Human-AI Interaction, Lecture 1

What Is Human-AI Interaction?

Three image models

A glass of red wine filled all the way to the brim



FLUX.1 schnell

Three image models

A glass of red wine filled all the way to the brim



FLUX.1 schnell



Nano Banana (Google)

Three image models

A glass of red wine filled all the way to the brim



FLUX.1 schnell



Nano Banana (Google)



Seedream 4 (ByteDance)

Rewording the prompt

A wine glass filled completely to the very top of the rim with red wine, so full it is about to spill over





The glass is still not full.

FLUX.1 schnell

Windows Central's report

ChatGPT's new image-generation upgrade is stunning, and OpenAI has seemingly loosened its grip on its capabilities. A Reddit user recently shared an interesting discovery, indicating that the tool can now generate a glass of wine filled to the brim.

Windows Central, March 26, 2025

Three questions from the wine glass

Understand: why can't a model draw it?

Steer: how do you get the glass you want?

Delegate: which model would you give the job to?

Alex O'Connor, “Why Can't ChatGPT Draw a Full Glass of Wine?”, February 18, 2025

How do we design AI that people can **understand**, **steer**, and **delegate** to?

Today

1. Why AI is different
2. Norman's two gulfs
3. Agency plus automation
4. Discussion
5. Map a tool you use
6. How the term runs

Three runs of the same request

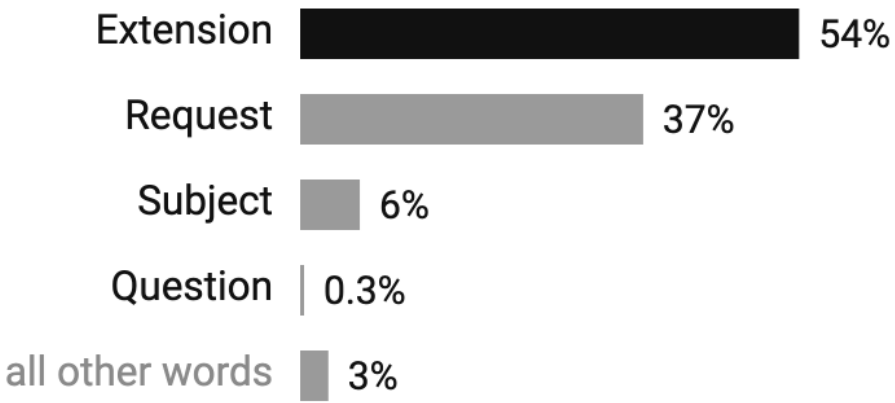
“Write a subject line for an email asking my professor for an extension”

Run 1	Extension Request: [Assignment Name] – [Your Full Name]	3 options
Run 2	**Request for Extension**: [Assignment Name] – [Your Full Name]	4 options
Run 3	Extension Request: **[Course Name]** – [Assignment Name]	4 options

Qwen3.8-27B

Next-word probabilities

...**Option 1: Clear & Professional (Recommended)**
>



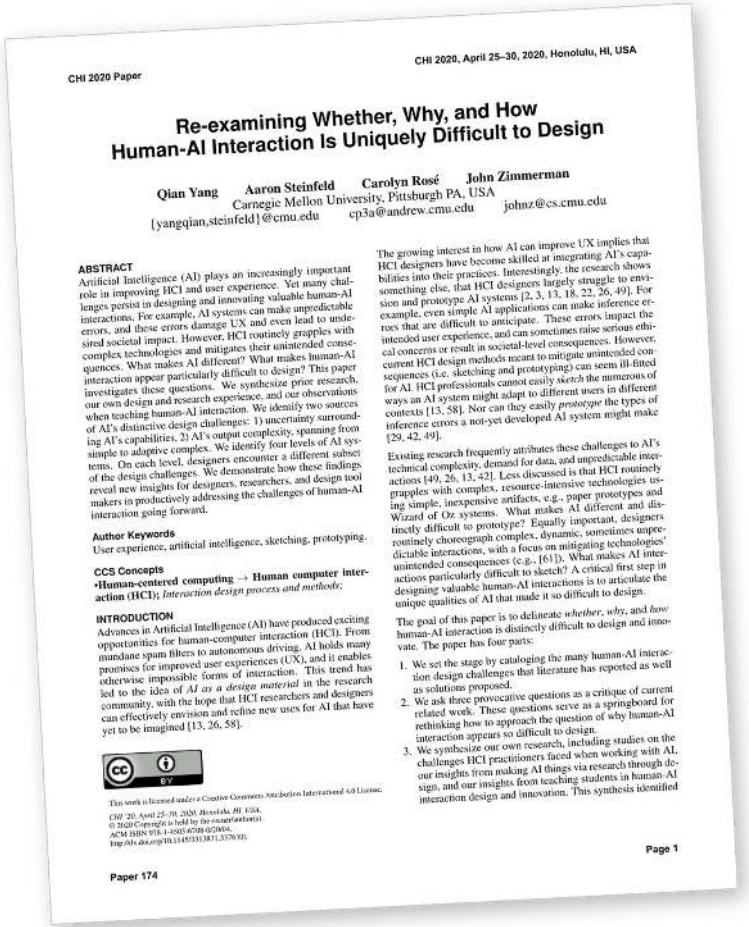
Run 1 picked “Extension.” Run 2 started with “Request.”

Qwen3.8-27B

How sampling works: Jurafsky & Martin, Speech and Language Processing, 3rd ed. draft, Sec. 7.6

“there exists no catalog of available AI capabilities”

Yang et al., CHI 2020, p. 6



CHI 2020 Paper

CHI 2020, April 25–30, 2020, Honolulu, HI, USA

Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design

Qian Yang Aaron Steinfeld Carolyn Rosé John Zimmerman
Carnegie Mellon University, Pittsburgh PA, USA
{yangqian,steinfeld}@cmu.edu cp3a@andrew.cmu.edu johnz@cs.cmu.edu

ABSTRACT
Artificial Intelligence (AI) plays an increasingly important role in improving HCI and user experience. Yet many challenges persist in designing and innovating valuable human-AI interactions. For example, AI systems can make unpredictable errors, and these errors damage UX and even lead to undesired societal impact. However, HCI routinely grapples with complex technologies and mitigates their unintended consequences. What makes AI different? What makes human-AI interaction appear particularly difficult to design? This paper investigates these questions. We synthesize prior research, our own design and research experience, and our observations when teaching human-AI interaction. We identify two sources of AI's distinctive design challenges: 1) uncertainty surrounding AI's capabilities, 2) AI's output complexity, spanning from simple to adaptive complex. We identify four levels of AI systems. On each level, designers encounter a different subset of the design challenges. We demonstrate how these findings reveal new insights for designers, researchers, and design tool makers in productively addressing the challenges of human-AI interaction going forward.

Author Keywords
User experience, artificial intelligence, sketching, prototyping.

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI);** *Interaction design process and methods;*

INTRODUCTION
Advances in Artificial Intelligence (AI) have produced exciting opportunities for human-computer interaction (HCI). From mundane spam filters to autonomous driving, AI holds many promises for improved user experiences (UX), and it enables otherwise impossible forms of interaction. This trend has led to the idea of *AI as a design material* in the research community, with the hope that HCI researchers and designers can effectively envision and refine new uses for AI that have yet to be imagined [13, 26, 58].

The growing interest in how AI can improve UX implies that HCI designers have become skilled at integrating AI's capabilities into their practices. Interestingly, the research shows something else, that HCI designers largely struggle to envision and prototype AI systems [2, 3, 13, 18, 22, 26, 49]. For example, even simple AI applications can make inference errors that are difficult to anticipate. These errors impact the intended user experience, and can sometimes raise serious ethical concerns or result in societal-level consequences. However, current HCI design methods meant to mitigate unintended consequences (i.e. sketching and prototyping) can seem ill-fitted for AI. HCI professionals cannot easily *sketch* the numerous of ways an AI system might adapt to different users in different contexts [13, 58]. Nor can they easily *prototype* the types of inference errors a not-yet developed AI system might make [29, 42, 49].

Existing research frequently attributes these challenges to AI's technical complexity, demand for data, and unpredictable interactions [49, 26, 13, 42]. Less discussed is that HCI routinely grapples with complex, resource-intensive technologies using simple, inexpensive artifacts, e.g., paper prototypes and Wizard of Oz systems. What makes AI different and distinctly difficult to prototype? Equally important, designers routinely choreograph complex, dynamic, sometimes unpredictable interactions, with a focus on mitigating technologies' unintended consequences (e.g., [61]). What makes AI interactions particularly difficult to sketch? A critical first step in designing valuable human-AI interactions is to articulate the unique qualities of AI that made it so difficult to design.

The goal of this paper is to delineate *whether*, *why*, and *how* human-AI interaction is distinctly difficult to design and innovate. The paper has four parts:

1. We set the stage by cataloging the many human-AI interaction design challenges that literature has reported as well as solutions proposed.
2. We ask three provocative questions as a critique of current related work. These questions serve as a springboard for rethinking how to approach the question of why human-AI interaction appears so difficult to design.
3. We synthesize our own research, including studies on the challenges HCI practitioners faced when working with AI, our insights from making AI things via research through design, and our insights from teaching students in human-AI interaction design and innovation. This synthesis identified

This work is licensed under a Creative Commons Attribution International 4.0 License.

CHI '20, April 25–30, 2020, Honolulu, HI, USA.
© 2020 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-6708-0/20/04.
http://dx.doi.org/10.1145/3313831.3376301

Norman's two gulfs

Today's reading: Don Norman, *The Design of Everyday Things* (2013), Chapter 2



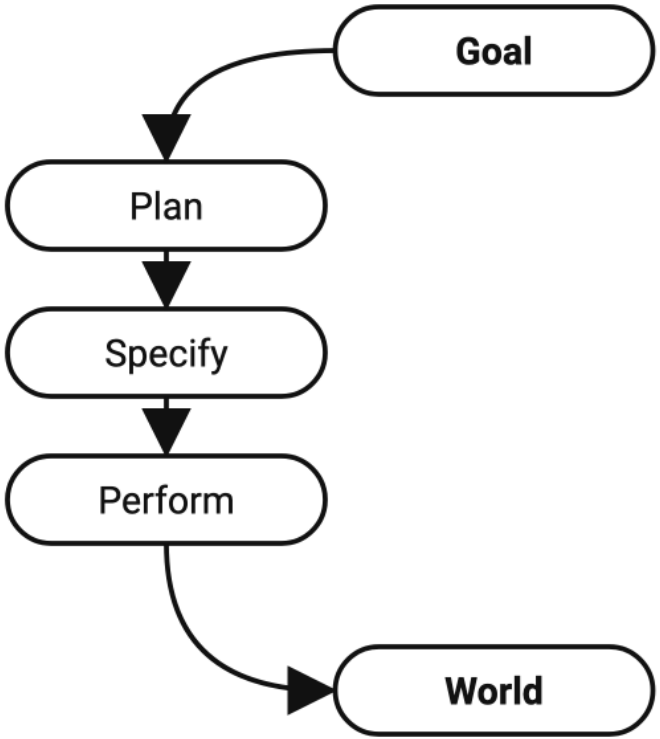
Don Norman in 2005
Photo: jordanfischer, CC BY 2.0
via Wikimedia Commons

Norman's seven stages of action



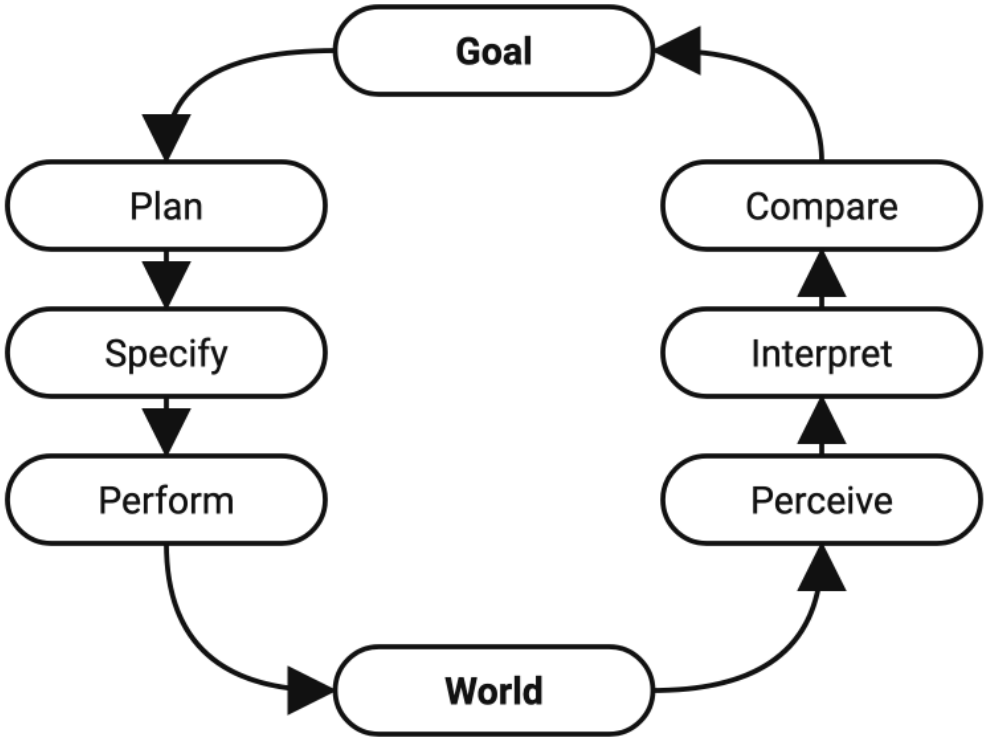
Seven stages of action: Norman, The Design of Everyday Things, 2013, p. 41

Three stages of execution



Seven stages of action: Norman, The Design of Everyday Things, 2013, p. 41

Three stages of evaluation

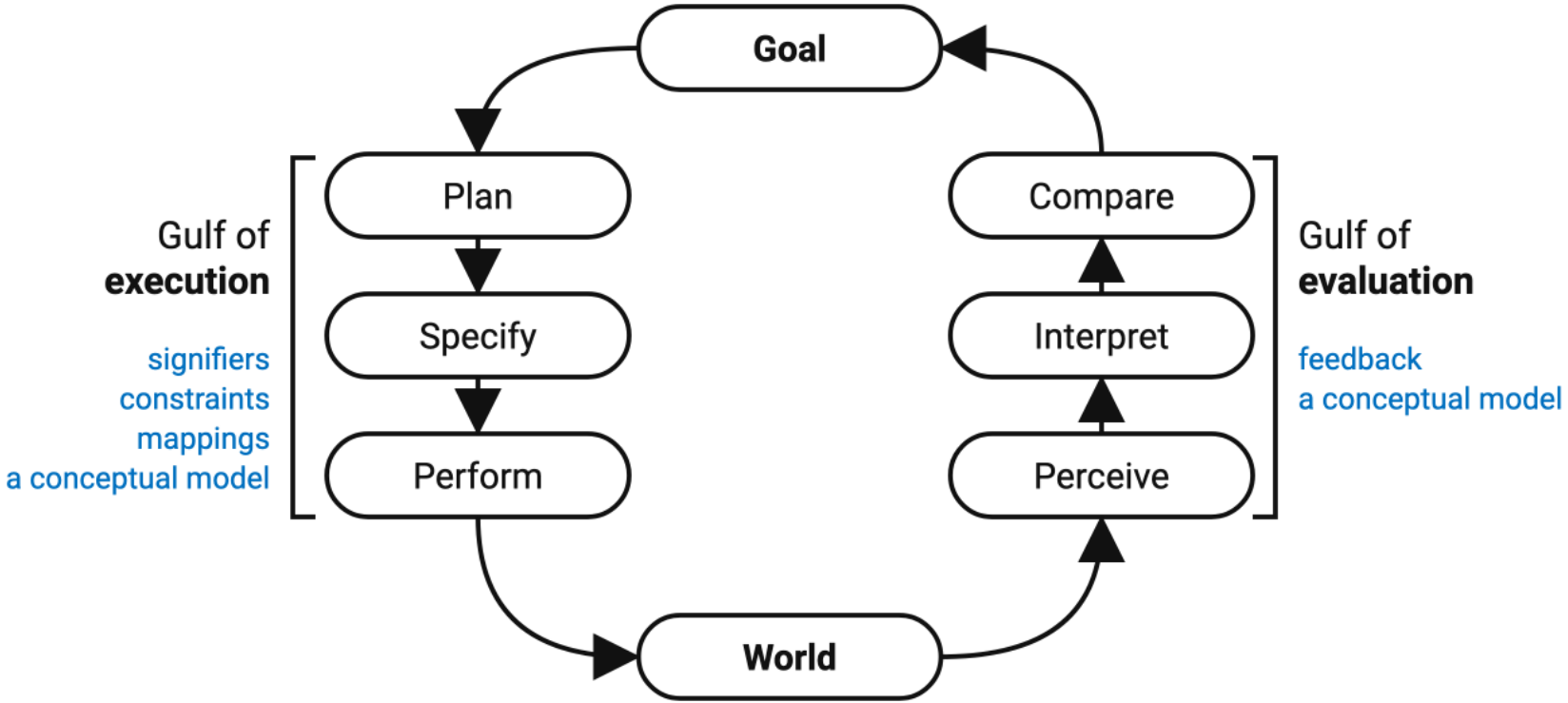


Seven stages of action: Norman, The Design of Everyday Things, 2013, p. 41

Gulf of **Execution**: “where they try to figure out how it operates”
Gulf of **Evaluation**: “where they try to figure out what happened”

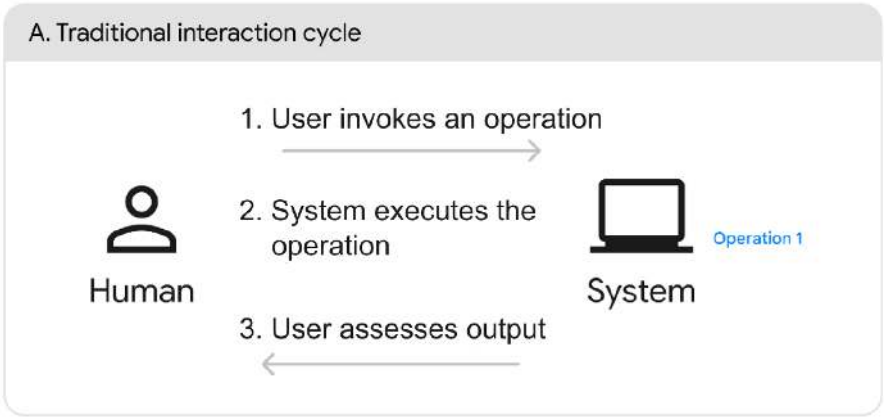
Norman, The Design of Everyday Things, 2013, p. 38

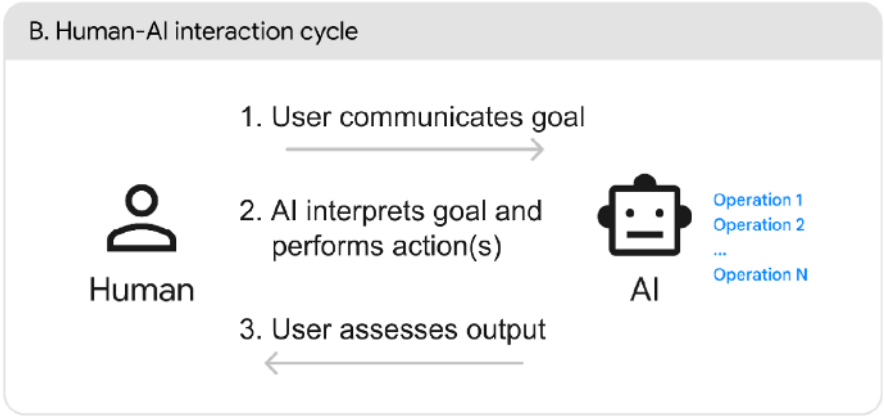
Bridging the gulfs



Bridges for each gulf: Norman, The Design of Everyday Things, 2013, p. 40

Describing the outcome



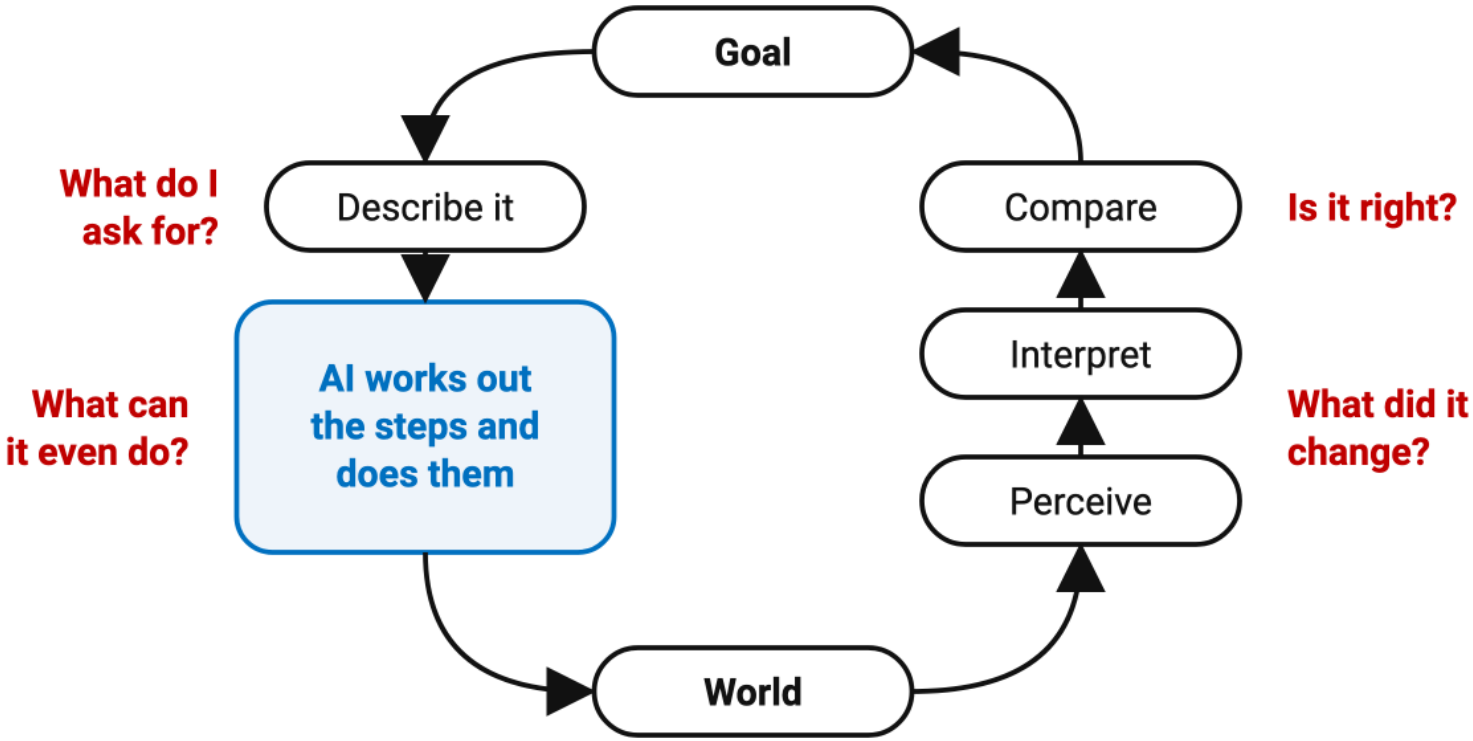


You pick the operation

You describe the result

Terry et al., arXiv 2023, Fig. 1 (panels A and B)

Norman's two gulfs with AI



Terry et al., arXiv 2023; Subramonyam et al., CHI 2024

Gulf of execution: the wine glass



FLUX.1 schnell

The prompt says
exactly what you want.

Rewording it did not get a full glass.

Making an email more polite

Make this email more polite:

Hi Dana, send me the slides by Friday at 5 pm. The client meeting moved to Monday at 10 am in Room 204. Thanks, Alex

Here are a few ways to rewrite the email to be more polite and professional, depending on your relationship with Dana:

Option 1: Professional & Courteous (Recommended)

> Hi Dana,
> Could you please send over the slides by 5:00 PM this Friday?
> Also, please note that the client meeting has been moved to Monday at 10:00 AM in Room 204.
> Best regards, Alex

Option 2: Slightly Softer & Collaborative

> Hi Dana,
> Hope you're having a good week.
> When you have a moment, could you please share the slides with me by Friday at 5:00 PM? I'd like to make sure we're ready for the client.
> Additionally, the client meeting has been rescheduled to Monday at 10:00 AM in Room 204.
> Thanks, Alex

Making an email more polite

Make this email more polite:

Hi Dana, send me the slides by Friday at 5 pm. The client meeting moved to Monday at 10 am in Room 204. Thanks, Alex

The times and the room are right.

Option 2 added a reason Alex never gave.

Here are a few ways to rewrite the email to be more polite and professional, depending on your relationship with Dana:

Option 1: Professional & Courteous (Recommended)

> Hi Dana,
> Could you please send over the slides by 5:00 PM this Friday?
> Also, please note that the client meeting has been moved to Monday at 10:00 AM in Room 204.
> Best regards, Alex

Option 2: Slightly Softer & Collaborative

> Hi Dana,
> Hope you're having a good week.
> When you have a moment, could you please share the slides with me by Friday at 5:00 PM? I'd like to make sure we're ready for the client.
> Additionally, the client meeting has been rescheduled to Monday at 10:00 AM in Room 204.
> Thanks, Alex

Qwen3.8-27B

Agency plus automation

Today's paper: Heer, *PNAS* 2019



PNAS

Agency plus automation: Designing artificial intelligence into interactive systems

Jeffrey Heer[a,1]

[a]Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, WA 98195-2350

Edited by Ben Shneiderman, University of Maryland, College Park, MD, and accepted by Editorial Board Member Eva Tardos October 15, 2018 (received for review May 22, 2018)

Much contemporary rhetoric regards the prospects and pitfalls of using artificial intelligence techniques to automate an increasing range of tasks, especially those once considered the purview of people alone. These accounts are often wildly optimistic, understating outstanding challenges while turning a blind eye to the human labor that undergirds and sustains ostensibly "automated" services. This long-standing focus on purely automated methods unnecessarily cedes a promising design space: one in which computational assistance augments and enriches, rather than replaces, people's intellectual work. This tension between human agency and machine automation poses vital challenges for design and engineering. In this work, we consider the design of systems that enable rich, adaptive interaction between people and algorithms. We seek to balance the often-complementary strengths and weaknesses of each, while promoting human control and skillful action. We share case studies of interactive systems we have developed in three arenas—data wrangling, exploratory analysis, and natural language translation—that integrate proactive computational support into interactive systems. To improve outcomes and support learning by both people and machines, we describe the use of shared representations of tasks augmented with predictive models of human capabilities and actions. We conclude with a discussion of future prospects and scientific frontiers for intelligence augmentation research.

visualization | data science | artificial intelligence | human–computer interaction | automation

Although sharing overlapping origins in midcentury computer science, research programs in intelligence augmentation (IA: using computers to extend people's ability to process information and reason about complex problems) and artificial intelligence (AI: developing computational methods for perception, reasoning, and action) have to date charted largely separate trajectories. The former, beginning with the work of early luminaries such as Douglas Engelbart and Alan Kay—and continuing in the modern field of human–computer interaction (HCI)—led to now-ubiquitous technologies such as the personal computer, graphical user interfaces, the web browser, and touch-driven mobile devices. Meanwhile, AI has similarly contributed to numerous innovations, from spam filters, fraud detection, and medical diagnostics to recent developments in autonomous vehicles. Advances in the AI subfield of machine learning (ML), including the resurgence of neural network approaches trained on increasingly large datasets, has led a frenzy of both research and industrial activity.

However, current rhetoric around AI often exhibits a myopic focus on fully-automated methods, for example, as a replacement for human labor. Certainly, people have their lapses. We do not "scale" to large repetitive tasks, we may commit errors in judgment due to insufficient information or cognitive biases, and we may have ambiguous intents that AI systems might help refine. Nevertheless, a sole focus on automation is concerning for multiple reasons. Automated approaches may mislead due to inappropriate assumptions or biased training data and may optimize fixed objectives that do not adapt in the face of changing circumstances or adversarial manipulation. In addition, users of such systems may come to overly rely on computational suggestions, leading to a potential loss of critical engagement and domain expertise. Real-world examples of such pitfalls range from flawed disease predictions (1) to preventable airplane crashes resulting from failed coordination among automated systems and the flight crew (2, 3). As these examples illustrate, safeguarding human agency is not only a value to uphold, but in many cases a necessity for effective and appropriate use.

This state of affairs has prompted prominent ML researchers to call for more "human-centered" approaches to AI. For example, Li (4) writes that a major goal of AI should be "enhancing us, not replacing us," while Jordan (5) notes that we "need well-thought-out interactions of humans and computers to solve our most pressing problems." AI experts are thus "rediscovering" HCI and recognizing the need to integrate AI and IA perspectives.

Although both exciting and overdue, we should note that the goal of integrating agency and automation is hardly new. In the early 1960s, Bar-Hillel (6) wrote (in the context of natural language translation): "The decisive problem becomes to determine the region of optimality in the continuum of possible divisions of labor" between people and computers. In the 1990s, HCI and AI researchers engaged in debate over direct manipulation vs. interface agents (7), arriving at a consensus in support of "increased automation that amplifies the productivity of users and gives them increased capabilities in carrying out their tasks, while preserving their sense of control and their responsibility." Meanwhile, research on creativity support tools (8) has explored nonanthropomorphic modalities for interacting with automated methods, including simultaneous interaction with multiple design alternatives (9) and integration of manual input and procedural programming (10).

Still, the need to productively couple agency and automation poses vital challenges for design and engineering. How might we effectively integrate automated reasoning into user-centered interactive systems? We examine one promising strategy: the design of shared representations of possible actions, enabling computational reasoning about people's tasks alongside interfaces with which people can review, select, revise, or dismiss algorithmic suggestions.

Rather than jumping right to the goal of imitating "human-level intelligence," we can take inspiration from a more humble

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Creativity and Collaboration: Resisting Cybernetic Serendipity," held March 13–14, 2018, at the National Academy of Sciences in Washington, DC. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/Cybernetic_Serendipity.

Author contributions: J.H. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

Conflict of interest statement: This article includes discussion of two academic research projects that were subsequently commercialized as start-up companies: Trifacta Inc. and Lyft Inc. J.H. has financial interests in both of these companies.

This article is a PNAS Direct Submission. B.S. is a guest editor invited by the Editorial Board.

Published under the PNAS license.

¹Email: jheer@uw.edu.

Published online February 4, 2019.

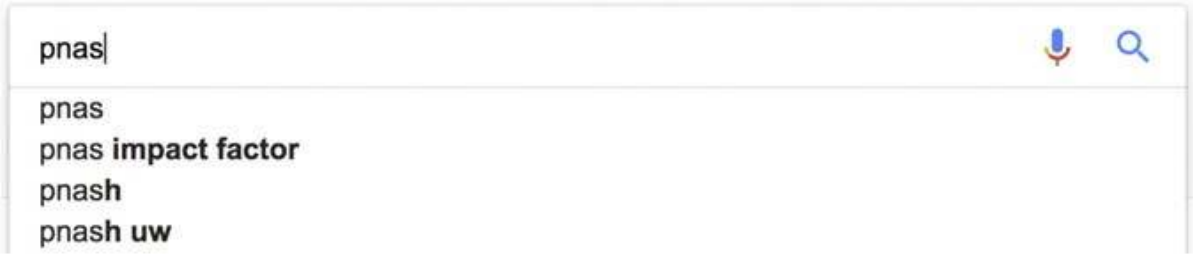
www.pnas.org/cgi/doi/10.1073/pnas.1807184115

Heer: a design space “in which computational assistance **augments and enriches, rather than replaces**, people's intellectual work”

Heer, PNAS 2019, p. 1844

Spell check and autocomplete

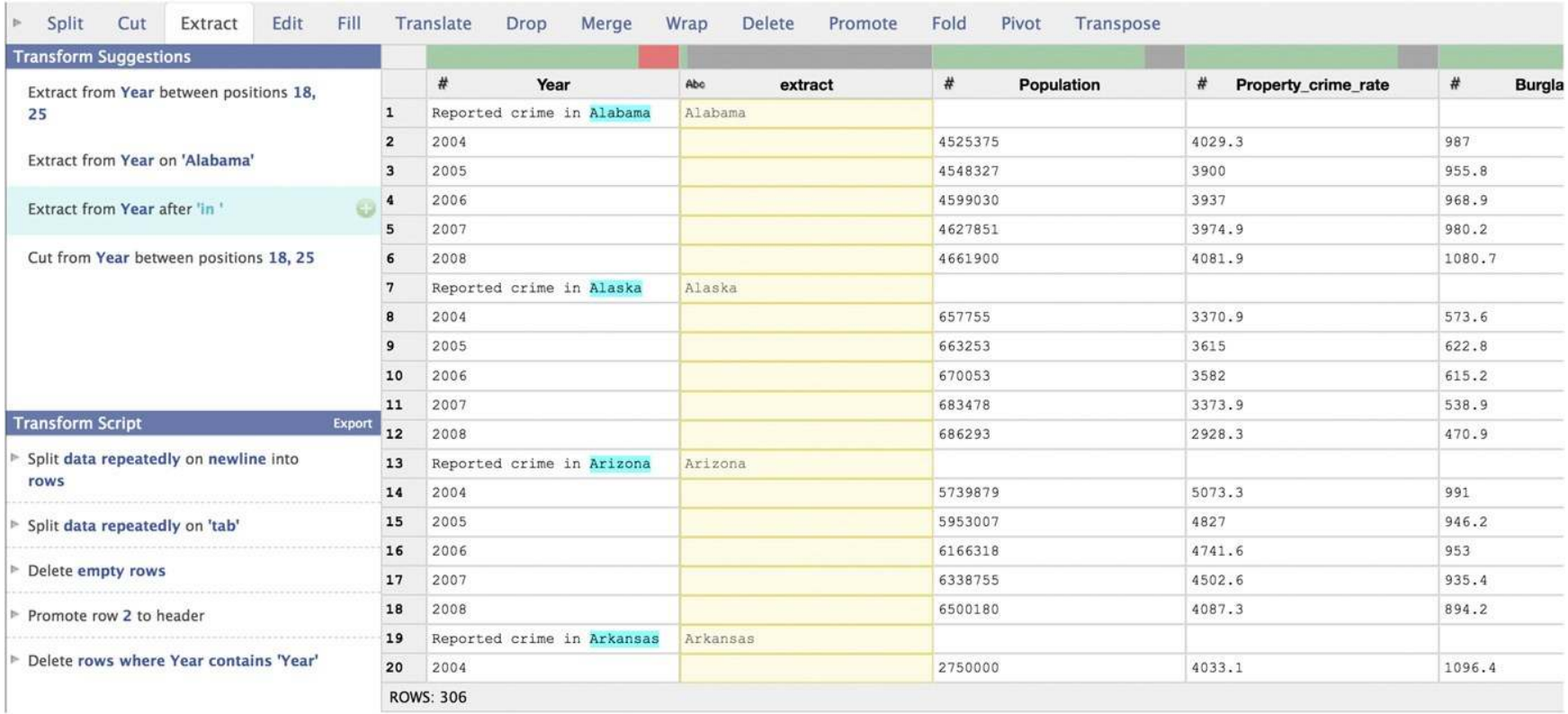
Human beings face ever more complex and urgent problems, and there effectiveness in dealing with these problems is a matter that is critcal to the stability and progress of society. Humans are effective in part because they make use of efficeint tools, methods, and strategies.



You can ignore both, and they are useful even when they are wrong.

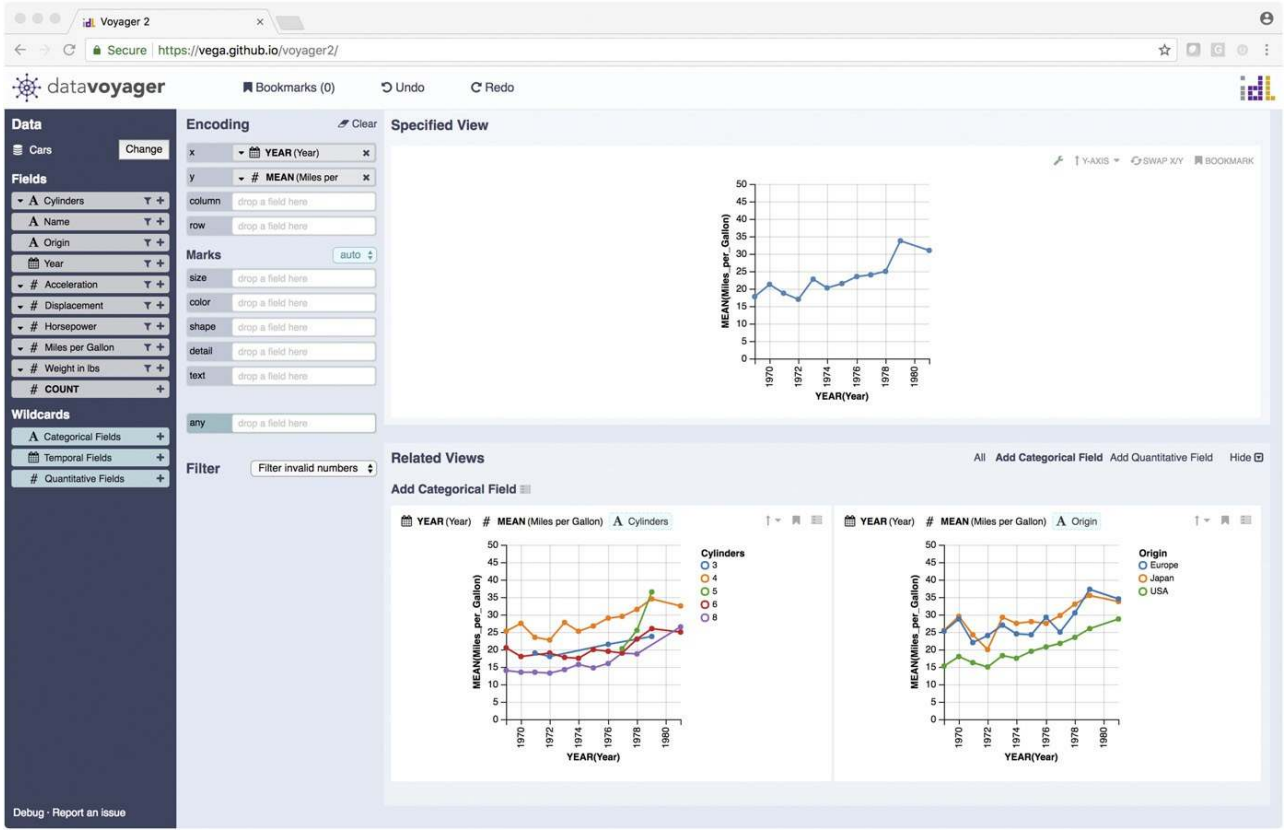
Heer, PNAS 2019, Figs. 1 and 2

Wrangler



Heer, PNAS 2019, Fig. 3

Voyager



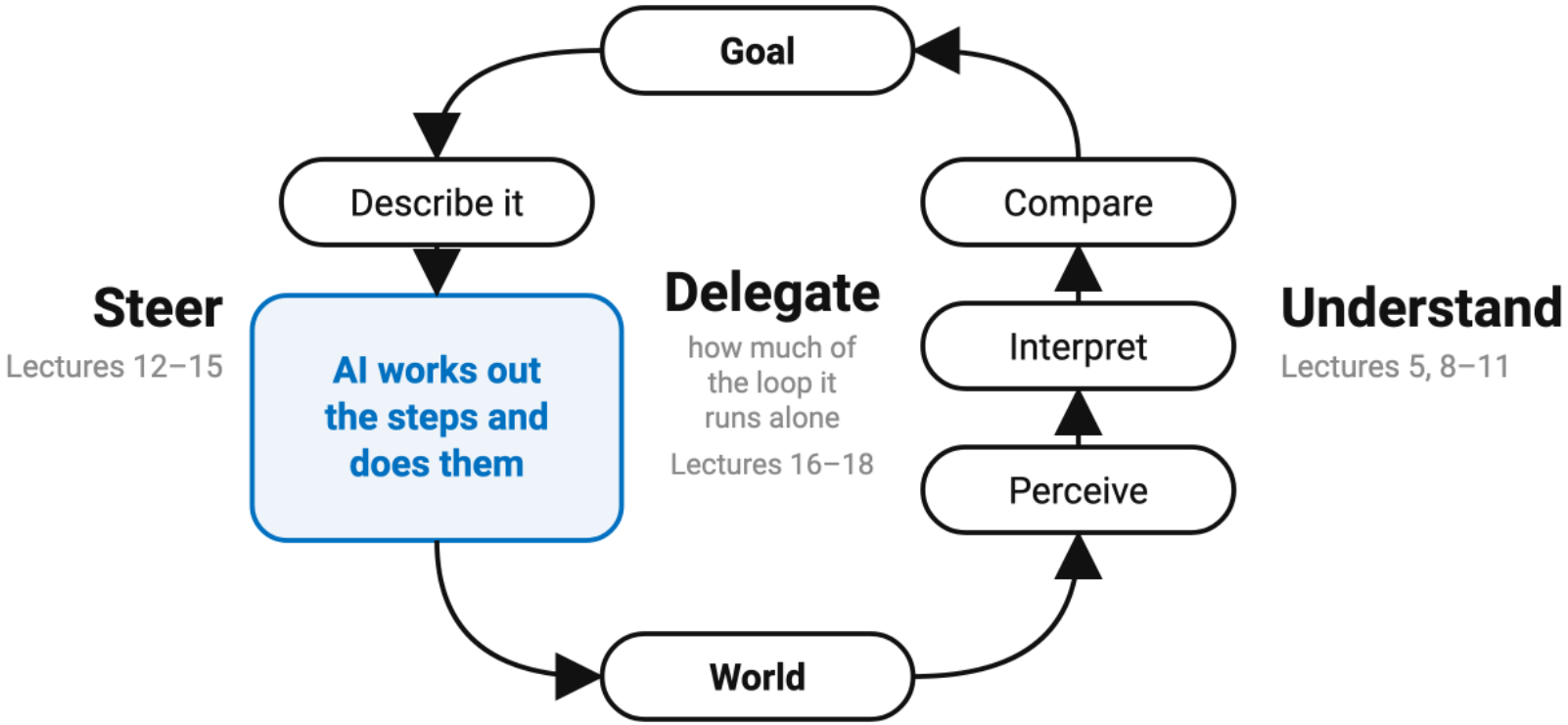
Top: your chart.
Below: charts
it suggests.

Heer, PNAS 2019, Fig. 4

“These related views are so good but it's also spoiling that I **start thinking less**.”

A Voyager study participant, quoted in Heer, PNAS 2019

The three parts of the course



Discussion: Norman and Heer

1. Which of Norman's bridges is hardest to build for a chatbot?

2. Which tasks would you hand to
AI completely, and which never?

3. When is it a problem that AI makes
people's work more alike, and for whom?

Map a tool you use

In pairs, 12 minutes

1. Pick an AI tool you have both used, and one real task
2. Fill in both sides of the map
3. Circle where the gulf was widest

The gulf map

Tool: _____ **Task:** _____

Gulf of execution

1. What did you want?

2. What did you type or do?

3. Where was it hard to get it to do that?

Gulf of evaluation

1. What did it give back?

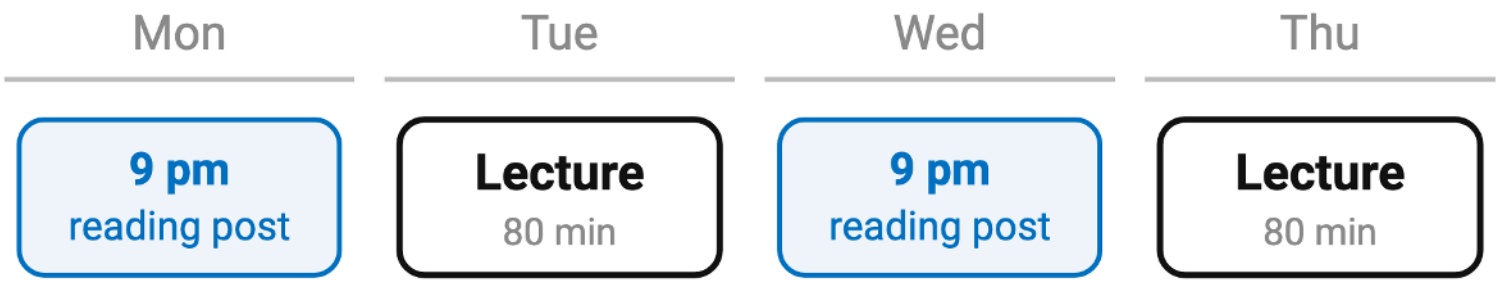
2. How did you check it?

3. Where was it hard to tell what it did?

Circle the place the gulf was widest. **One change** that would make it smaller:

Keep your sheet for Lecture 20.

How the term runs



Reading post: due **9 pm the night before** each lecture

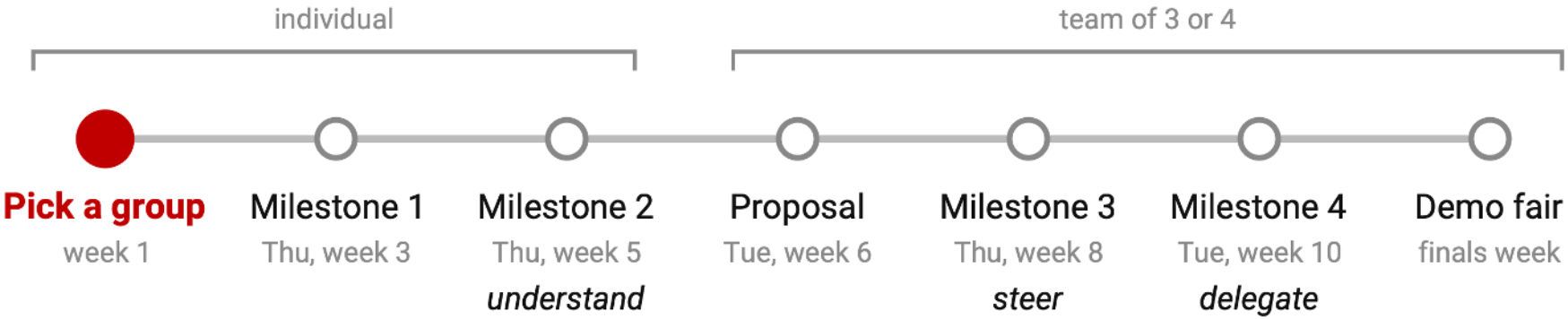
Using AI tools in this course

+ allowed, including for code and prototypes

+ note each use in your milestone report

- not for reading posts or the final paper

Your project: four milestones



Milestone 1: Watch people use AI

Due Thursday of week 3

Next: Lecture 2

History of Human-AI Interaction

Read before class:

- Licklider, “Man-Computer Symbiosis” (1960)
- Shneiderman & Maes, “Direct manipulation vs. interface agents” (1997)