



Creating Human-Computer Partnerships

Wendy E. Mackay

► To cite this version:

Wendy E. Mackay. Creating Human-Computer Partnerships. CHIRA 2023 - 7th International Computer-Human Interaction Research and Applications, Nov 2023, Rome, Italy. pp.3-17, 10.1007/978-3-031-49425-3_1 . hal-04440357

HAL Id: hal-04440357

<https://hal.science/hal-04440357>

Submitted on 6 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Creating Human-Computer Partnerships

Wendy E. Mackay^(✉)

Université Paris-Saclay, CNRS, Inria Saclay, Gif-sur-Yvette, France
wendy.mackay@inria.fr

Abstract. How can we design “human-computer partnerships” that take optimal advantage of human skills and system capabilities? Artificial Intelligence research is usually measured in terms of the effectiveness of an algorithm, whereas Human-Computer Interaction research focuses on enhancing human abilities. I argue that better AI algorithms are neither necessary nor sufficient for creating more effective intelligent systems. Instead, we need to focus on the details of interaction and how to successfully balance the simplicity of the user’s interaction with the expressive power of the system. After describing our approach to “generative theories of interaction”, I illustrate how to create interactive intelligent systems where users can discover relevant functionality, express individual differences and appropriate the system for their own personal use.

Keywords: Generative theories of interaction · Human-centered artificial intelligence · Human-computer partnerships · Participatory design · Upskilling

1 Introduction

My research in the field of Human-Computer Interaction (HCI) focuses on *human-centered AI*, i.e. how to create intelligent interactive systems with a measurably positive impact on human users. We seek to design “*human-computer partnerships*” that combine human skills and system capabilities so that humans and intelligent agents perform better together than either individually. This requires not only taking advantage of the best characteristics of each, but also considering how to improve human capabilities over time, thus “upskilling” rather than deskilling or replacing the users of these systems [24].

Unfortunately, most current AI research is measured in terms of improving the algorithm, not its effect on human users. Systems such as ChatGPT are rightly hailed as major technical breakthroughs, but carry many embedded assumptions about how to interact with them. Asking people to converse with a simulated person appeals to our very human tendency to anthropomorphize the objects and animals we interact with. However, this can lead to severe problems when the intelligent agent is wrong or “hallucinates” [12].

1.1 The Cost of System-Induced Errors

Consider the recent fiasco surrounding Meta’s *Galactica* whose language-based AI model was trained on a large body of scientific literature to support writing scientific papers. Unfortunately, *Galactica* was prone to generating dangerous pseudo-science. The journalist Tristan Green got it to generate “*well-written research papers on the benefits of committing suicide, practicing antisemitism, and eating crushed glass.*” [10]. *Galactica* was shut down after only three days, not because of the algorithm per se, but rather because the designers had not sufficiently considered how humans might interact with it. Legitimate science takes expertise, discipline and time, but *Galactica* allowed non-scientists to produce fake science with a highly authoritative voice. Although designed to save time, it was equally ripe for misuse.

Of course, not all intelligent systems exhibit such dramatic failures. We are all confronted daily with more a more innocuous version of this problem: the smartphone’s “auto-correct” feature. Here, the designer’s assumption is that the user made an error — which may or may not be true — and that the correction is actually correct, which also may or may not be true. The designer may also assume that any errors that are produced by the system are innocuous and that the benefits of using the system outweigh the cost of introducing new errors.

Another more subtle effect is the transformation of the user’s role from “author” to “error corrector”. Although a touch typist can simultaneously type and read text on a laptop screen, phone typing requires continually shifting attention between the soft keyboard, the suggested words and the text output window. Producing error-free text requires constant monitoring to ensure that the system has not introduced errors, a notoriously poor use of human skills [9].

Worse, incorrect words require immediate action from the user. I usually catch errors only after I press “send”, if I notice them at all. I must then decide whether or not to shift the topic of conversation to a discussion of the perils of auto-correct. This may be amusing when chatting with friends and family, but is less acceptable for work-related texts.

Finally, people and intelligent agents produce very different kinds of mistakes. If a person misspells a word, other people are very good at inferring the correct one. By contrast, some auto-corrected errors are impossible to guess. For example, my phone always transforms the verb “is” into “OSS”. I cannot inform the system that this is incorrect and must either remember to slow down my typing and explicitly choose “is” from the suggestion window or else waste time explaining my phone’s odd quirk to the message recipient. A good human speller may question the advantage of using “auto-correct” if the act of correcting the corrections wastes more time than simply typing in the first place.

1.2 Shaping the User’s Behavior

Not only do humans *adapt* their behavior to accommodate the system, but the system also *shapes* [25] the user’s actions. For example, auto-correct clearly affects how we type over time. Instead of fully typing each word, we may just type

the beginning of a long word and wait for the system to propose the complete word. Some people take this even further, typing the fewest possible letters and waiting for the system to produce a good-enough suggestion. This can be efficient and save keystrokes but risks producing only pat, stereotypical phrases.

We also might choose to accept a proposed word even if it was not quite what we wanted to say. For example, if I type an informal “congrats”, the system might replace it with the more formal “congratulations” or a highly informal champagne emoji. Now I must not only decide the “right” level of formality but also whether or not to take time correcting it.

Human conversations operate differently from chats with intelligent agents. Conversation partners establish what Clark [7] calls “common ground” and rely on mutual knowledge, beliefs and assumptions to communicate efficiently. Intelligent agents and people have access to different, overlapping sets of data but only humans can share beliefs and assumptions. Humans react to each others’ conversation styles and shape each other’s behavior during the conversation. However, humans also tend to treat intelligent agents as if they were human.

I witnessed an early example of this in the mid 1980s when I worked on Digital Equipment Corporation’s *IVIS*, the first commercial interactive video system. My research group developed an intelligent tutor, “Dwayne”, an animated claymation character who spoke with *DECtalk* as he explained how to fix a printer. We found that field service technicians typed much simpler, requests when Dwayne spoke with a child’s rather than an adult’s voice. The presentation and interaction mode shaped users’ behavior, encouraging them to limit their word choices as if they were talking to a child, which made it much easier for us to interpret them.

1.3 Presentation Details Matter

Sometimes, showing additional information reduces human skills. For example, my studies of air traffic controllers [17] led to a system called *Caméléon* [19] that augmented controller’s interactions with paper flight strips. They greatly appreciated being able to tap on a flight strip to visually locate that aircraft on the RADAR screen. Yet they objected to another feature that calculated and displayed the aircraft’s expected trajectory, since it would interfere with developing and maintaining their mental model of the air traffic.

A related example involves an intelligent system designed to help ship captains plot an acceptable course when entering a harbor¹. The first version presented the AI’s suggestion *before* the captain plotted their own course. The captains rejected this version because they feared it would reduce their expertise while still holding them responsible if it made an error. However, they accepted the same system with a reversed presentation order: the captain first plotted a course and only then saw suggestions that they could choose to adopt. Here, the same algorithm with a different interface produced very different results: the first risked deskilling users whereas the second helped improve their skills.

¹ Personal communication.

Users may have difficulty determining the effectiveness of their interactions with intelligent agents. A recent study of 54 cybersecurity experts at Stanford University found that programmers who used an intelligent code assistant (*Github’s Co-Pilot*) wrote significantly less secure code but believed they wrote better code [23]. Another study of 49 students in a Behavioral Economics class at the HEC Business school showed that students who wrote their own analytical essays performed 28% better than those who corrected an essay written by ChatGPT [11].

1.4 Interacting with AI

Traditional characterizations of intelligent systems treat users as just another component of the system, either as a useful source of data or an unwanted source of error. Rafner et al. [24] distinguish among “*Human-In-The-Loop*” (HITL) systems that plan, execute or evaluate data acquisition tasks [21]; “*Human-On-The-Loop*” (HOTL) systems that require a human to check the results [22]; and “*Human-Out-Of-The-Loop*” (HOOTL) systems that do not include human users at all [26].

Note that each frames the user’s role in terms of the system’s requirements, not the user’s needs. These characterizations also raise important questions: For example, if humans are “out of the loop” does that save them time and free them for more interesting cognitive tasks or will it replace them altogether? If the system takes over the repetitive and “boring” elements of a task, will this enhance the user’s job satisfaction or reduce their ability to learn and enhance their expertise over time? A concert pianist does not become a virtuoso by performing only the most interesting elements of a piece — practicing scales is also necessary for honing motor skills and developing one’s ear. The assumption that intelligent systems will provide users with more interesting cognitive tasks is just that: an assumption. It simply will not happen unless intelligent system designers adopt the goal of enhancing the user’s experience, which requires focusing on the interaction.

Humans and intelligent systems also fail in different ways, each with associated risks. For example, under normal conditions, the French national energy grid’s automated systems balance loads somewhat better than human experts. However, when the COVID-19 pandemic caused radical shifts in energy requirements, the automated systems made strange, costly errors and the human experts performed significantly better.² Consider too what will happen with almost-autonomous vehicles where drivers need not pay attention until the car suddenly faces a crisis that the system cannot handle. Such systems will both deskill drivers and blame them when they lack the expertise necessary to immediately avert disaster.

We also need to consider interaction at multiple levels of scale. Designers of intelligent systems must consider the back-and-forth interaction between the user and the system, not only when everything works well but also when the system makes errors. Does the system limit the user’s power of expression to make their

² Personal communication from staff.

inputs easier to interpret? Does the system transform the user from an author into a system monitor who detects and corrects errors? What are the longer term effects on the user’s behavior? Do users learn and enhance their skills when using the system? Or does removing “easy” tasks and suggesting answers reduce their ability to develop and maintain expertise, effectively deskilling them?

This paper describes our approach to creating human-computer partnerships. I begin with a brief overview of our approach to “generative theory”, which offers a principled approach for designing effective interaction. I then present a series of examples that illustrate how we can significantly improve human-computer interaction with intelligent systems by shifting the focus from the *algorithm* to the *interaction*.

2 Generative Theories of Interaction

I argue that we need to re-frame how we think about artificial intelligence research. Rather than treating human users as mere “input” to a computer algorithm and measuring success in terms of the algorithm’s performance, we must explicitly design intelligent systems so that they increase *human* capabilities.

My research group, ExSitu, explores how to design *human-computer partnerships*, where humans share agency with intelligent, interactive technologies, while explicitly retaining control. Our research is deeply influenced by what we call *Generative Theories of Interaction* [5]. The key insight is that *interaction* can be studied as a phenomenon in its own right [6], with short- and long-term impacts that can be characterized, measured and evaluated. When we apply this approach to the design of intelligent systems, we see that an excellent algorithm can perform poorly if embedded in an inappropriate interface while a simpler, more sustainable algorithm can perform better if embedded in an effective interface.

Generative Theories of interaction (see Fig. 1) begin with concepts drawn from established theories from the natural and social sciences. We then decompose each concept into one or more actionable principles that can be applied to the study and design of interactive systems. We can examine existing interactive systems with an analytical lens, to see if the principles apply and with a critical lens, to see if the system can be improved by applying those principles. Finally, we can explore how to improve existing systems or create new ones by applying a constructive lens that applies the principles in innovative ways.

Much of our work builds on the concept of *co-adaptation* [17] inspired by evolutionary biology, where humans both *adapt to* or learn how to interact with the technology as well as *adapt* or appropriate the system for their own purposes. Darwin’s classic text *The Origin of Species* [8] describes *co-evolution*, where species are both affected by and affect the environments in which they live as well as *co-adaptation*, which describes the interactions among individual organisms affect each other: “*Co-adaptation* emphasizes this on-going, potentially asymmetrical process of mutual influence between organisms and the environment, where *survival of the fittest* is not simply a matter of an organism adapting to a changing environment, but also of it physically changing that environment to

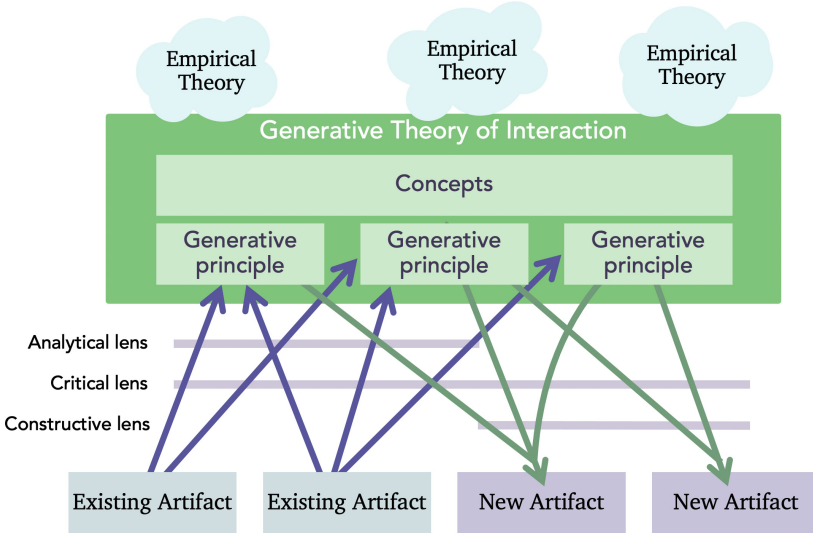


Fig. 1. Generative theories of interaction generate actionable design principles drawn from the natural and social sciences.

ensure its survival.” We can examine the phenomenon of co-adaptation with respect to human users and intelligent agents. If we take the user’s perspective, we see that users establish an ongoing dual relationship with the technology, where they discover how it works but also modify it for new purposes. If we switch to an intelligent system’s perspective, the dual relationship is similar, although not identical, since some systems adapt to the user while others seek to change the user’s behavior [5, 17]. We call this set of inter-related interactions *reciprocal co-adaptation*, where both users and intelligent systems *adapt to* or learn from each other, as well as *adapt* or shape each other’s behavior.

Of course, human users and intelligent agents access, interpret and present information differently. Intelligent systems, especially those based on statistical models such as machine learning and neural networks, are essentially impossible for human users to understand. The examples in the next section illustrate how we have employed generative theories of interaction to address the challenge of building effective interactive intelligent systems.

3 Creating Effective Human-Computer Partnerships

How can we design effective “human-computer partnerships” that make the most of human skills and system capabilities? The first examples explore how to increase a smartphone user’s power of expression while maintaining simplicity of execution [18]. This involves creating intelligent co-adaptive systems that are discoverable — easy to learn and adapt to — as well as expressive and appropriable — easy to personalize and adapt. The next examples illustrate how human

users can successfully share agency with intelligent agents. The final examples explore how we can exploit principles from behavioral psychology [25] where humans and agents shape each others' behavior.

3.1 Discoverability

Even though today's smartphones are extremely powerful, the interaction is extremely limited and takes little advantage of the expressive power of our fingers. Users press buttons, select from menus, fill out forms and enter text with tapping, swiping or pinching gestures. The advantage of this approach is the very low barrier to entry — almost anyone can use the system with little training. The disadvantage is that users have very little power of expression. They can only react to the buttons and controls displayed on the screen, rather than generating commands directly.

What if we could express commands by drawing gestures, for example, drawing a “pigtail” gesture to delete a phrase? The phone is certainly capable of identifying such gestures, so why is this not an option? This is a problem of discovery: users must be aware that such gesture-commands exist as well as recall the mappings between gestures and commands. The design challenge is how to provide in-context help for novice users when they need it, without slowing down experts.

We introduced the concept of “dynamic guides” [4] which combines on-screen feedforward and feedback to help users learn, execute and remember gesture sets. We take advantage of the fact that users who are unsure tend to pause, whereas experts simply draw the gesture. Figure 2 shows what happens when the user starts to draw but hesitates: a visual guide appears around the cursor or their finger, with different paths for each gesture-command combination. The classification model updates dynamically as the user follows the desired template, providing feedback as to what the user has already drawn and feedforward as other alternatives diminish or disappear. Novices can learn these gesture commands over time and smoothly transition into experts. If they forget a gesture, they can always pause to display the guide.

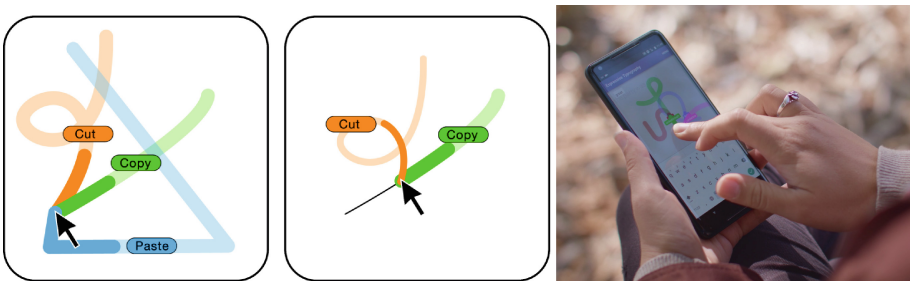


Fig. 2. *Octopocus* displays a dynamic, in-context guide that shows how to draw different command gestures. Continuously updated feedforward and feedback appear as the user draws.

3.2 Expressivity

A fully co-adaptive system does not simply help users adapt to the existing system, but also lets them adapt it to meet their individual needs. For example, human-produced handwriting is highly personal. We can recognize who wrote what and make educated guesses about the conditions under which it was written, such as in a rush or slowly and carefully. By contrast, text messages all look the same.

How can we offer users the expressive power of handwriting, but on a phone? Zhai and Kristensson [27] introduced “gesture typing”, where users glide their fingers between letters, lifting only between words, and showed that it is 40% faster than tapping on a soft keyboard. Today’s smartphones all offer this feature, with extremely proficient algorithms that identify the correct word and discard all extraneous human variation as unwanted “noise”.

By contrast, we were interested in the details of this human variation and how to transform it into expressive output, such as color and font. Figure 3 shows the *Expressive Keyboard* [2], which maps the variation in how users glide between letters into different colors or fonts. For example, we can create the full RGB color space by mapping “curviness” to red, size to “green” and speed variation to “blue”. Users can then consciously control both the content and the color of gesture-typed words. We thus “recycle” otherwise ignored gesture variation to create rich output without sacrificing accuracy. We also created the *CommandBoard* [1], which offers a full range of custom-designed fonts and the *MojiBoard* [3], which transforms gesture-typed words into emojis with a single stroke.

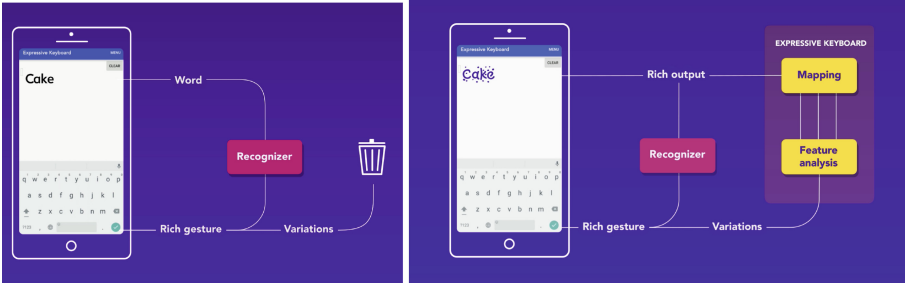


Fig. 3. Standard gesture typing recognizers discard user variation as “noise”. The *Expressive Keyboard* transforms individual variation into rich output, such as color or font style, without sacrificing accuracy.

3.3 Appropriability

In addition to creating expressive output, users also want to personalize their input by choosing their own gestures and mappings to their preferred commands.

Here, the challenge is to balance the needs of the human user, who wants to create gestures that are easy to recall and the system, which needs to reliably recognize and distinguish each gesture within the set of gesture-commands.

We developed the *Fieldward* [20] dynamic guide as an interactive colored heat map that visualizes the “negative space” of unused gestures and verifies whether or not the current gesture is unique (see Fig. 4). The user simply enters a command to define and begins to draw the desired gesture.

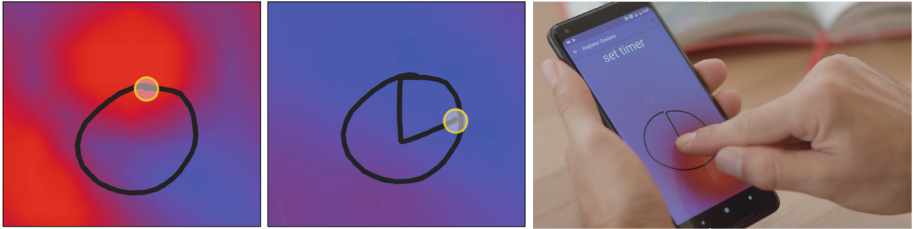


Fig. 4. *Fieldward* displays a dynamic heat map that shows if the user’s preferred gesture will be recognized (gesture stops in a blue zone) or not (gesture stops in a red zone). (Color figure online)

As the user draws, the display changes color: Ending the gesture in a red zone indicates that gesture already exists. Ending it in a purple zone means that the gesture is ambiguous and it is safer to either continue or try another option. Ending it in a blue zone means that the current gesture can be recognized reliably and is suitable as a new gesture-command.

Users enjoyed exploring different gesture possibilities, such as drawing a heart to phone a user’s boyfriend. Another user not only designed individual gesture-commands but also created his own personal gesture syntax. He associated certain gestures with commands, e.g., “call” or “text” and others with people, e.g. “Mom” and “Bob”. He could then create fluid gesture combinations to create compound commands, such as “Call Mom”.

3.4 Sharing Agency

One of the major challenges in creating human-computer partnerships is how to share agency between human users and intelligent agents, broadly defined as exerting control over the outcome. We worked closely with designers to create *Semantic Collage*, a digital mood board creation tool that attaches semantic labels to images [14]. Designers propose images that *Semantic Collage* translates into semantic labels, which designers can then adjust to improve the search. This back-and-forth approach allows designers to share agency with the algorithm, where each proposes possibilities and provides specific feedback to the other.

A subsequent project, *ImageSense* [13] (see Fig. 5) combined individual and shared work spaces as well as collaboration with multiple forms of intelligent

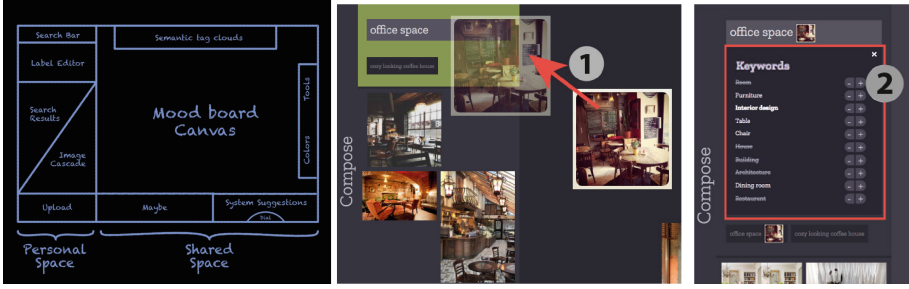


Fig. 5. *ImageSense* lets designers vary their level of agency with an intelligent mood board (left). *Image Cascade* supports full user control, *Semantic Collage* supports shared user-agent control and *Select-o-matic* offers primarily agent control.

agents. *ImageSense* allows users to transition smoothly between different levels of agency, from user-directed selection of curated images (*ImageCascade*), to alternating between user-initiated and system-initiated search (*Semantic Collage*) to responding to system-driven recommendations (*Select-o-matic*). Designers can tweak the agent’s behavior, e.g. by specifying “fewer images like this” and modify the types of generated semantic labels, color palettes or tag clouds and adjust their level of importance.

Intelligent systems such as those described above are intended to facilitate the interaction. But what if we instead ask an intelligent agent to challenge the user in order to gain efficiency?

We created *BIGNav* [15], a multi-scale navigation technique based on Bayesian Experimental Design. The goal is to maximize the information-theoretic concept of mutual information, also known as “information gain” (see Fig. 6.) Rather than simply executing user navigation commands, *BIGNav* interprets the user’s input to update its knowledge about the user’s intended target. It then produces a new view that maximizes the expected information gain i.e. that

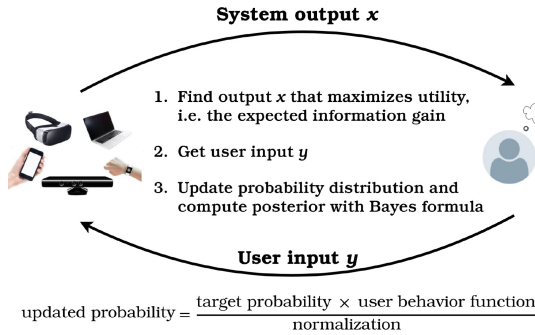


Fig. 6. *BIGNav* presents the user with choices selected to maximize the information gained from each user response and then updates the probability distribution accordingly.

should provide the most information from the user’s next input. *BIGFile* [16] uses a similar approach for file navigation. The split interface combines standard file system navigation with an adaptive area that includes shortcuts calculated with *BIG*. Both *BIGNav* and *BIGFile* were 40% faster than the standard interface, but each increases the user’s cognitive load. This type of human-computer partnership highlights the trade-off that can arise between increasing human efficiency and creating a satisfying user experience. A remaining challenge is how to let users choose an appropriate level of difficulty based on their needs in the moment.

3.5 Shaping Human and Agent Behavior

We also explored how humans and intelligent agents can shape each other’s behavior using full-body interaction. In the context of an art-science collaboration sponsored by *Curiositas* and *La Diagonale* at the Université Paris-Saclay, we collaborated with the *N+1 Theater Group* to produce a wall-sized interactive Christmas display, first for a month in a shopping center and then at an art-science fair (see Fig. 7).

The goal was to create an entertaining interactive experience where users share control with the system. The theater company created a mechanical interactive Santa-making factory that would operate only if a user had successfully engaged with an animated Santa Claus character. Our job was to teach passers-by to perform a specific movement, without explaining the movement or even that they were supposed to learn a movement. We designed the interaction so that the user thought they controlled the animated character but the character



Fig. 7. Users interact with the Santa Claus figure. After first mirroring the user’s movements, Santa starts to act independently. Users are rewarded for successive approximations to the final behavior: If the user performs an appropriate movement, Santa moves up another step. If the user performs the final movement, Santa pulls a level and launches the mechanical “Santa-making factory”.

also “controlled” them. We used the concept of “reinforcement” from behavioral psychology, where the reinforcer (or reward) is Santa’s attention.

When the user first arrives, Santa mirrors her movements, which keeps her engaged, but only for the first 30 s or so. Next, Santa starts to act independently and waits until she starts to copy Santa’s movements according to the behavioral phenomenon of “extinction”. This begins the shaping process: a stair with three steps appears and Santa moves up each step as the user makes successive approximations to the correct movement. If the movement is too far from desired movement, Santa looks bored and taps his foot. If she does not respond appropriately, the stairs disappear and Santa drops through the floor. If she successfully discovers the correct movements through trial and error, Santa climbs to the top step and jumps up to grab a virtual lever. This in turn operates a mechanical lever that launches the machine.

The system was a great success and a number of people returned over successive days to play with Santa. Although it was created primarily as a source of entertainment for the public, it also helped us to better understand how an intelligent agent can both respond to the user and also guide (or control) them.

4 Conclusion

In a 1984 interview, Steve Jobs famously observed that, even though he could not, as a human, outrun many animals, if you gave him a bicycle, he could beat them all. He viewed computers as tools for augmenting human capabilities and argued that computers should act as “bicycles for the mind”. In the intervening decades, computers have touched every aspect of our lives, including commerce, entertainment and education. After multiple highs and lows, recent advances in artificial intelligence offer exciting new possibilities for significantly improving our lives, but remain controversial, with conflicting views as to their risks and benefits.

I argue that, in order to truly benefit from AI, we need to reconsider how we conduct AI research and build intelligent systems. We need to shift from measuring the efficiency of the *algorithm* to measuring the effectiveness of the *interaction* over time. We need to create successful “human-computer-partnerships” that augment rather than deskill or replace human users such that the combination of human and AI capabilities together are better than either alone.

We must view human users, not as a source of errors or as cogs in the machine, but rather as the system’s *raison d’être*. Designers of intelligent interactive systems should ask themselves whether or not their users are forced to:

- Monitor the system over long periods of time?
- Remove interesting tasks and replace them with boring, repetitive ones?
- Jump into action to solve a crisis that the system cannot solve, after passively letting the system do everything else?
- Feed data into the system in the system’s terms rather than the user’s?
- Generate yes/no answers to complex questions?
- Behave in prescribed ways to achieve a particular goal?

Clearly, AI researchers must continue developing new algorithms. However, we also need extensive research into how those algorithms are embedded into intelligent interactive systems. We need to carefully design the way information is presented and how users control their interaction with these systems. We must recognize that the same algorithm embedded in a different interface may affect users in extremely different ways. At the same time, we can choose among different algorithms to produce a particular interaction style and develop more “sustainable” AI that achieves the same effect with less data or computing power. Focusing on developing effective human-computer partnerships instead of more powerful algorithms will improve not only how we approach AI research but also AI policy, ethics and law.

Acknowledgments. My thanks to Michel Beaudouin-Lafon, Janin Koch and members of the ExSitu research lab for numerous productive discussions about these ideas. This work was supported by the European Union Horizon 2020 research and innovation program under grant agreement No 952091 (*ALMA: Human-Centric Algebraic Machine Learning*) and grant agreement No 952026 (*HumanE-AI Network*), as well as the *ERC European Research Council* grant No 321135 (*CREATIV: Creating Co-Adaptive Human-Computer Partnerships*).

References

1. Alvina, J., Griggio, C.F., Bi, X., Mackay, W.E.: Commandboard: creating a general-purpose command gesture input space for soft keyboard. In: Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST 2017), pp. 17–28. Association for Computing Machinery, New York (2017). <https://doi.org/10.1145/3126594.3126639>
2. Alvina, J., Malloch, J., Mackay, W.: Expressive keyboards: enriching gesture-typing on mobile devices. In: Proceedings of the 29th ACM Symposium on User Interface Software and Technology (UIST 2016), pp. 583–593. ACM, Tokyo (2016). <https://doi.org/10.1145/2984511.2984560>
3. Alvina, J., Qu, C., McGrenere, J., Mackay, W.: Mojiboard: generating parametric emojis with gesture keyboards. In: CHI 2019 Extended Abstracts (CHI EA 2019), Glasgow (2019). <https://doi.org/10.1145/3290607.3312771>
4. Bau, O., Mackay, W.E.: Octopocus: a dynamic guide for learning gesture-based command sets. In: Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST 2008), pp. 37–46. Association for Computing Machinery, New York (2008). <https://doi.org/10.1145/1449715.1449724>
5. Beaudouin-Lafon, M., Bødker, S., Mackay, W.E.: Generative theories of interaction. *ACM Trans. Comput. Hum. Interact.* **28**(6), 1–54 (2021)
6. Beaudouin-Lafon, M., Mackay, W.E.: Rethinking interaction: from instrumental interaction to human-computer partnerships. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–5 (2018)
7. Clark, H.H.: *Using Language*. Cambridge University Press, Cambridge (1996)
8. Darwin, C. (ed.): *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London (1859)

9. Fitts, P.M., et al.: Human engineering for an effective air navigation and traffic-control system. Tech. rep., Defense Technical Information Center (1951)
10. Green, T.: Meta takes new AI system offline because twitter users are mean. In: TNW (2022). <https://thenextweb.com/news/meta-takes-new-ai-system-offline-because-twitter-users-mean>
11. Hill, B.: Taking the help or going alone: Chatgpt and class assignments. In: HEC Research Paper (2023). <https://doi.org/10.2139/ssrn.4465833>
12. Klein, N.: AI machines aren't 'hallucinating' but their makers are. *The Guardian* (2023)
13. Koch, J., Taffin, N., Beaudouin-Lafon, M., Laine, M., Lucero, A., Mackay, W.E.: Imagesense: an intelligent collaborative ideation tool to support diverse human-computer partnerships. *Proc. ACM Hum. Comput. Interact.* **4**(CSCW1), 1–27 (2020)
14. Koch, J., Taffin, N., Lucero, A., Mackay, W.E.: Semanticcollage: enriching digital mood board design with semantic labels. In: *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, pp. 407–418 (2020)
15. Liu, W., D'Oliveira, R.L., Beaudouin-Lafon, M., Rioul, O.: Bignav: Bayesian information gain for guiding multiscale navigation. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI 2017)*, pp. 5869–5880. Association for Computing Machinery, New York (2017). <https://doi.org/10.1145/3025453.3025524>
16. Liu, W., Rioul, O., McGrenere, J., Mackay, W.E., Beaudouin-Lafon, M.: Bigfile: Bayesian information gain for fast file retrieval. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018)*, pp. 1–13. Association for Computing Machinery, New York (2018). <https://doi.org/10.1145/3173574.3173959>
17. Mackay, W.: Responding to cognitive overload?: co-adaptation between users and technology. *Intellectica* **30**(1), 177–193 (2000). <https://doi.org/10.3406/intel.2000.1597>
18. Mackay, W.E.: Réimaginer nos interactions avec le monde numérique. *Leçons inaugurales du Collège de France, Collège de France — Fayard* (2022)
19. Mackay, W.E., Fayard, A.L., Frobert, L., Médini, L.: Reinventing the familiar: exploring an augmented reality design space for air traffic control. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1998)*, pp. 558–565. ACM Press/Addison-Wesley Publishing Co. (1998). <https://doi.org/10.1145/274644.274719>
20. Malloch, J., Griggio, C.F., McGrenere, J., Mackay, W.E.: Fieldward and pathward: dynamic guides for defining your own gestures. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI 2017)*, pp. 4266–4277. Association for Computing Machinery, New York (2017). <https://doi.org/10.1145/3025453.3025764>
21. Monarch, R.M.: *Human-in-the-Loop Machine Learning: Active Learning and Annotation for Human-Centered AI*. Manning, New York (2021)
22. Nahavandi, S.: Trusted autonomy between humans and robots: Toward human-on-the-loop in robotics and autonomous systems. *IEEE Syst. Man. Cybernet. Magaz.* **3**(1), 10–17 (2017). <https://doi.org/10.1109/MSMC.2016.2623867>
23. Perry, N., Srivastava, M., Kumar, D., Boneh, D.: Do users write more insecure code with AI assistants? (2022)
24. Rafner, J., et al.: Deskillling, upskilling, and reskillling: a case for hybrid intelligence. *Moral. Mach.* **1**, 24–39 (2021). <https://doi.org/10.5771/2747-5174-2021-2-24>

25. Reynolds, G.S.: A Primer of Operant Conditioning. Scott Foresman, Glenview (1968)
26. Steelberg, C.: The path to an AI-connected government. Tech. rep., Data Center Dynamics (2019). <https://www.datacenterdynamics.com/en/opinions/path-ai-connected-government>
27. Zhai, S., Kristensson, P.O.: The word-gesture keyboard: reimagining keyboard interaction. *Commun. ACM* **55**(9), 91–101 (2012). <https://doi.org/10.1145/2330667.2330689>