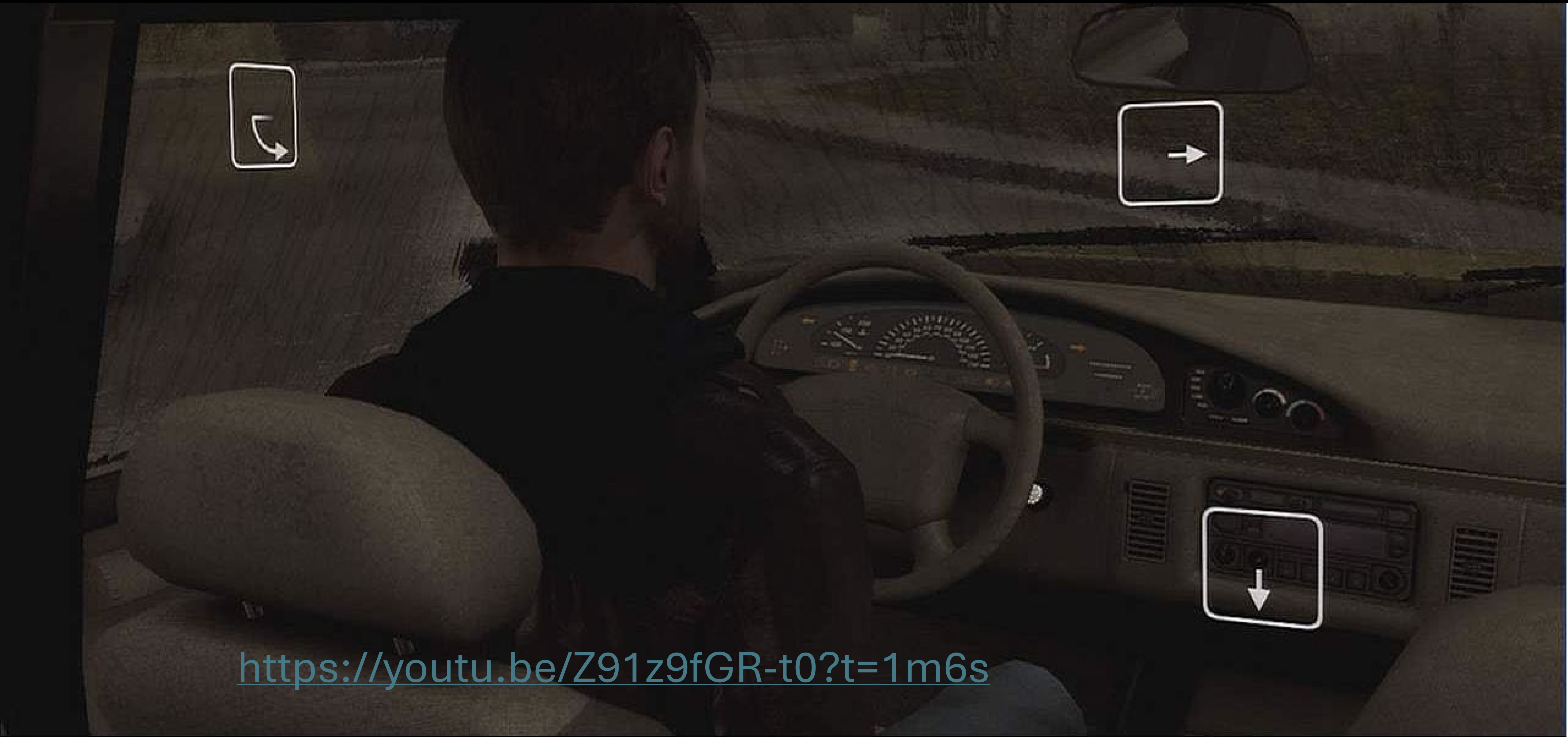
Figure 2.9 The Stroop effect. When ink color and form of stimulus conflict, reaction time to name the ink color is slowed.

Control-response compatibility





Heavy Rain



https://youtu.be/Z91z9fGR-t0?t=1m6s

Importance of "standard" controls

- One's reaction time and task completion time is heavily affected by learning
- Avoid making your players learn a new control scheme

=> Study games your players are probably familiar with, use same defaults

2.1. Perceptual-Motor Skills

Let us start with the historical case of Snoddy (1926). As remarked earlier, the task was mirror-tracing, a skill that involves intimate and continuous coordination of the motor and perceptual systems. Figure 1 plots the log of performance on the vertical axis against the log of the trial number for a single subject.

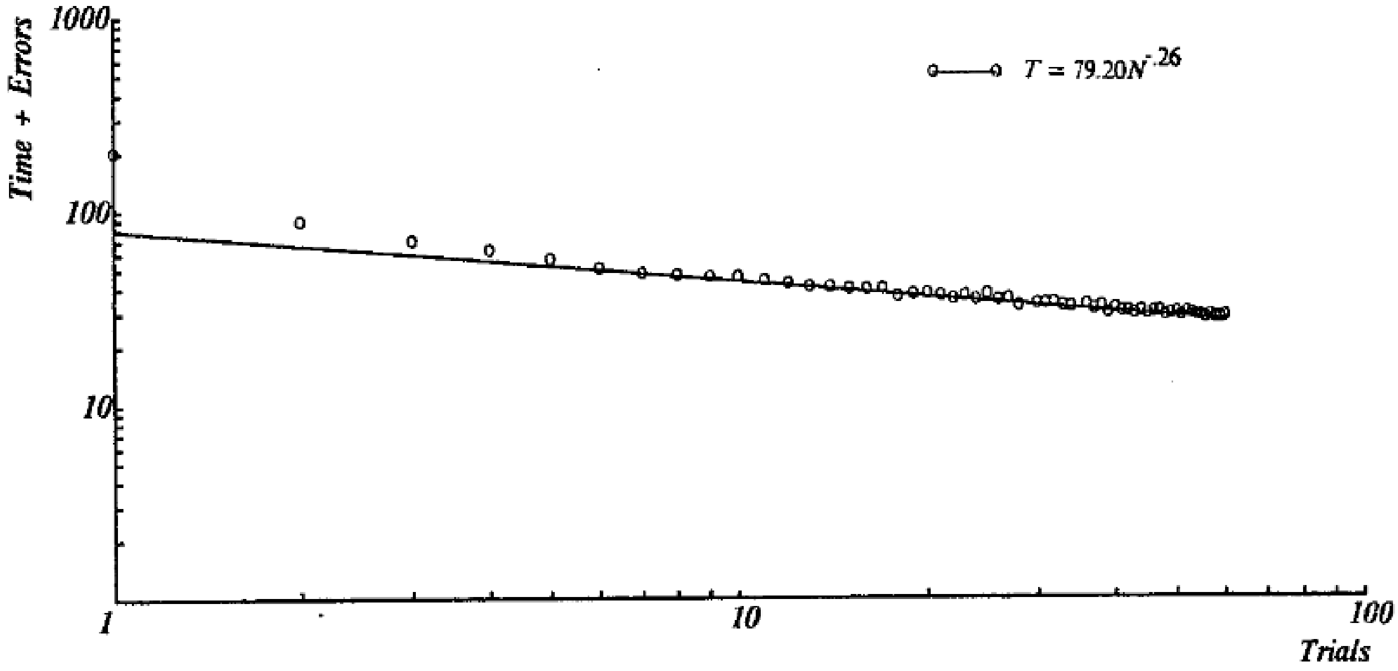
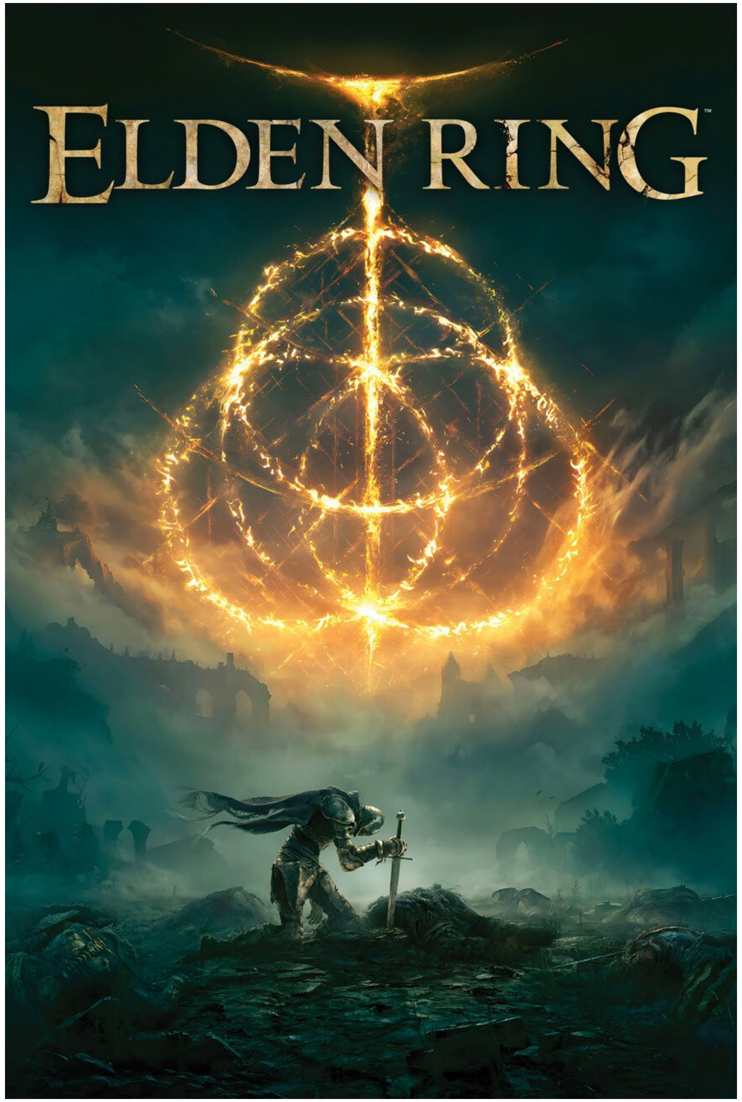


Figure 1: Learning in a Mirror Tracing Task (Log-Log Coordinates). Replotted from Snoddy (1926).

Newell & Rosenbloom 1981: Mechanisms of Skill Acquisition and the Law of Practice, http://shelf2.library.cmu.edu/Tech/7897434.pdf

See also: https://en.wikipedia.org/wiki/Power_law_of_practice

How to run and control the camera at the same time?



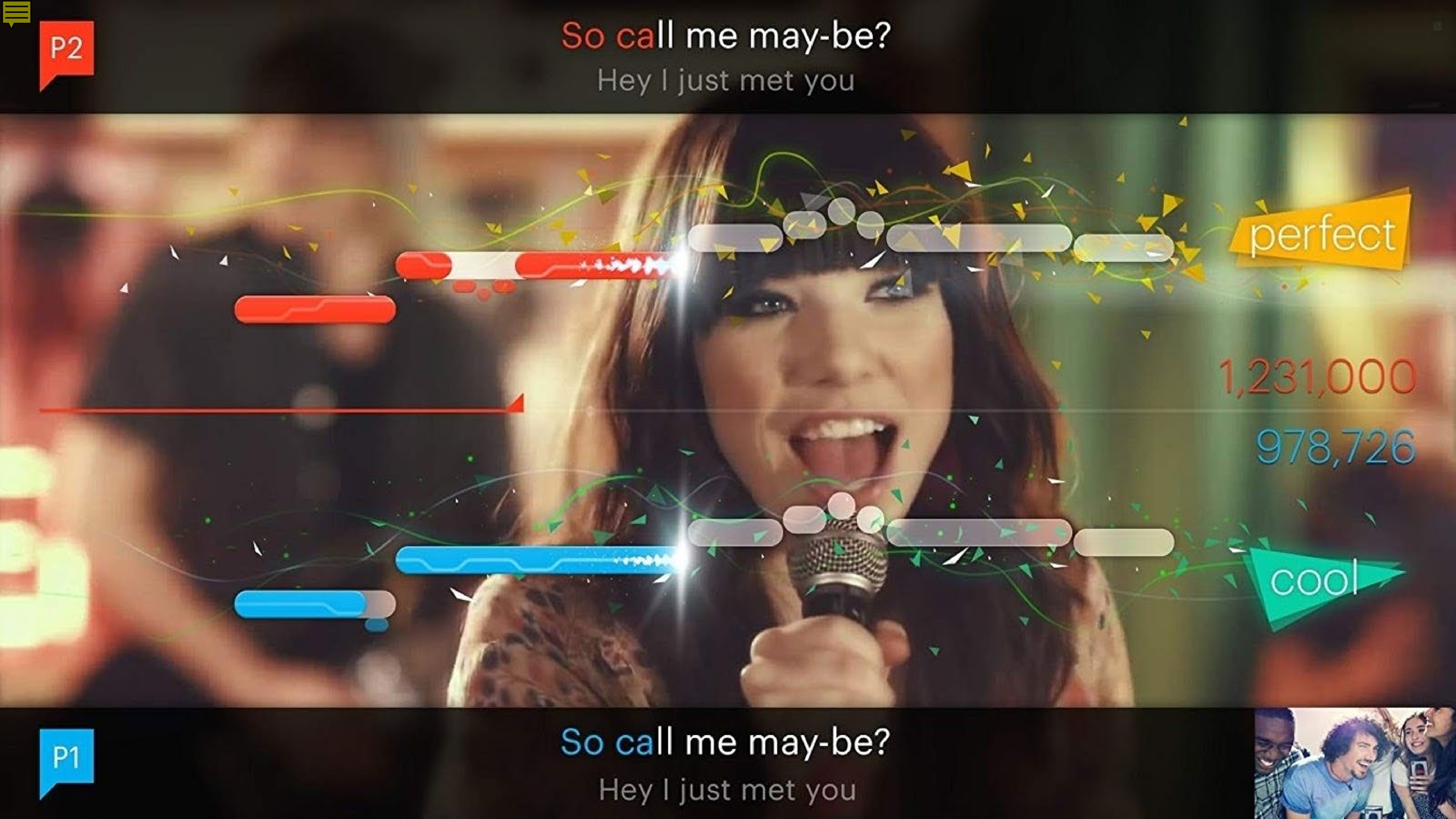
Things affecting reaction time

- Hick’s Law (number of alternative reactions)
- Intuitiveness (stimulus-response compatibility)
- **Anticipation**

Anticipation

- Spatial and temporal anticipation can make reactions both fast and controlled
- Orchestra starting synchronously (notes, count-in)

Supporting anticipation using a timeline


P2
So call me may-be?
Hey I just met you
perfect
1,231,000
978,726
cool
P1
So call me may-be?
Hey I just met you

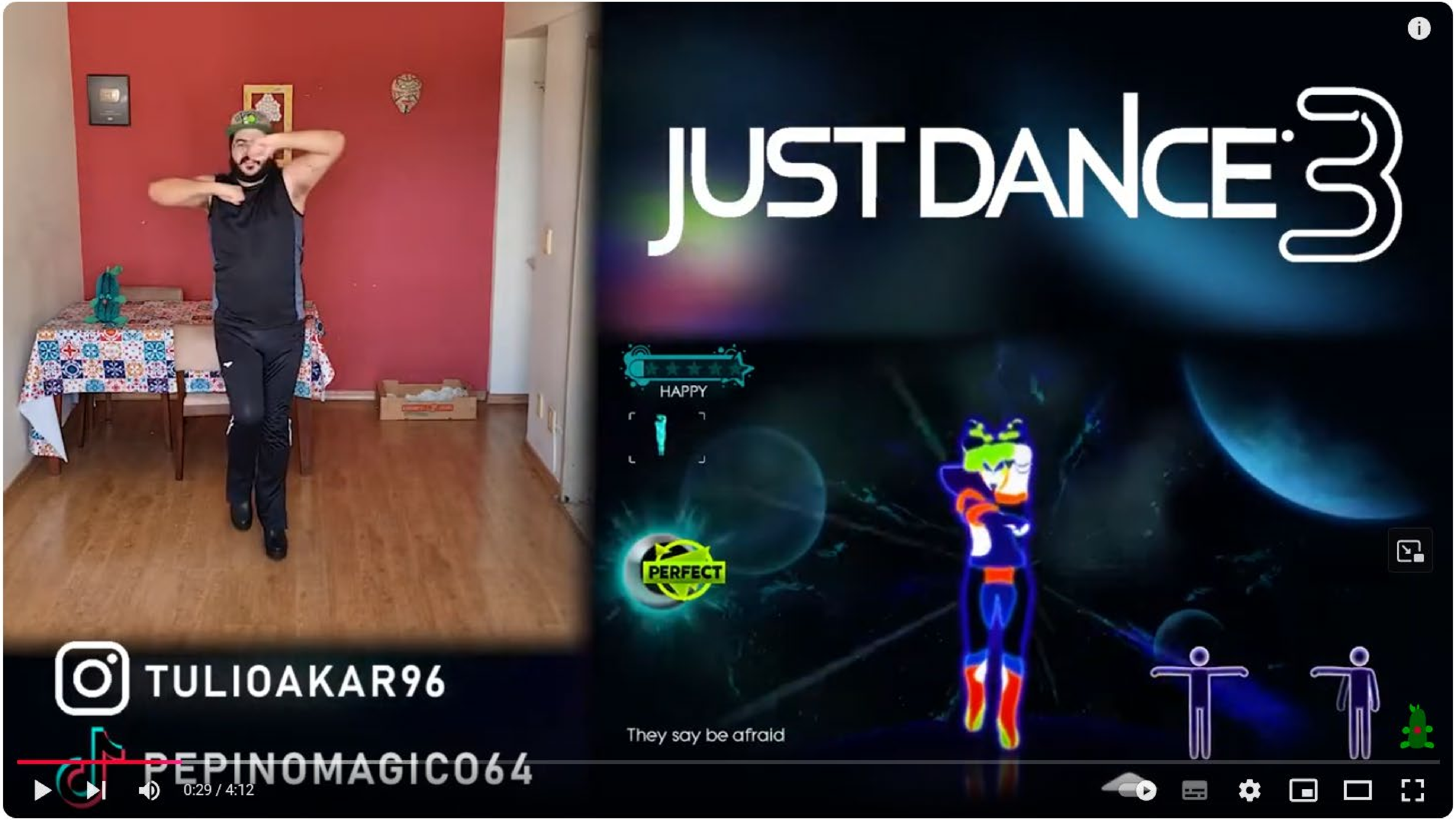
Dance games

- Without knowing the choreography in advance, can’t simply follow a model dancer
- Reacting with the whole body is even slower than reacting with button presses
- The human body is complex – can’t have SingStar –like curves for each degree of freedom
- How to provide anticipation?

Dance Dance Revolution

- Timeline only guides steps
- No guidance for the rest of the body

https://www.polygon.com/features/2017/12/11/16290772/the-rise-fall-and-return-of-dance-dance-revolution-in-america



Just Dance 3 - E.T. by Katy Perry [5 Stars] Gameplay Xbox 360 Kinect https://www.youtube.com/watch?v=PKBxoCIenpA

Dance Central - Lady Gaga Poker Face HARD 5 Gold Stars https://www.youtube.com/watch?v=TY2pA48yGkg

WAVE (CHI 2024)

- A new kind of full-body dance timeline by doctoral researcher Markus Laattala, an Aalto Game Design graduate
- Developed and validated through multiple user studies

WAVE: Anticipatory Movement Visualization for VR Dancing

Markus Laattala
markus.laattala@aalto.fi
Aalto University
Finland

Roosa Piitulainen
IT University of Copenhagen
Denmark

Nadia M. Ady
Aalto University
Helsinki Institute for Information Technology
Finland

Monica Tamariz
Edinburgh Napier University
Heriot-Watt University
United Kingdom

Perttu Hämäläinen
perttu.hamalainen@aalto.fi
Aalto University
Finland

Figure 1: Our proposed VR dance visualization technique. The user sees several model dancers with different time offsets, effectively becoming part of a crowd "making waves". The user can, therefore, mimic the moves of nearby dancers and anticipate upcoming movements. The image is a 3rd-person mixed-reality visualization with the user added inside the virtual world. See the supplementary video for a 1st-person view captured directly from a VR headset.

ABSTRACT

Dance games are one of the most popular game genres in Virtual Reality (VR), and active dance communities have emerged on social VR platforms such as VR Chat. However, effective instruction of dancing in VR or through other computerized means remains an unsolved human-computer interaction problem. Existing approaches either only instruct movements partially, abstracting away nuances, or require learning and memorizing symbolic notation. In contrast, we investigate how realistic, full-body movements designed by a professional choreographer can be instructed on the fly, without prior learning or memorization. Towards this end, we describe the design and evaluation of WAVE, a novel anticipatory movement visualization technique where the user joins a group of dancers performing the choreography with different time offsets, similar to spectators making waves in sports events. In our user study (N=36), the participants more accurately followed a choreography using WAVE, compared to following a single model dancer.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642145

CCS CONCEPTS

• Computing methodologies → Virtual reality; • Human-centered computing → Interaction techniques.

KEYWORDS

VR, dance instruction, dance game

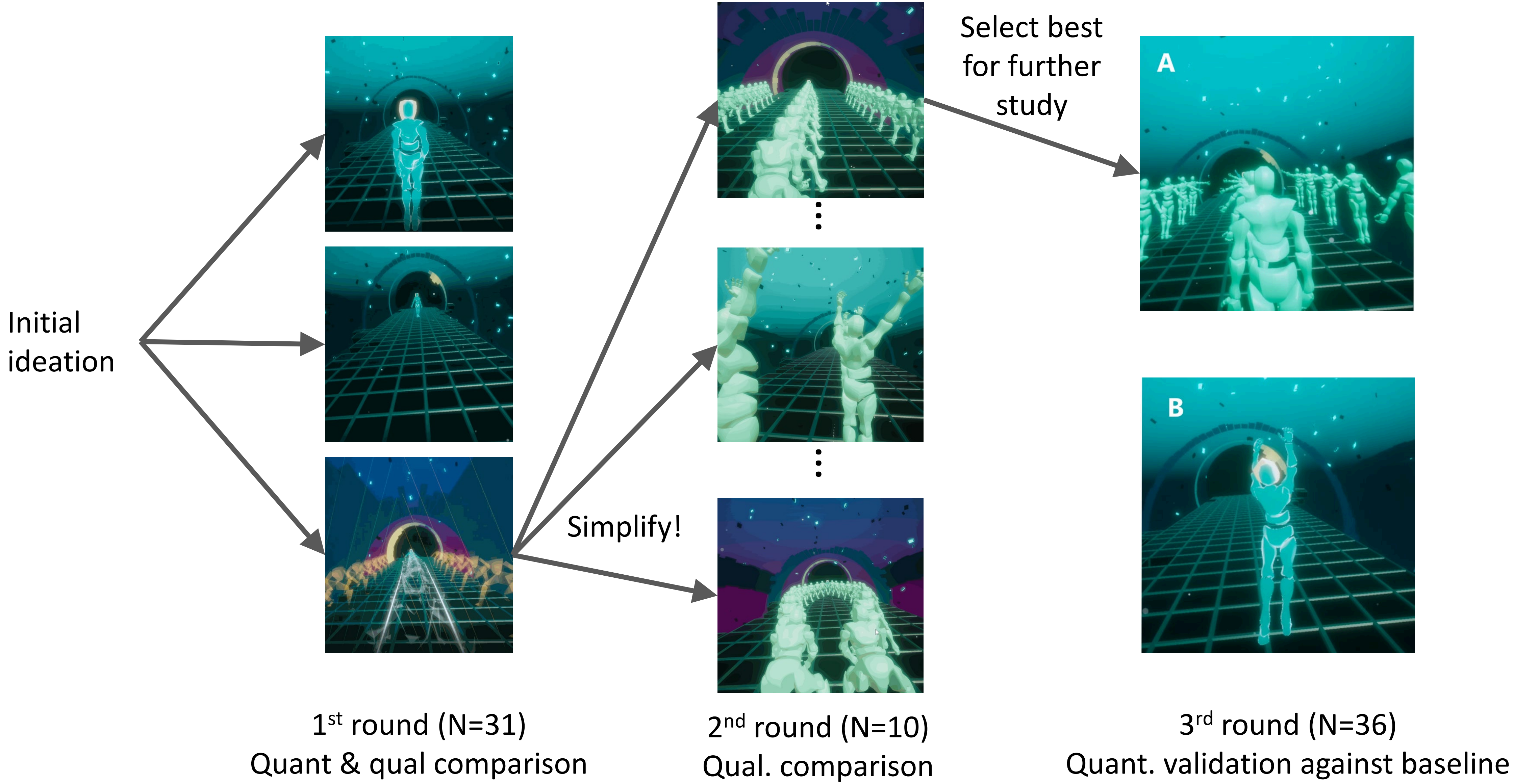
ACM Reference Format:
Markus Laattala, Roosa Piitulainen, Nadia M. Ady, Monica Tamariz, and Perttu Hämäläinen. 2024. WAVE: Anticipatory Movement Visualization for VR Dancing. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3613904.3642145

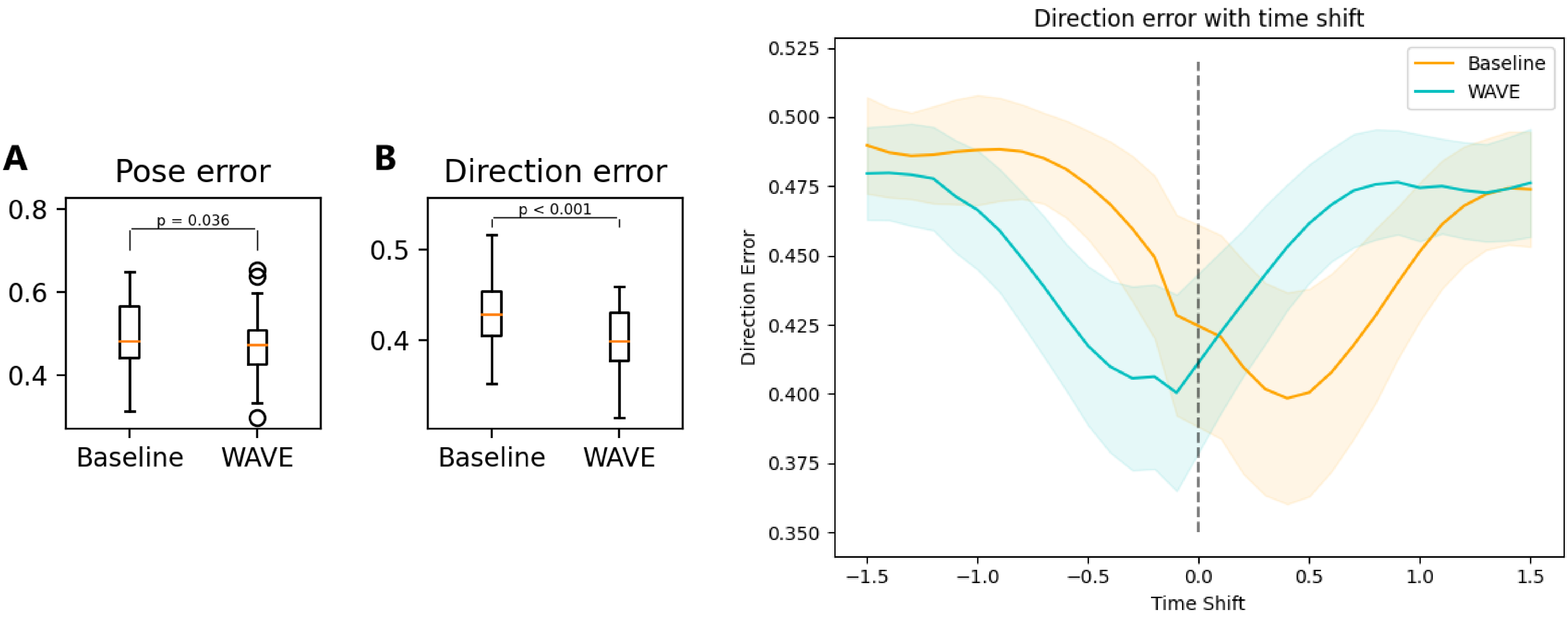
1 INTRODUCTION

Dance and rhythm games have emerged as one of the most popular genres in consumer Virtual Reality (VR). A prime example is Beat Saber [14], the best-selling VR title of all time [5], in which the player wields dual lightsabers to slice targets in rhythm. Dance has

WAVE: Anticipatory Movement Visualization for VR Dancing (CHI 2024 paper video) https://www.youtube.com/watch?v=xC4vm1unpC0

Iterative ideation, prototyping & evaluation

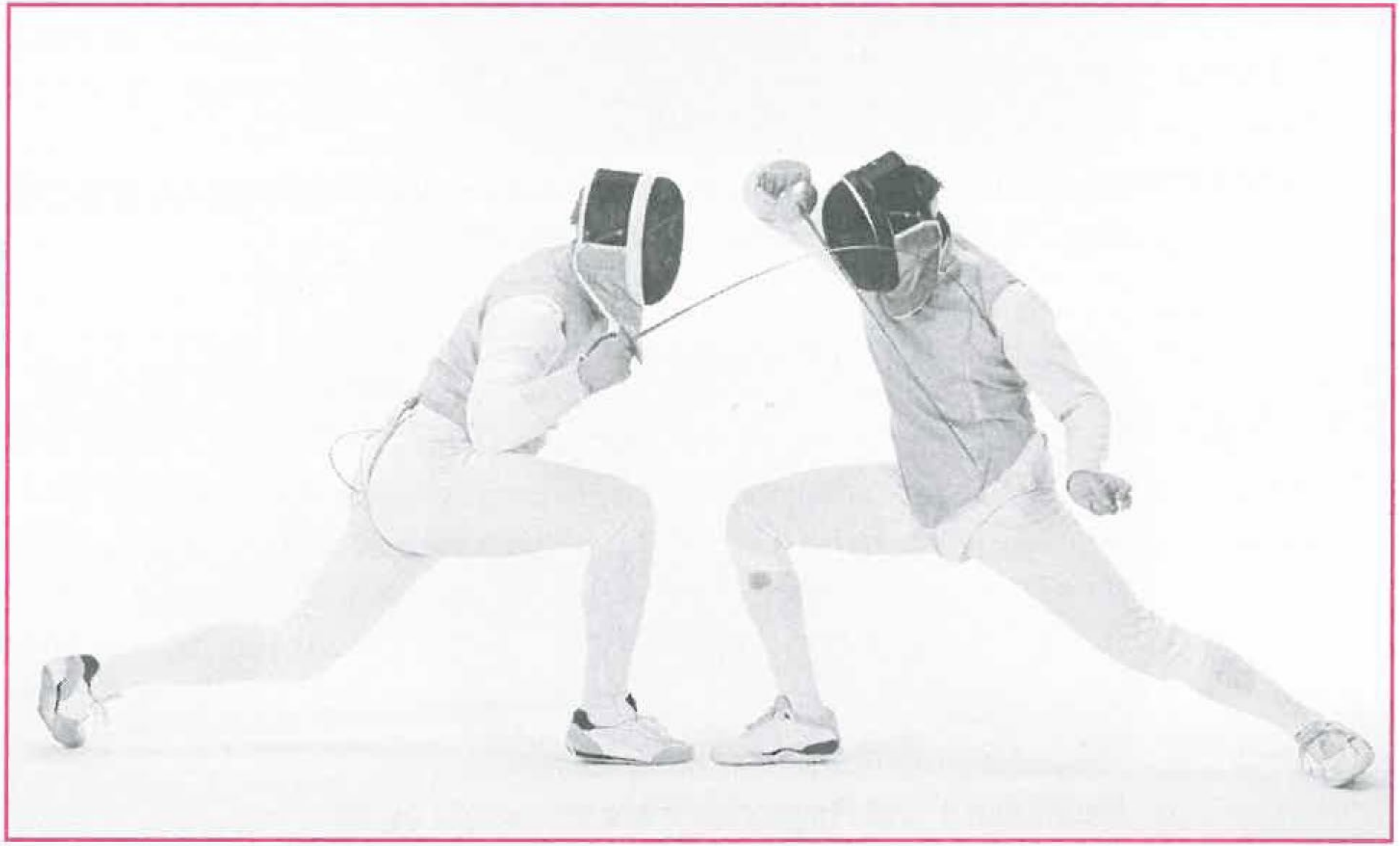



A
Pose error
0.8
0.6
0.4
p = 0.036
Baseline
WAVE
B
Direction error
0.5
0.4
p < 0.001
Baseline
WAVE
Direction error with time shift
Baseline
WAVE
Direction Error
0.525
0.500
0.475
0.450
0.425
0.400
0.375
0.350
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
Time Shift

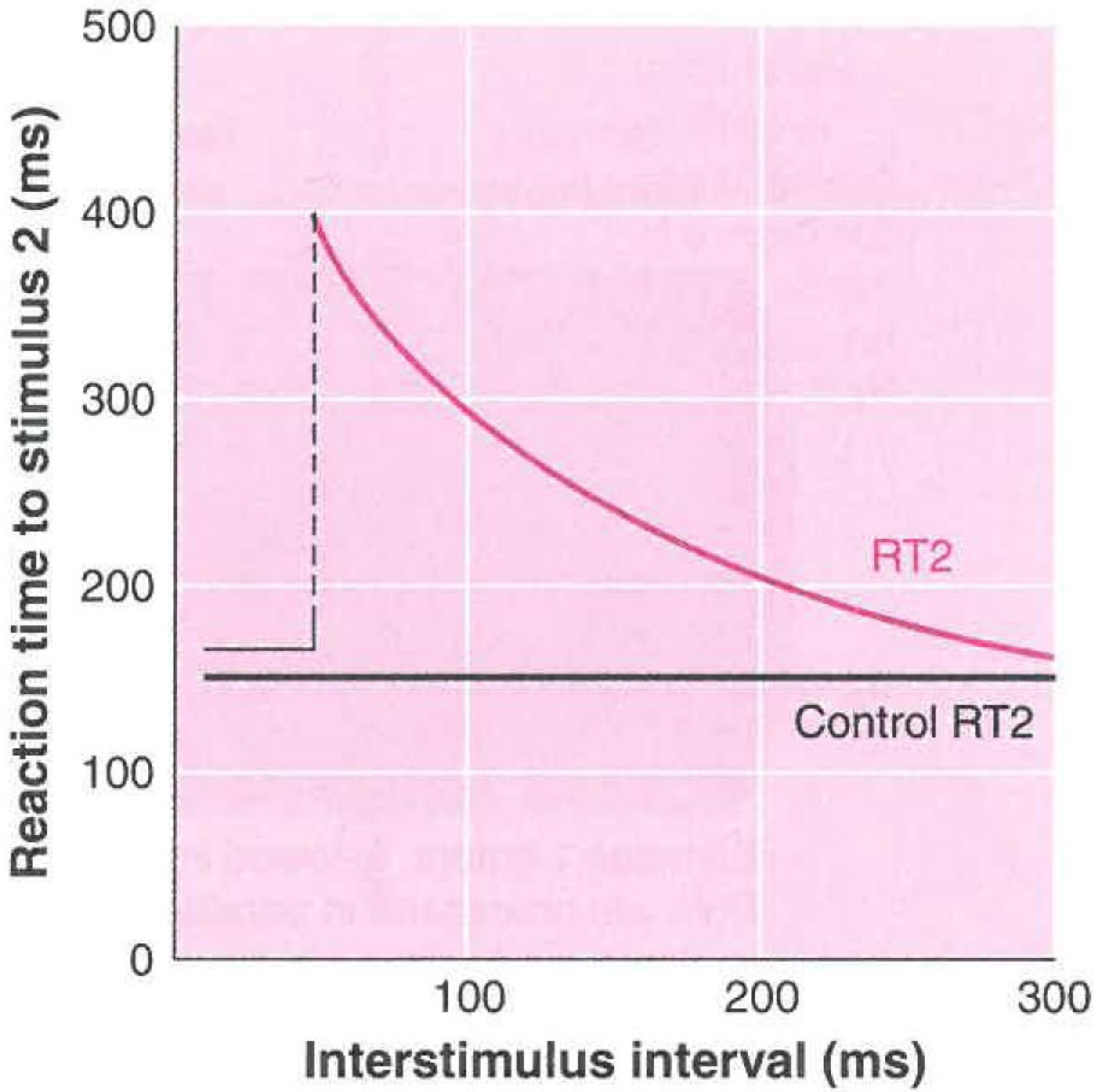
Problem solved?

- For this specific style of dancing, maybe.
- Can you reimagine dancing games in some other way?

Exploiting anticipation

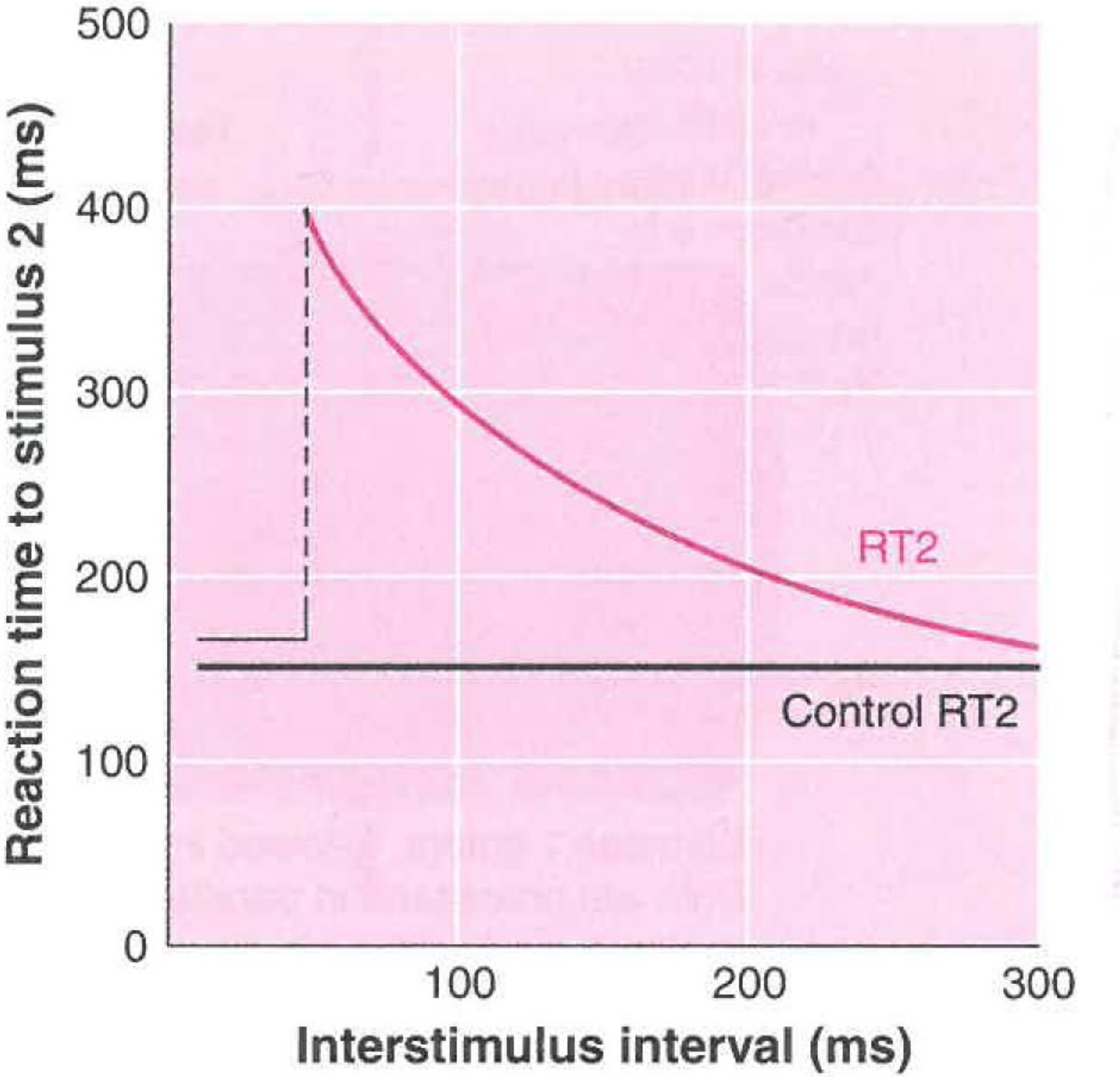


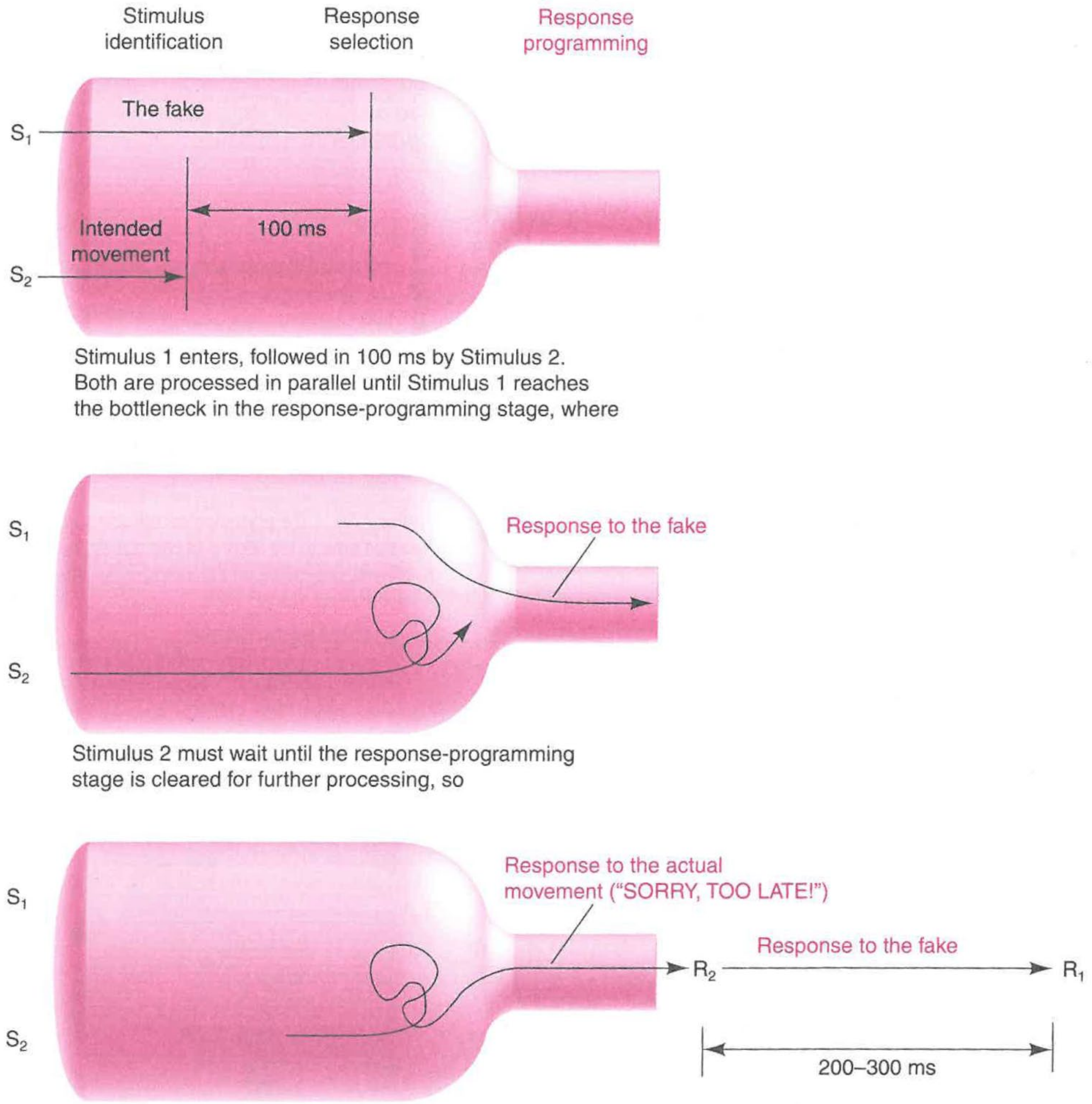
A faked movement followed closely in time by an actual movement can lure an opponent into an incorrect anticipation.


500
400
300
200
100
0
Reaction time to stimulus 2 (ms)
RT2
Control RT2
100
200
300
Interstimulus interval (ms)

Optimal fake timing

- The optimal fake duration is 50-60ms (in the simplified experiment with a short stimulus identification time)
- Shorter: no effect
- Longer: opponent may have completed the response to the fake and can react again





Stimulus 1 enters, followed in 100 ms by Stimulus 2. Both are processed in parallel until Stimulus 1 reaches the bottleneck in the response-programming stage, where

Stimulus 2 must wait until the response-programming stage is cleared for further processing, so

Response 1 and Response 2 are separated by far more than 100 ms.

The fake & games

- How common is faking in games?
- Can and should we provide opportunities and tutorials for faking in multiplayer games?
- How do the game animation systems support faking? Blends/transitions need to be optimized
- Faking and automatic fake detection as upgrades?

Faking in Clash Royale

- Optimal fake duration is longer than in, e.g., real-life boxing
- Card deploy latency (1s) increases pressure to react fast => fakes can be more powerful
- The slowly replenishing elixir means that reacting to a fake blocks one's further actions for a longer time.
- Optimal fake: Wait for the opponent to react, select the fake such that the opponent commits a large amount of elixir

Recap: Things affecting reaction time

- Hick’s Law (number of alternative reactions)
- Intuitiveness (stimulus-response compatibility)
- Anticipation (animations, UI design with anticipatory elements such as a timeline visualization)

Exercise: Fight choreography

- How to allow the player to execute semi-interactive pre-choreographed fight with high tempo?
- Nice choreography but slow: Uncharted 3 bar fight: https://youtu.be/7Z-fipEWLUU?si=GfY0Au0liHO9W0t5&t=216
- Goal: Fast-paced Asian martial arts such as Once Upon Time in China: https://youtu.be/bD87CECfi8Q?si=BP_M86JXhCYxHmby&t=34

Solutions (fast-paced action choreography)

- Slow down time and let the player see the future (reaction time, anticipation), then show a real-time replay
- Some form of timeline display that allows anticipation and planning of movements in advance

Topics

- Reaction time
- **Aiming, hitting, catching**
- Importance of automation
- Limits of attention

Fitt’s law

- Related to all kinds of aiming and reaching movements
- Movement time is proportional to $\log_2(2A/W)$, where A is the distance and W is width of target
- If distance is doubled, need to double the target width to maintain the same movement time

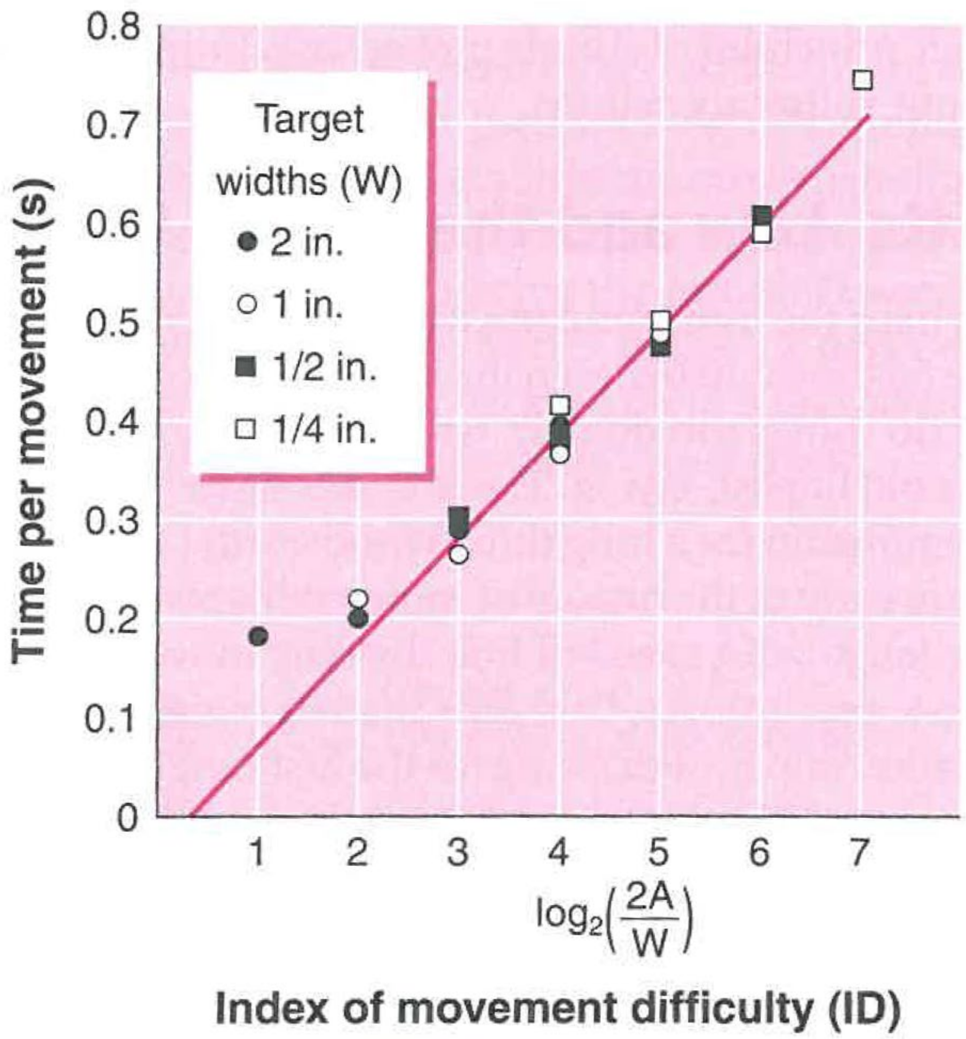


Figure 5.6 Fitts' law. The average time per movement is linearly related to the index of difficulty (ID). Target widths (W) are shown by the various symbols. For any given symbol, the movement distances (A) from left to right are 2, 4, 8, and 16 in., respectively.

Exercise: Sniper mode

- A basic sniper scope doesn't change the ratio of (motor space) distance and target size. Therefore, it does not improve aiming time (provided that the target is originally more than a few pixels wide)
- How can one design an aim assist system or sniper mode that provides an advantage based on Fitt's law? (Presemo)

Some solutions

- A chemical weapon that actually makes the heads of enemies bigger – cheap and funny
- Magnetic reticule – reduces distance by snapping to target when close enough
- Autotargeting – cycle targets with a button
- Steam controller: two parallel control methods (gyro and touchpad) that can use different sensitivities
- Fitt's law says that mouse sensitivity has no effect! (but: see slide notes)

https://archive.org/details/GDC2013Weihs

Topics

- Reaction time
- Aiming, hitting, catching
- **Importance of automation**
- Limits of attention

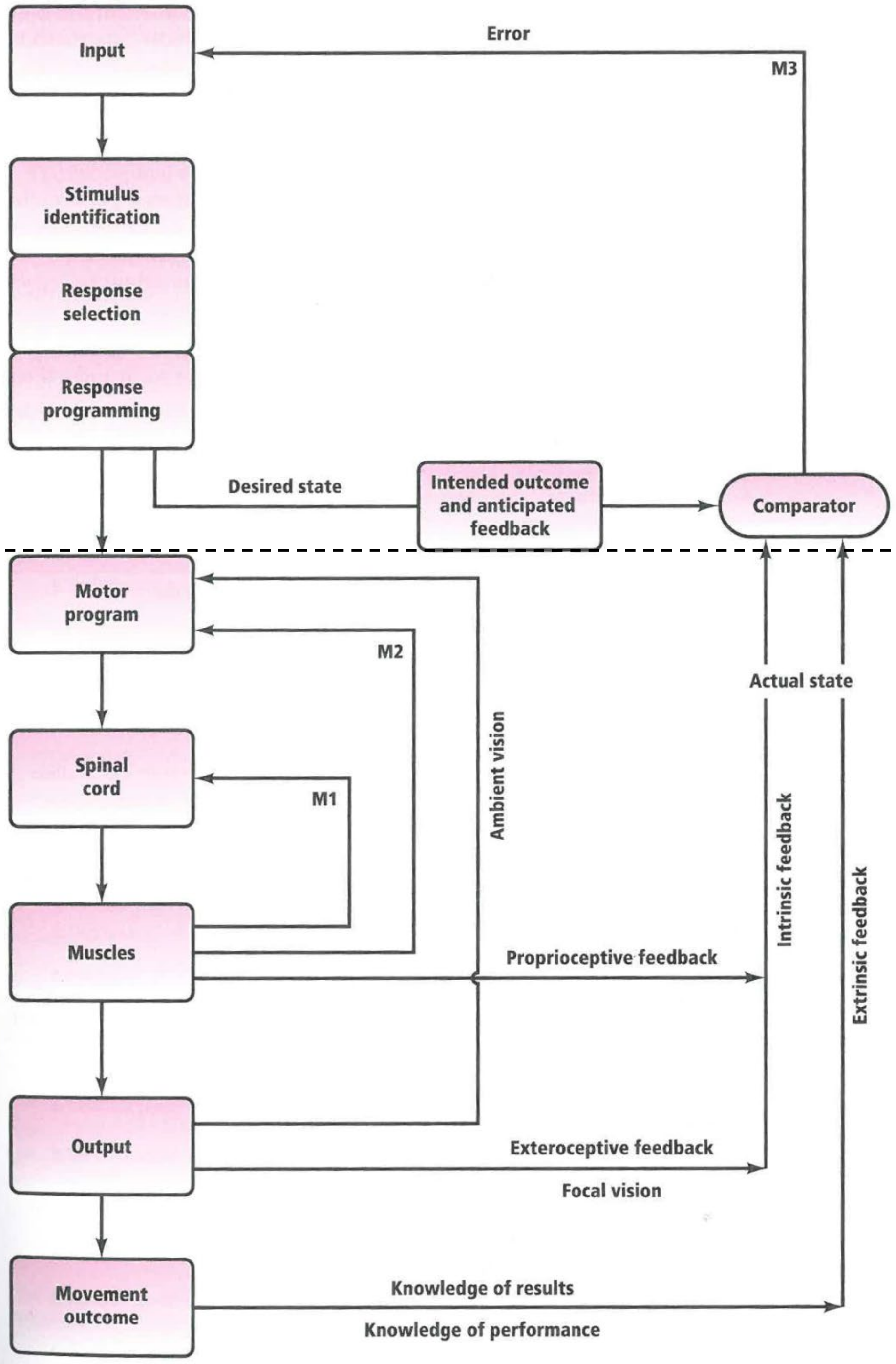
Automatic vs. conscious

- In both cognitive and motor learning, things become automated through repetition
- Conscious = slow, effortful, requires attention
- Automated = fast, effortless, requires no attention
- Again, it’s a continuum, not a clear-cut division

Example

Slow

Fast



Input
Error
M3
Stimulus identification
Response selection
Response programming
Desired state
Intended outcome and anticipated feedback
Comparator
Motor program
M2
Actual state
Spinal cord
M1
Ambient vision
Intrinsic feedback
Extrinsic feedback
Muscles
Proprioceptive feedback
Output
Exteroceptive feedback
Focal vision
Movement outcome
Knowledge of results
Knowledge of performance

Automatic vs. controlled/conscious

- Example: reading.
- Extreme example: reading aloud while thinking of something altogether different

Automatic vs. controlled/conscious

- Example: choking under pressure
- Skilled athletes perform best under automation
- Pressure may cause one to think too much, attempt to consciously control what should be automated. Focusing/trying too much may cause a failure (”lost move syndrome”)

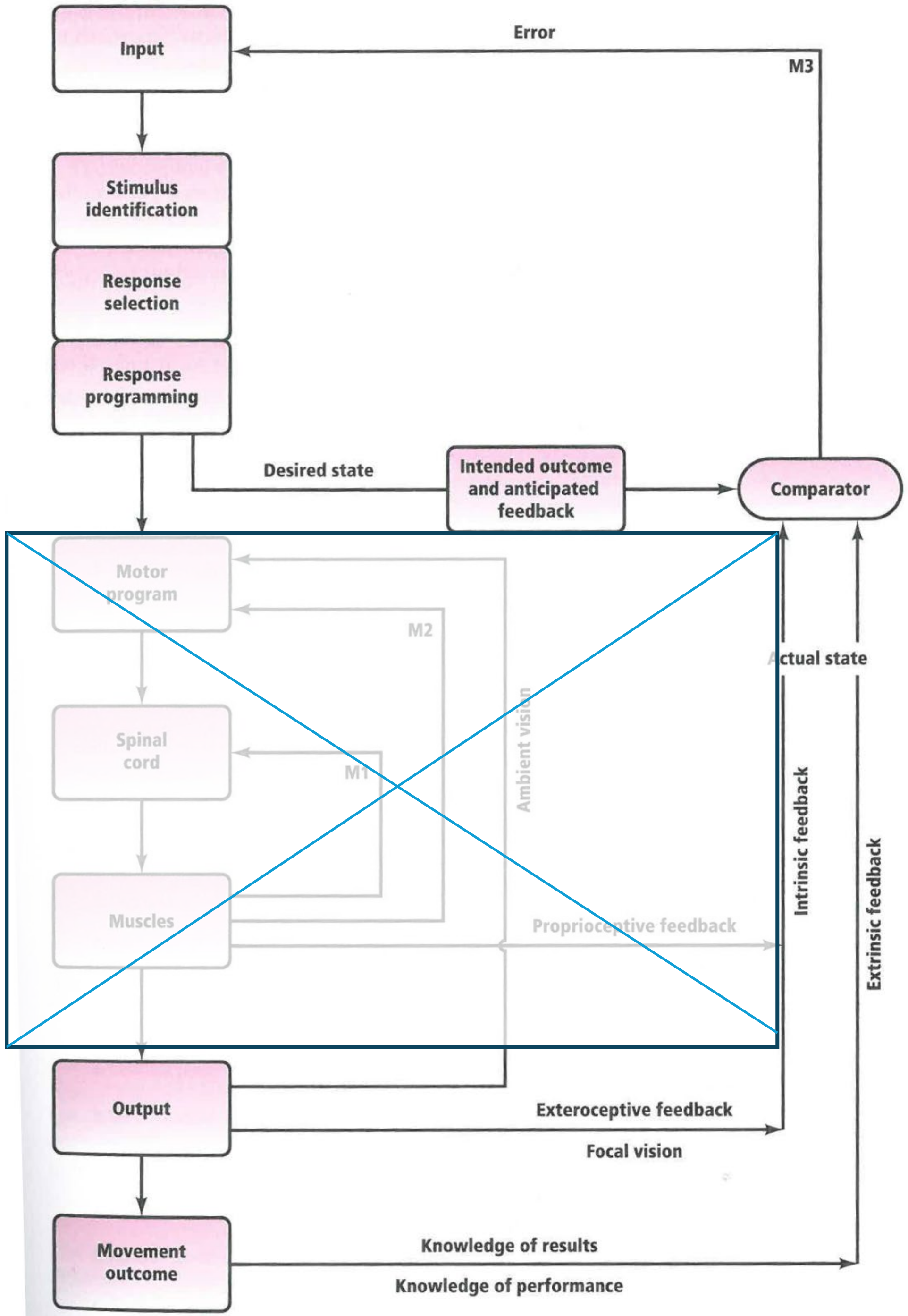
Games vs. Real life

- Games lack many sources of sensory information that contribute to automatic actions and unconscious control of motor skills
- Other than 1st person immersive VR games lack or distort ambient vision => must rely on consciously controlled movements or relearn control on an automatic level.
- Input latency increases reaction time
- Input-output latency slows down the closed control loop

Focal and ambient vision

- Focal: center of the visual field. Answers: What is it? Processed consciously.
- Ambient: the whole visual field. Used both consciously and on an automatic/unconscious way. Information about movement, e.g., time to target.
- Optical flow: the movement of light on the retina because of movement. Can be used to compute things about movement without identifying the moving object.

Optical flow


Input
Error
M3
Stimulus identification
Response selection
Response programming
Desired state
Intended outcome and anticipated feedback
Comparator
Motor program
M2
Actual state
Spinal cord
M1
Ambient vision
Intrinsic feedback
Extrinsic feedback
Muscles
Proprioceptive feedback
Output
Exteroceptive feedback
Focal vision
Movement outcome
Knowledge of results
Knowledge of performance

Example: catching and intercepting

- Catching and intercepting rapidly flying objects is not possible using closed-loop control
- It’s learned on an automatic level based on ambient vision (optical flow on retina)

Catching in a 3rd person view

Exercise: catching with impaired vision

- Together with a pair, throw and catch and object with a single hand in the following three ways:
  - Both eyes open (stereo and ambient vision)
  - Only one eye open (ambient but no stereo)
  - Looking through your fist with one eye (no ambient or stereo vision)
- 5 throws each. Count and write down the number of successful catches

Catching and VR camera design

- Master's thesis and conference paper by Felipe Marjalizo Alonso, now a senior game designer at Rovio
- We systematically varied the VR camera between 1st person and bird's eye 3rd person
- Test task: catching virtual balls (intercept a ball flying towards you with the VR controller)

Virtual Ball Catching Performance in Different Camera Views

Felipe Marjalizo Alonso
Aalto University
Finland
felipemarjalizo@gmail.com

Raine Kajastila
Aalto University
Finland
raine.kajastila@aalto.fi

Tuukka Takala
Aalto University
Finland
tuukka.takala@aalto.fi

Mikael Matveinen
Aalto University
Finland
mikael.matveinen@gmail.com

Mikko Kytö
Aalto University
Finland
mikko.kyto@aalto.fi

Perttu Hämäläinen
Aalto University
Finland
perttu.hamalainen@aalto.fi

ABSTRACT

Virtual camera design is an important but tricky part of creating virtual reality experiences; interaction can feel awkward if the camera is not placed exactly at the user's eyes, but on the other hand a 3rd person perspective (3PP) can provide a better view of the environment and/or the avatar. To inform camera design, we contribute the first study that systematically explores and quantifies how interaction difficulty changes when the camera is moved between a natural 1st person perspective (1PP) and a typical 3PP where the camera is behind and above the user. In our experiment, 24 participants catched flying virtual balls in seven different camera views. Catching performance degraded almost linearly as a function of camera distance from 1PP, and adaptation to non-1PP was slow or non-existent after a quick initial and partial adaptation. Our result suggest that natural 1PP should be used whenever possible, and transitions between views should be minimized to minimize the user constantly struggling to adapt. We also discuss how our results can be explained by the relation of camera perspective and retinal optical flow, and what interaction techniques can mitigate 3PP interaction problems.

CCS Concepts

•**Human-centered computing → Empirical studies in HCI;** *Virtual reality;*

Keywords

catching, intercepting, virtual reality, 3D user interfaces, camera design, 1st person, 3rd person

AcademicMindtrek'16, October 17 - 18, 2016, Tampere, Finland

ISBN 978-1-4503-4367-1/16/10…$15.00

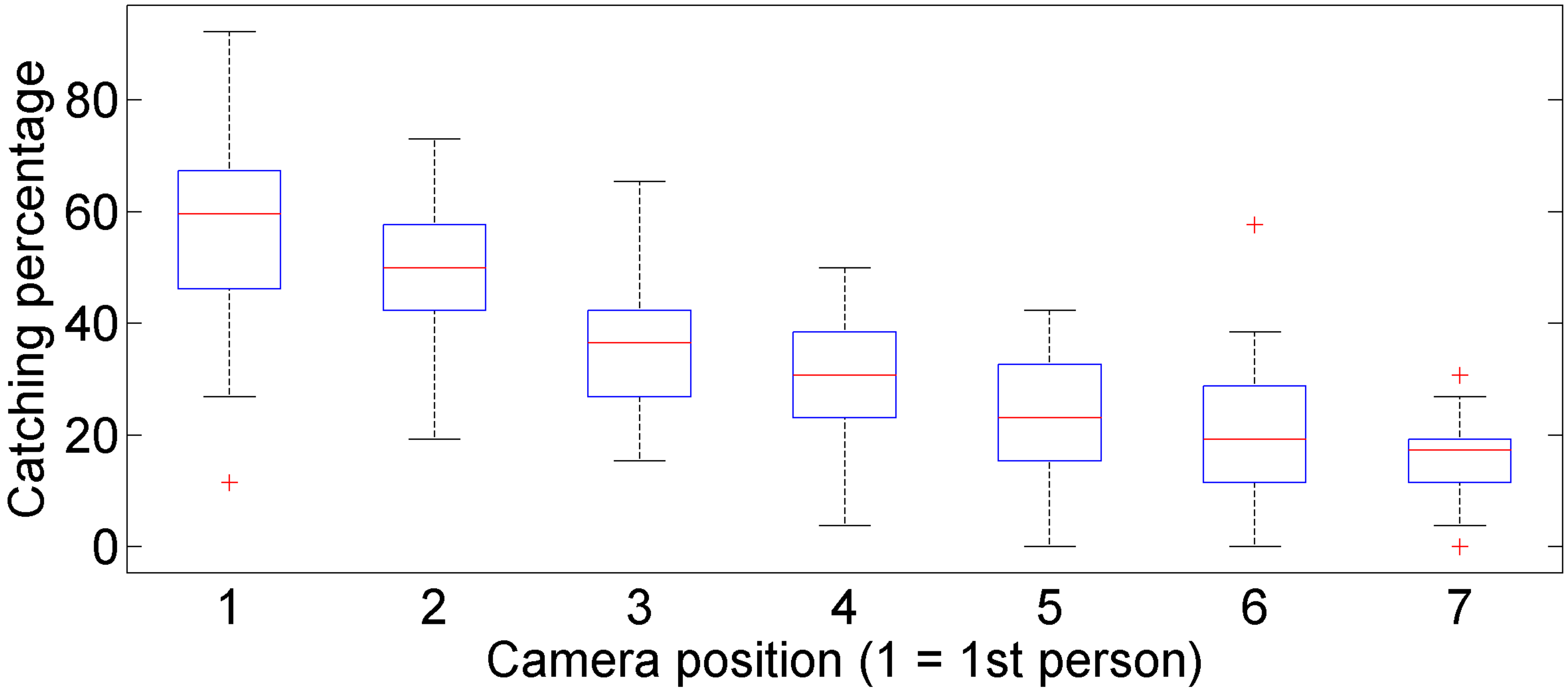
DOI: http://dx.doi.org/10.1145/2994310.2994335

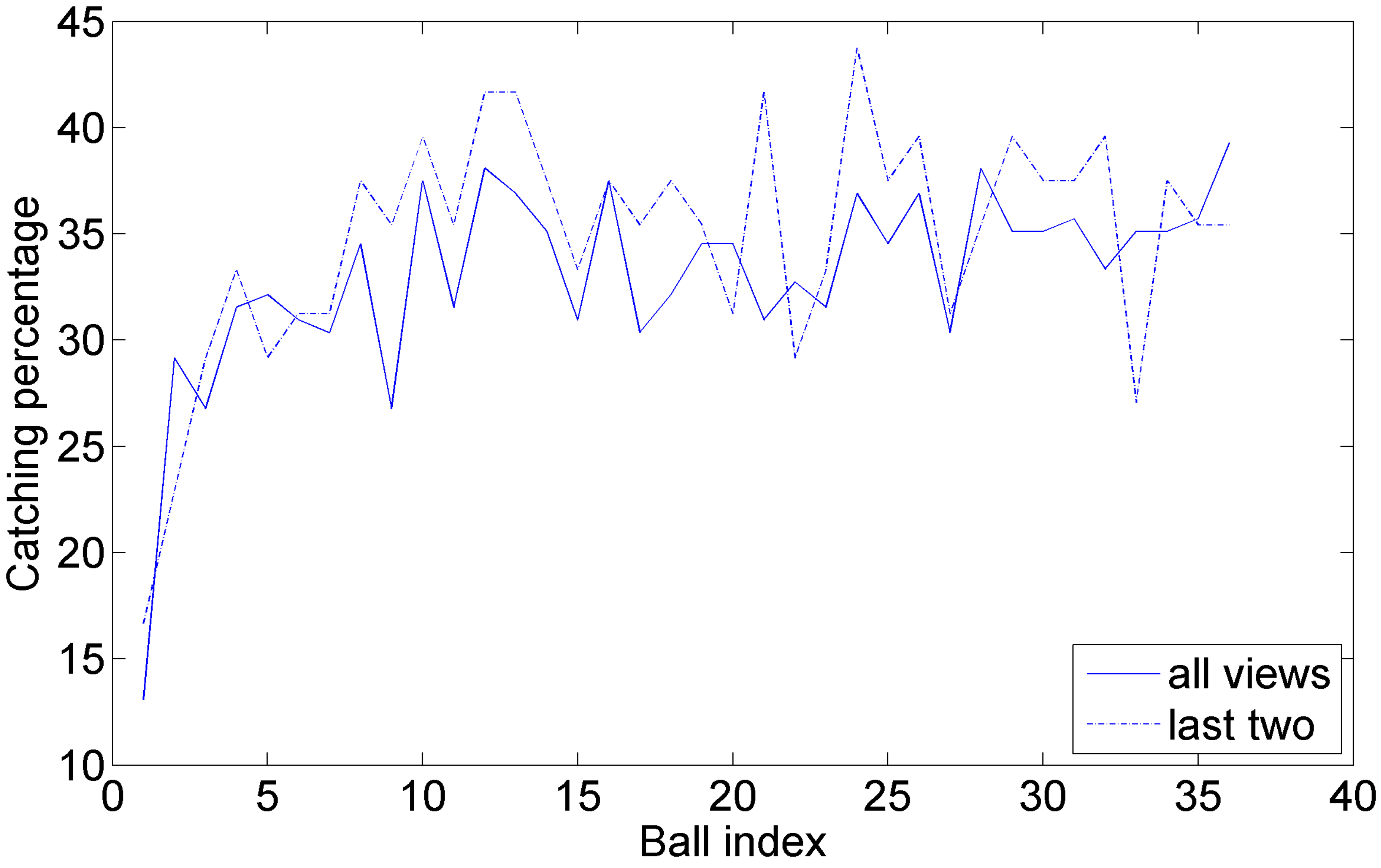
1. INTRODUCTION

3D user interfaces have become mainstream thanks to low-cost body tracking and virtual reality (VR) hardware such as Microsoft Kinect, Oculus Rift, and HTC Vive. Although low-latency head-tracked 1st-person perspective (1PP) VR is now possible with a modest budget, 1PP is not optimal for all game types. For example, navigating complex environments may be easier in a bird's eye 3rd-person perspective (3PP). The aesthetic experience of many action games also relies to a large extent on the player being able to see the appearance and movements of the avatar in 3PP. According to the review of gameplay videos of Kinect games by Takala et al. [36], 37% of the games rely primarily on 3PP, and there are also 3PP games among the Oculus Rift and HTC Vive launch games, e.g., Lucky's Tale, Chronos, and Edge of Nowhere.

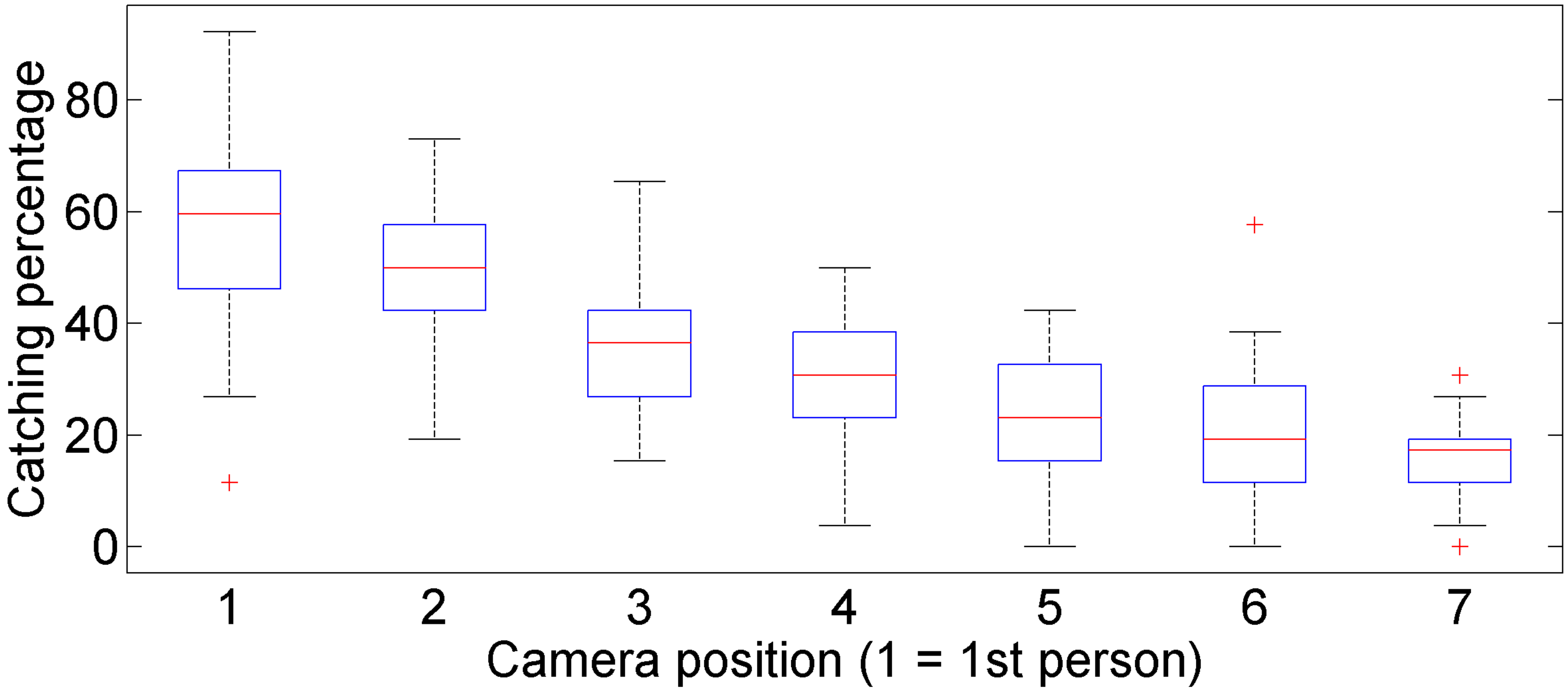
The downside of a 3PP viewpoint is that full-body human-computer interaction can be surprisingly difficult. For example, none of the eight participants in a study by Salamin et al. [31] could catch a ball thrown at them in 3PP. However, there is little research on how interaction difficulty changes when camera perspective is adjusted. We contribute such a study in the case of catching/intercepting flying objects, a task common in many ball sports and action games. The novelty of our study is in the combination of 1) comparing several viewpoints that form a continuum, in contrast to comparing just a few different views such as in Bateman et al. [2], and 2) measuring the effect of camera distance from 1PP instead of 3PP camera rotation, the latter studied previously by Ustinova et al. [40].

We used the setup shown in Figure 1 and Figure 2 which allowed us to freely move the camera between 1PP and 3PP. In our experiment, participants tried to catch flying balls in seven different camera views, with two examples shown in Figure 2. Our results show a gradual degradation of catching performance as the viewpoint is moved away from natural 1PP. The degradation is almost linear with no thresholds; there appears to be no specific distance from 1PP beyond which catching a ball would become significantly easier or more difficult. Further, adaptation to new camera settings was slow or non-existent after a quick initial and partial adaptation. Based on our results, we provide guidelines for designing camera perspectives, and discuss possible tech-


Catching percentage
0
20
40
60
80
1
2
3
4
5
6
7
Camera position (1 = 1st person)



Catching percentage
Ball index
all views
last two
45
40
35
30
25
20
15
10
0
5
10
15
20
25
30
35
40


Catching percentage
0
20
40
60
80
1
2
3
4
5
6
7
Camera position (1 = 1st person)

Dealing with impaired ambient vision

- For precisely controlled action, the player typically needs more time and artificial visual aids compared to real life
- Realistically blockable punches & kicks are not enough

=> Need to provide anticipatory cues!

Dealing with impaired movement precision

- Snap (slide) to target
- Anticipatory motion trails

Anticipatory movement trails

Snap (slide) to target

Snap (slide) to target

- Common in fighting games
- Contributes to effortlessness and competence
- Makes AI and animation design easier
- Allows more dynamic movement and use of a larger space

Topics

- Reaction time
- Aiming, hitting, catching
- Importance of automation
- **Limits of attention**

Multitasking and attention

- We can’t really multitask consciously
- In reality, we time-slice with our ”conscious CPU”
- Only automatic, unconscious actions can run in parallel

http://www.youtube.com/watch?v=vJG698U2Mvo

Kahneman: ” The gorilla study illustrates two important facts about our minds: we can be blind to the obvious, and we are also blind to our blindness.”

Attentional blindness and games

- Be conscious and cautious of where you guide the player’s attention and gaze. E.g., speech bubbles vs. subtitles
- Avoid making player blind to relevant feedback
  - Keep feedback visible until the player can pay attention
  - Provide feedback where the player is looking
  - If feedback not where the player is looking, draw attention to it. EXAGGERATE, ANIMATE!

How is it done right?

- Any successful F2P game: coins/gems/X fly from action to the numerical display
- Curved flight path => harder to predict, holds gaze (a trick magicians often use)
- Also consider panning/zooming/cutting the game camera to what the player needs to notice.

Topics

- Intro
- Reaction time
- Aiming, hitting, catching
- Importance of automation
- Limits of attention