

H20	H19	H18	H17	H16	H15	H14	H13	H12	H11	H10	H9	H8	H7
Taggert											Harrison		
G20	G19	G18	G17	G16	G15	G14	G13	G12	G11	G10	G9	G8	G7
Anahita		Brandon			Ziyan		Hesper			Yuan		Jason	
F20	F19	F18	F17	F16	F15	F14	F13	F12	F11	F10	F9	F8	F7
Alexis	Cheng			Joshua		Rachel		Don		Brian		Isabella	
E20	E19	E18	E17	E16	E15	E14	E13	E12	E11	E10	E9	E8	E7
Yang		Henry			Tyden		Andres		Mokshita		Thanh-Nha		Nilesh
D20	D19	D18	D17	D16	D15	D14	D13	D12	D11	D10	D9	D8	D7
Julian	Kayla			Nicholas		Tuan		Jacob		Michael		Ethan	
C19	C18	C17	C16	C15	C14	C13	C12	C11	C10	C9	C8	C7	C6
Ye		Nishanth			Anirudh		Yiteng		Liangchun		Yunyi		Zoe
B14	B13	B12	B11	B10	B9	B8	B7	B6	B5	B4	B3	B2	B1
Ross	Qisen			Abeer		Shreyas		Eric		Qingyang		Ainaz	
A14	A13	A12	A11	A10	A9	A8	A7	A6	A5	A4	A3	A2	A1

Row A = empty

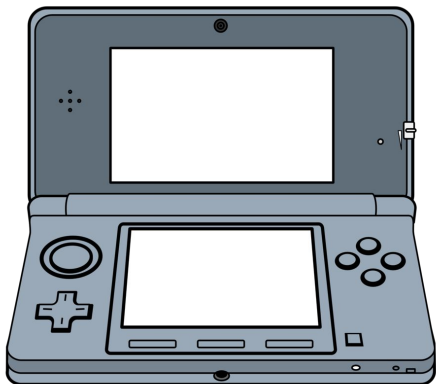
Week 7: Social Computing

Abeer Alessa, Ani Canumal, Ross DeVito



Labeling Images with a Computer Game

Luis von Ahn and Laura Dabbish



Motivation

- 2004 ACM Conference on Human Factors in Computing Systems (CHI)
- Old image retrieval systems were *text-based*
- They “annotate images with text derived from the HTML documents that display them” [Ahn 7]
- **Problem:** Traditional image labeling methods scale poorly
 - Manual Labeling was expensive and tedious
 - Q1: Why not use computer vision techniques?



Context: 2004 vs Today

Figure 5: Google query of “car” in 2004



Google query of “car” today



Motivation

- **Goal:** Build an interactive system that gamifies the image labeling process. It should:
 - Be fun!
 - Produce high quality labels
- **Key Concept:** Instead of manual labeling, leverage human instinct for stimulus and entertainment
- Paper was ahead of its time! Predates crowdsourcing platforms (Upwork and AWS Mechanical Turk)



duolingo



Discussion Question

What do you think are the main limitations of the traditional image labeling methods?

How could a gamified approach address these limitations in terms of engagement and efficiency?

The Extra-Sensory Perception (ESP) Game Overview

- Game Design
- describe image
- Teams are ranked
- **Taboo Words**
- **Skip!** Player
- **Sequence-k**



tch words to

y selected words

el

ESP Game Mechanics

- **Player Agreement:** Players guess words for an image until both agree on the same string, creating a “label” for the image
- **Entertainment First:** Players participate for entertainment, without explicitly being told they are contributing labor.
 - Ethical considerations?
- **High Fidelity Labels / “Good Label Threshold”:** game admins can account for meaningful labels by making sure a label is agreed on by at least **X pairs** before marking as taboo!
- *Q1: How does the game account for a label being “done”?*
- *Q2: Why are labels, never really “done”?*



Player 1 guesses: purse
Player 1 guesses: bag
Player 1 guesses: brown

Success! Agreement on “purse”



Player 2 guesses: handbag

Player 2 guesses: purse
Success! Agreement on “purse”

Figure 1. Partners agreeing on an image. Neither of them can see the other's guesses.

Evaluation

Usage / Fun

- Over 80% of players returned!
- 33 players spent over 50 hours playing

Labeling Rate

- How long would it take to label all of Google? (AVG=3.89 labels/minute)
- Paper (2004): 425,000,000 images in 31 days

Labeling Precision

- Manual assessment over subset indicates 85+% of labels are useful



Dog
Leash
German
Shepard
Standing
Canine

Figure 4. An image with all its labels.



Discussion Question

Can you think of other real-world examples that could benefit from a gamified approach to technology? What aspects of the gamified approach would you borrow/change from the ESP game?

Feel free to use examples from class or your own life.

4 month telemetry process

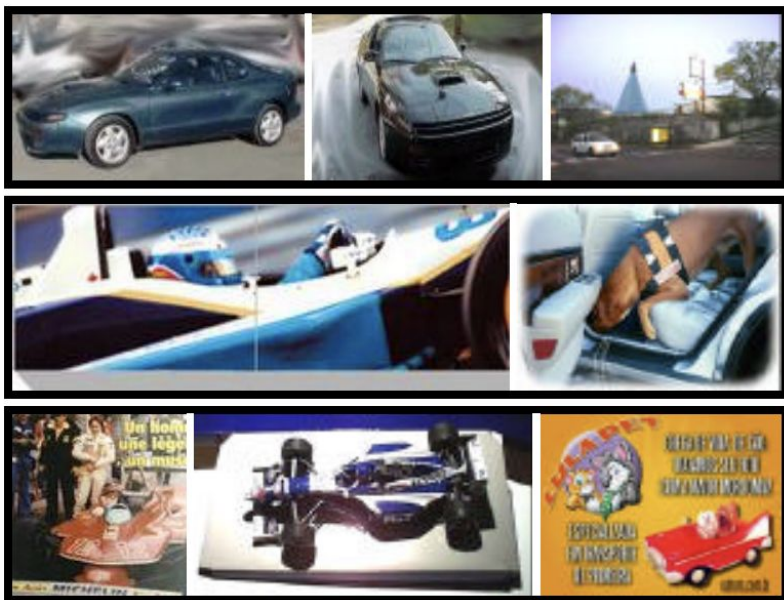


Figure 3. The first 14 images that **had** the label “car” associated to them by the ESP game (some of them have been slightly cropped to fit the page better).

Key Contributions

1. Free, scalable image labeling: predates crowdsourcing platforms
2. Pioneer of using gameplay to facilitate boring/difficult tasks
3. High-fidelity labeling: Introduced taboo words, random pair matching, and good-label thresholds to improve label quality

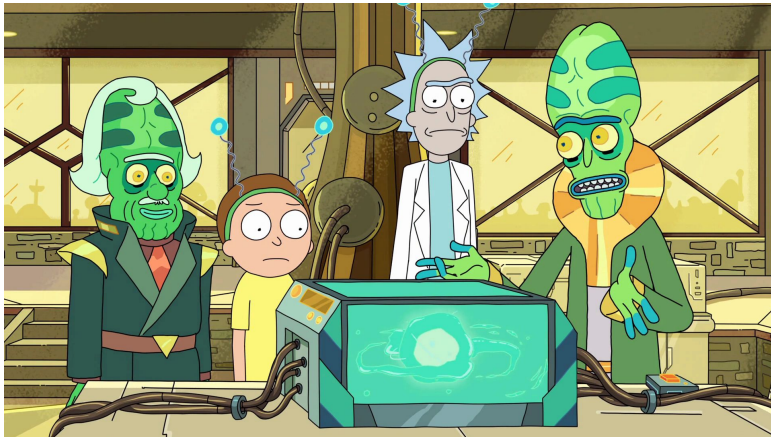
3:00

Discussion Question

The paper totally omits ethical considerations!

What are the social and ethical implications of extracting free labor from gamified image labeling?

How could designers make the game more transparent without compromising engagement or quality?



Limitations

- **Limited Incentive:** game relies on players not being drawn away by the myriad of other free, online games
- **Practical Limitations**
 - AVG=3.89 labels / minute
 - Paper: 425,000,000 images in 31 days
 - Today: 4.7 trillion images → 900 years!?!
- **Ethical Considerations**



Discussion Question(s)

Would something like the ESP Game still be effective today, given the massive increase in online data? If not, how might you scale such an approach today?

Compare and contrast the benefits/drawbacks of ESP with a non-gamified crowdsourcing platform like Mechanical Turk.

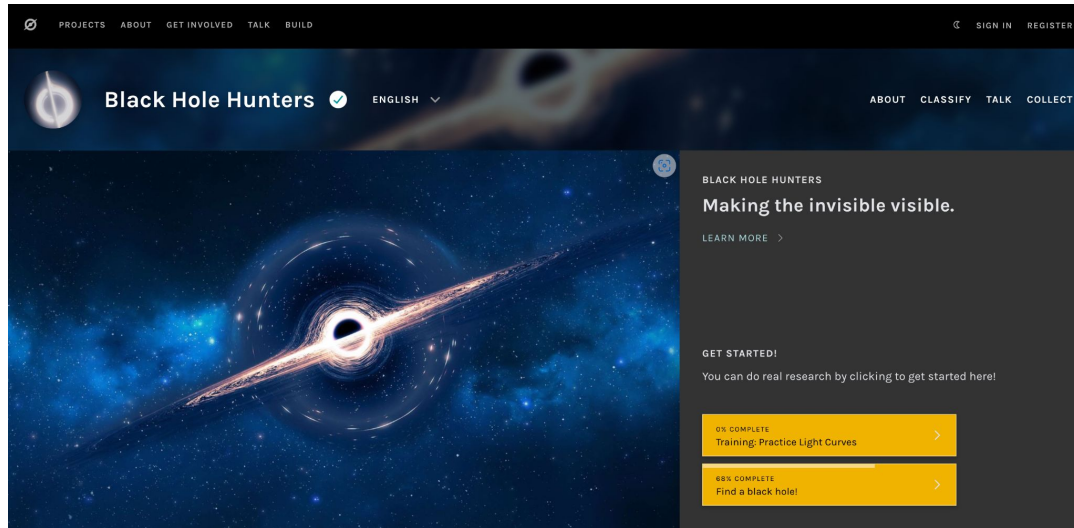
Discussion Question(s)

3:00

Would something like the ESP Game still be effective today? Why or why not?

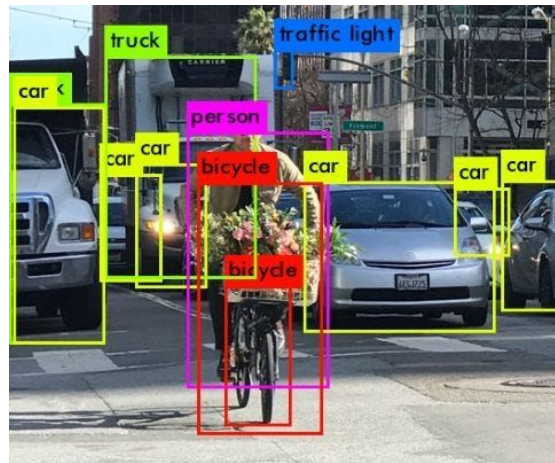
Have you ever participated in something like this before? If not, would you?

Consider efforts like <https://www.zooniverse.org/projects>



Broader Impact and Key Takeaways

- Gamified image labeling proven to be **cost-effective** and **high quality**
- Foundational work that led to future crowdsourcing platforms like MTurk
- Long-Term Influence on AI Systems: widespread availability of labeled datasets (ImageNet)



Your Takeaways

- Assumes “all of the users use the same language”, ignores “slang differences between geographical regions”
-



Discussion Question

What factors do you think contributed to the high engagement of players in the ESP Game? To what extent might these factors apply to social computing applications broadly?

Starting point: think about how competitive or cooperative game mechanics can steer player motivation and effort.

Using Social Psychology to Motivate Contributions to Online Communities

Gerard Beenen, Kimberly Ling, Xiaoqing Wang, Klarissa Chang, Dan Frankowski, Paul Resnick, Robert E. Kraut

Using Social Psychology to Motivate Contributions to Online Communities

- 2004 ACM conference on computer supported cooperative work (CSCW)
- **Goal:** Motivate contributions in online communities using social psychology
 - Typically just a small minority contribute
 - Communities die from under-contribution
 - ***Social loafing:*** people exert less effort on collective tasks than individual ones
- Performed two studies to evaluate hypotheses derived from social psychology

The role of designers

Encourage more contributions from a broader group of people by:

- Designing technical features
- Seeding the social practices of the community

For studies in paper, designed email campaigns to motivate contributions in different ways

MovieLens

- Web-based movie recommendation service
 - Run by University of Minnesota, University of Michigan, and Carnegie Mellon University
- Members rate movies and are recommended new movies based on the ratings of others
- Requires contributions to improve individual and overall suggestions
 - 20% of movies have too few ratings for accurate prediction

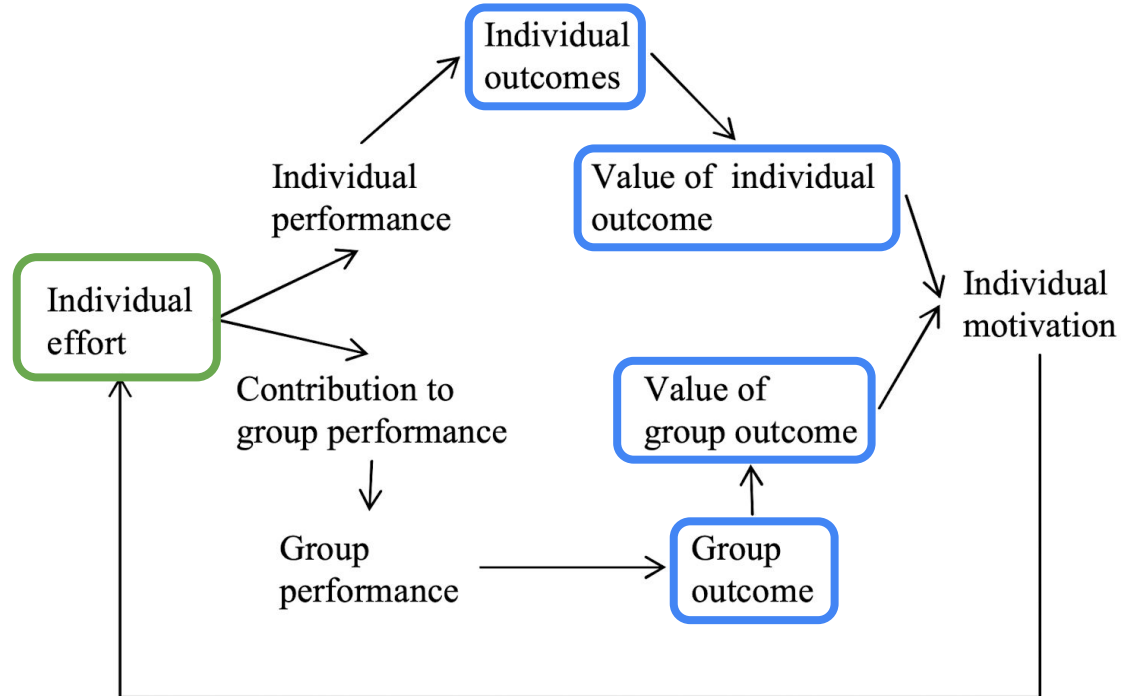
The logo for MovieLens, featuring the word "movielens" in a white, lowercase, sans-serif font on an orange rectangular background.

Non-commercial, personalized movie recommendations.

Study 1: Framing uniqueness and benefit

Figure 1: The collective effort model (adapted from [13])

“people work hard when they think their effort will help them achieve outcomes that they value”



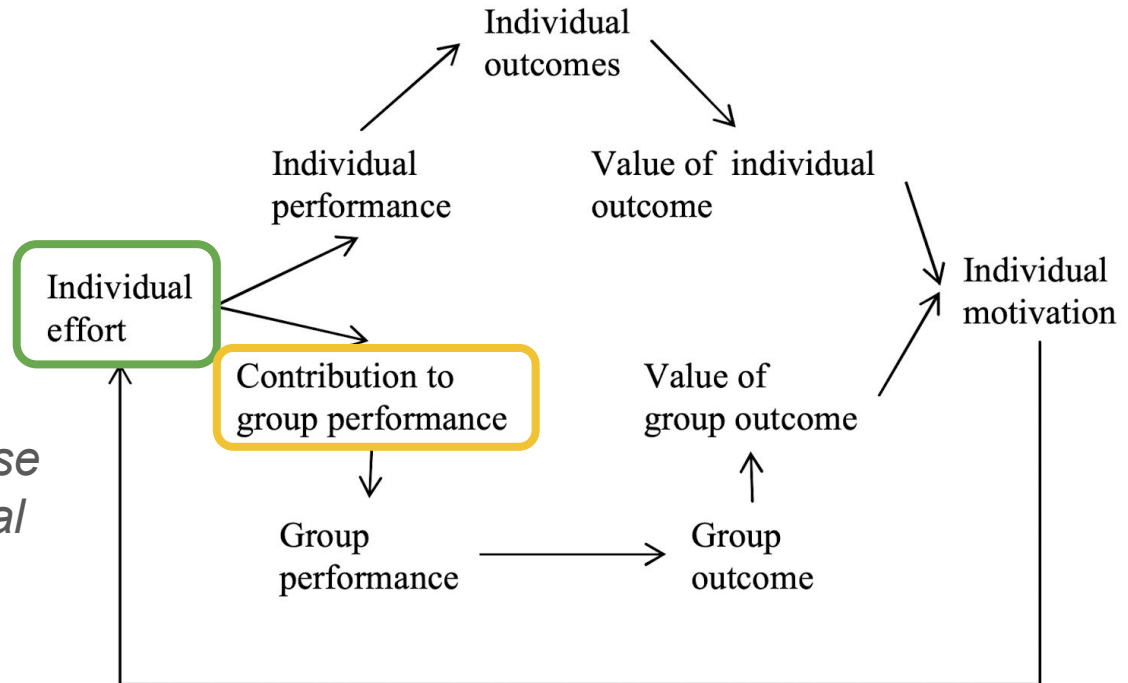
Study 1: Salience of uniqueness

Hypothesis 1:

Users will rate more movies when the uniqueness of their contributions is made salient

“We are contacting you because as someone with fairly unusual tastes, you have been an especially valuable user of MovieLens.”

Figure 1: The collective effort model (adapted from [13])

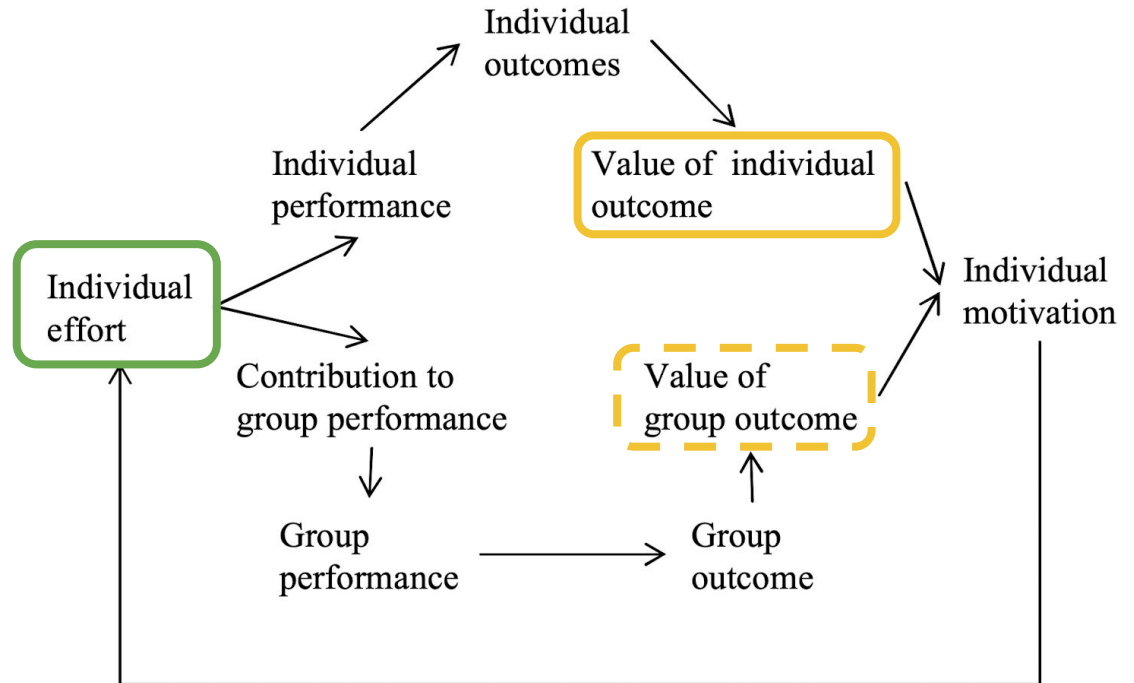


Study 1: Salience of benefit and the beneficiary

Figure 1: The collective effort model (adapted from [13])

Hypotheses 2 a & b:

Users will rate more movies when the personal or community benefit they receive from doing so is made salient

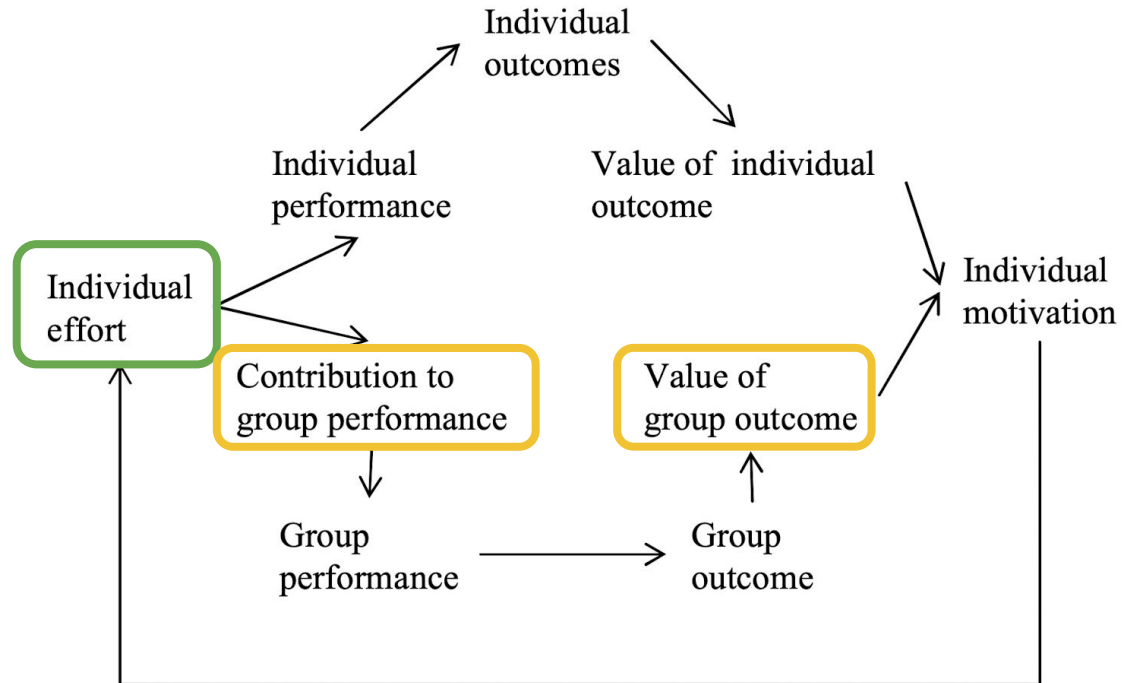


Study 1: Interaction between uniqueness and benefit

Hypothesis 3:

Users will rate more when the perception of both unique contribution and benefits to the community are made salient than when only unique contribution or benefits are made salient

Figure 1: The collective effort model (adapted from [13])



Study 1: Results

Discussion questions:

Other than uniqueness, what elements of users' contributions might motivate further contributions? Why?

Reminding users to contribute was generally effective. Are there limits to this approach?

Hypothesis	Results	
	All movies	Rarely-rated movies
H1: Unique group rates more than non-unique group	Supported	Supported
H2a: Benefit-to-self group rates more than no-benefit groups	Disconfirmed (significant in the opposite direction)	Not supported
H2b: Benefit-to-others group rates more than no-benefit group	Disconfirmed (significant in the opposite direction)	Not supported
H3: Interaction between Uniqueness and Benefit	Not supported	Not supported

Study 2: Motivating through goal setting

“Providing people with specific, high-challenge goals stimulates higher task performance than easy or ‘do your best’ goals”

Why?

- Belief in ability to perform task is motivating
- Normative information on expected performance
- Achieving the goal is satisfying

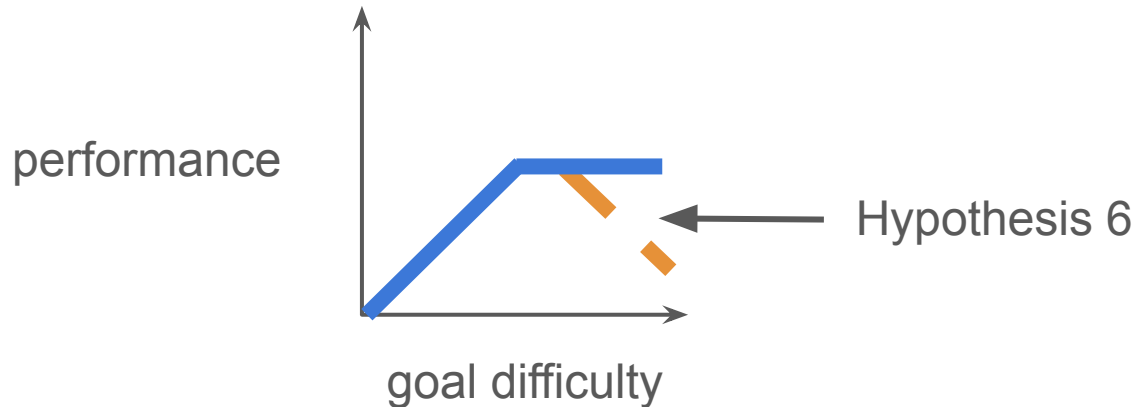
Hypothesis 4: Specific, numeric goals will motivate greater contributions than non-specific goals

Study 2: Group goals and the limits of challenge

Hypothesis 5: Members assigned individual goals will provide more contributions than members assigned group goals

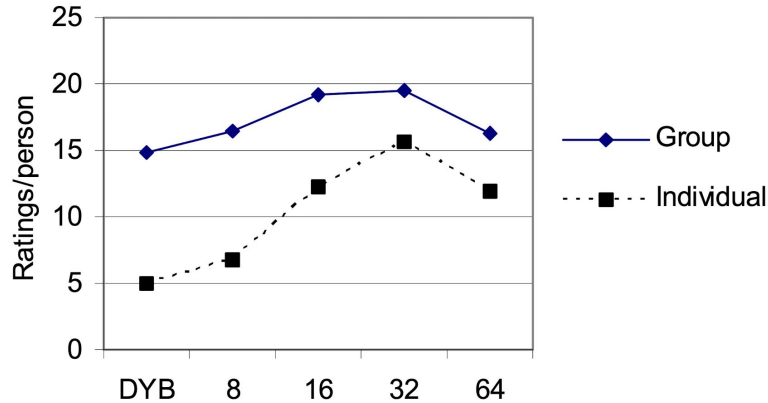
Hypothesis 6: In an online community, contribution will drop off when goals exceed some difficulty threshold

Locke and Latham Model



Study 2: Results

Figure 4: Effects of goals on number of ratings



Hypothesis	Result
H4: Specific goals more effective than non-specific goals	Supported
H5: Individual goals more effective than group goals	Disconfirmed (reverse of hypothesis at $p < .10$ level)
H6: Difficulty of goal has convex effect on contributions	Weak support (only for participants in the individual goals condition who rated)

Why do you think group goals outperformed individual goals, despite expected social loafing?

How would you integrate specific high-challenge goals into a platform you use to improve contributions? What are the potential downsides of this approach?

Why principals may fail in practice

Failures of implementation

- Theory correct, but implementation did not capture principles properly

Mismatches between science and engineering

- In engineering and HCI implementation, primary goal is problem-solving using many factors
- In behavioral science, goal is uniquely determining the causes for social phenomena
- e.g. What was the effect of calling groups “explorers”

Incomplete theories

The way forward

- Theories may fail in practice, but mining social psychology is a useful general strategy
- Recommends motivating contributions to groups using theories related to:
 - Group cohesion
 - Group identity
 - Interpersonal attraction
 - Altruism

Discussion Question: What are pros and cons of these potential options?

Closing Discussion

What elements of online platforms today, either by design or by the nature of their composition, utilize social psychology to motivate behavior? What else might you try?

What are the potential downsides of the use of social psychology to motivate online activity?

How might these results or their interpretation have changed with technological changes since 2004?



Closing Discussion

What are the potential effects of the use of social psychology to motivate online activity? What examples support your answer?

Theory used in this work:

- people work less hard to achieve some goal when they think they are working jointly with others except for when they
 - (a) believe their effort is important to the group's performance,
 - (b) believe that their contributions to the group are identifiable, and
 - they (c) like the group they are working with

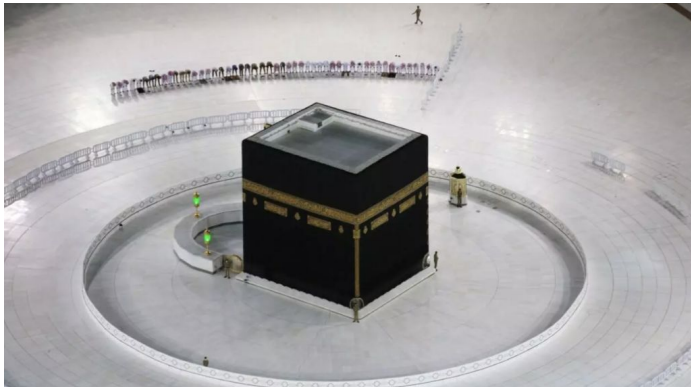
(Re)Capturing the Spirit of Ramadan: Techno-Religious Practices in the Time of COVID-19

Nadia Caidi, Cansu Ekmekcioglu, Rojin Jamali

COVID-19 Pandemic and Ramadan

Ramadan, as a recurring global religious month, is practiced by a diverse Muslim community and has well-defined codes and rituals.

- Fasting in Ramadan: One of the five pillars of Islam where physically able Muslim refrains from drinking or eating from sunrise to sunset.
- Major disruptions in the social, cultural, and economic lives.
- Major events around the world have been cancelled



Definitions

Technology appropriation: The ways in which a technology is not just adopted by people, but also integrated into everyday social, economic, or cultural practices.

Non-Use: people might choose to not use a technology despite it being readily available and accessible

Techno-Spirituality: Techno-spirituality examines how technology is integrated with religious and spiritual practices.

Three key themes in techno-spiritual HCI:

- Institutional
- Practical
- Experiential

Discussion question (1/5)



3:00

Can you think of an example of a design used to support or reshape existing religious practices or enhance individual spiritual and religious experiences?

Research Question

RQ: How were the religious practices around Ramadan mediated and reshaped by technologies during the pandemic?

Data collection: Interview

Focused on four data categories:

- **Demographics:** Participants' age, gender, and backgrounds.
- **Ramadan Memories and Rituals:** Emotional and ritualistic dimensions.
- **Pandemic Coping Strategies:** Everyday life adjustments and decision-making.
- **Technology Use:** Pre- and post-pandemic technology habits.

Findings (1/5): Digital Information Practices during the pandemic

Demands:

- Adjusting daily routines
- Nutrition.
- Cooking and sourcing Ramadan-specific foods
- Vaccine updates
- Fasting guidelines during the pandemic.
- Health risks for fasting, especially for older adults or those with preexisting conditions.
- Social distancing and mosque closures.

Limitations:

- Overwhelming information online. WHO declared COVID-19 as infodemic.

Discussion question (2/5)



3:00

What strategies can individuals and communities adopt to manage information overload and misinformation in the context of religious and public health practices?

Findings (2-3/5): The Digital Reimagining of Ramadan and Coping strategies during Ramadan

Participants described Ramadan's meaning in sensory terms.

Technology use:

- Recording of the call for prayer.
- Ramadan playlists Programs.
- Recipes online or images of childhood dishes
- Track prayer times and fasting cutoffs

Common Coping Strategies Across Both Ramadans:

- Social media to stay in touch with family.
- Focusing less on pandemic news and more on religious practices.

Limitations:

- No linguistically relevant content.
- Absence of localized content.



Discussion question (3/5)



3:00

How do sensory experiences, such as sounds, smells, and tastes, contribute to the significance of religious rituals?

How can digital tools replicate or enhance these sensory aspects to maintain the essence of the rituals?

How does isolation during events like the pandemic reshape the communal and spiritual dimensions of religious rituals?

What strategies can be used to maintain a sense of connection in such circumstances?

Findings (4/5): Ramadan Beyond the Individual

- Sense of togetherness :
 - WhatsApp/Facebook Groups
- Global Solidarity :
 - Black Lives Matter movement and discussions about mental health.

Limitations

- Data and privacy concerns.
- Appropriate concerns



- Local Muslims Rally in Support of Black Lives Matter -
CAIR-Ohio

**3:00**

Discussion question (4/5)

Some participants noted that these platforms are not inclusive or sometimes feature insensitive content. Consider methods researchers could employ to address these issues and make the platforms more inclusive.

Findings (5/5): On limited or non-use of technologies

Reason why participants limited their technology use

- Lack of Access
- Intentional Avoidance for Mental Well-being
- Shifting Digital Habits
- Dissatisfaction with Digital Platforms



Discussion question

Social media platforms can have an impact on people's mental well-being. What role should designers play in maintaining this balance, and how can they create platforms that effectively support it?

What do you think are some of the lasting impacts of COVID-19's digital transformation on faith in a post-pandemic world?

Discussion question



3:00

What do you think are some of the lasting impacts of COVID-19's digital transformation on religious practices in a post-pandemic world?

Conclusion and takeaways

- Digital technologies played a crucial role during the pandemic to sustain religious rituals.
- Digital tools could not fully replicate the embodied and communal experiences of in-person rituals.
- Platforms should cater to diverse communities while addressing cultural, linguistic, and socio-religious needs.
- Religious rituals are deeply tied to sensory experiences (sight, sound, smell, touch, taste), which are challenging to recreate digitally.

Thank you

	Week 7			Week 8			Week 9			Week 10		
						Finish Analysis			Finish Analysis Seriously	Develop Slides & Script		Edit Paper Before Submission
Assign Writing Tasks												
Intro												
Related Work												
Methodology												
Results												
Discussion												
Abstract/Conclusion												

Introduction (~1-page)

What is the problem? (establish societal need, context)

2. What is our solution? What did we do specifically? (1 paragraph together, add a little more relevant detail)

3. What are top 1-2 findings to take away? (biggest impact again but also include a summary of more detailed findings, try to include a figure too)

4. How will this help the field of HCI? (describe contributions in detail, bulleted list, what would be next steps)

Related Work

- What gap is this research filling?
- What research should you summarize for the authors so they can understand this gap?
- No more than 4 paragraphs (for this class)
- Try to do a bit more than simply summarize. Tell a story with the prior work.
- For this class, there is no background section in the paper, so feel free to add that in there as well

Related Work Outline (Example)

- Paper focused on end user use of llms for Security+ exam

P0: Might spend a few sentences explaining the exam and how it impacts job future (should really be in intro but might not be depending on RQs)

P1: Prior work on how people study and study strategies

P2: Prior work on how people interact with AI agents and the impact on learning

P3: Prior work on user beliefs about llms (trust, perceptions, etc.)

P0cont: A few sentences expressing the gap here we are filling

Related Work Outline (Example)

Bottom line - what do people need to know in order to full get the point you are making in the paper?

That information should be given in the introduction, background , and related work.

Keep in mind that for this class I do not want a background section. If you add it I will not read it

Methodology

- What methods did you use to do your work? (cite them)
- Participant recruitment
- Did you have IRB approval (not needed for this class)
- Ethics (what efforts did you make to keep things ethical?)
- Participants demographics

Results

What did you find?

Discussion

Explain your findings

Results & Discussion

Sometime the results section is used to present findings and then explain them.

In this case, the discussion section is used to explore deeper questions or conclusions.

Example -

Results & Discussion

Example - Paper focused on end user use of llms for Security+ exam. RQ: How does user trust in llms impact user behavior in high stakes situation?

Results - results from survey, usage data, Security+ outcome

- Discuss how one results impact another

Discussion

- What do these results mean for the future of AI as a resource?
- Confidence and LLM usage