Pilot Attitudes Toward AI in the Cockpit: Implications for Design

Zelun Tony Zhang
fortiss GmbH
Research Institute of the
Free State of Bavaria
Munich, Germany
zhang@fortiss.org

Yuanting Liu
fortiss GmbH
Research Institute of the
Free State of Bavaria
Munich, Germany
liu@fortiss.org

Heinrich Hußmann
Chair of Applied Informatics
and Media Informatics
LMU Munich
Munich, Germany
hussmann@ifi.lmu.de

***Abstract*—As the aviation industry is actively working on adopting AI for air traffic, stakeholders agree on the need for a human-centered approach. However, automation design is often driven by user-centered intentions, while the development is actually technology-centered. This can be attributed to a discrepancy between the system designers' perspective and complexities in real-world use. The same can be currently observed with AI applications where most design efforts focus on the interface between humans and AI, while the overall system design is built on preconceived assumptions. To understand potential usability issues of AI-driven cockpit assistant systems from the users' perspective, we conducted interviews with four experienced pilots. While our participants did discuss interface issues, they were much more concerned about how autonomous systems could be a burden if the operational complexity exceeds their capabilities. Besides commonly addressed human-AI interface issues, our results thus point to the need for more consideration of operational complexities on a system-design level.**

***Index Terms*—interviews, thematic analysis, intelligent cockpit assistant systems, human-AI interaction, imperfect AI**

I. Introduction

When microprocessors paved the way for advanced automation beginning with the 1970s, system designers aimed to replace human tasks with machines, without much consideration for human factors. The implicit assumption was that individual tasks in complex systems are independent from each other such that humans can be substituted easily with machines, the so-called *substitution myth* [1]. However, empirical evidence soon revealed the flaw in this assumption, showing that automation introduces unexpected qualitative changes to human roles within the overall system [2]. Driven by the experience of several high-profile automation-related accidents, aviation is a domain where these "automation surprises" [1] are particularly well studied. As a result, the aviation industry started to shift its automation philosophy from the technology-centered approach of automating as much as possible toward human-centered automation in the 1990s [3].

Today, the industry is actively working toward introducing artificial intelligence (AI) to commercial air transport (CAT). In doing so, the frequently cited goal is to maintain a human-centered approach to AI-driven flight automation [4]. However, since the rapid development of new algorithmic capabilities is the driver for AI innovation, AI development is currently predominantly technology-centered. Human-centered approaches to AI as for instance recently proposed by Shneiderman [5] are far from being mature yet. Instead, today's AI applications frequently exhibit the gap between user-centered intentions and actually technology-centered development described by Sarter et al. in the 1990s with respect to automation [1]. As they elaborate, this gap can be attributed to designers' tendency to oversimplify the complexity of real usage [1].

To bridge this gap, participatory design or co-design approaches [6] are important to ensure that designs are driven by users' needs and understanding of operational complexities. As an early contribution to this end, we conducted interviews with four experienced pilots to understand their concerns regarding automation, AI and possible usability issues with intelligent cockpit assistant systems. Based on the interviews, we derived high-level design guidelines for AI-based flight deck systems.

II. Background

In expectation of coming AI applications for aviation, the European Union Aviation Safety Agency (EASA)[1] has recently published a first version of its AI roadmap [4]. It outlines the agency's vision of human-centered AI for aviation, with the trustworthiness of AI at its core. From an operational and human factors perspective, the most relevant technical building block of the roadmap is *explainability*, which EASA considers *"a concept that is resolutely human-centric"* [4]. Establishing guidelines for explainability will be a major part of EASA's efforts around AI certification for aviation [7].

Beyond the aviation domain, *explainable AI* (XAI) is in general increasingly moving into the focus, especially among human-centered AI researchers [8]. In the last few years, human-computer interaction (HCI) researchers have conducted numerous studies to examine the effectiveness of explanations for various purposes, like trust calibration [9], [10], system comprehension [11], [12], or task performance [13], [14].

[1]EASA is responsible for certification, regulation, and standardization for civil aviation in the European Union.

978-1-6654-0170-8/21/$31.00

Taking explainability a step further is the idea of *human-autonomy teaming* (HAT) [15], which aims to enable semi-autonomous systems to work together with humans in teams. While XAI is mostly concerned with the communication from machine to human, HAT frameworks build on bi-directional communication [15], [16]: Humans should not only understand what the automation is doing, but should also be able to communicate their intentions to the machine. This concept is particularly popular in aviation-related automation research [17]–[19]. EASA also sees human-AI collaboration, which corresponds to the HAT vision, as an intermediary step before transitioning to full autonomy in the more distant future [4].

III. Research question

Works on both XAI and HAT are driven by human-centered intentions. However, they tend to focus on the human-AI interface for *preconceived* system designs. What is thereby easily neglected is the users' perspective on the operational environment and how it can inform how the system design should look like in the first place. Hence, we set out to engage those that know the operational complexities on the flight deck best—pilots. We formulate the following research question:

RQ: Based on the operational experience of pilots, which kinds of usability issues can be anticipated with the introduction of AI-driven cockpit assistant systems?

IV. Method

A. Interviews

We conducted semi-structured interviews with four experienced pilots (all male, avg. age: 55 years, avg. flying hours: 9,230 hours). Three are working as test pilots and have a background as fighter pilot, but two of them also regularly test-fly passenger aircraft. The fourth interviewee is a former airline pilot.

The interviews had a length of about 1.5 hours each and were conducted via Webex or phone calls with audio recording. Interviewees received no incentives for their participation. The interview script focused on three areas: the cooperation between crew members, participants' experience with flight deck automation, and their attitudes toward AI in the cockpit.

B. Analysis

We transcribed and coded the interview recordings and performed an inductive thematic analysis following the six phases recommended by Braun and Clark [20]. We chose an inductive analysis since our research question is exploratory in nature: Our explicit goal was to understand the pilots' perspective as opposed to basing our results on preconceived assumptions. We further concentrated on themes at the latent level as we were interested in what the deeper-lying motives behind the pilots' statements might reveal about AI usability issues. After several rounds of refinement, we arranged the data into two main themes, five subthemes, and ten sub-subthemes. The final thematic map is shown in Fig. 1.

V. Results

All interviews revolved around two main themes: For one, participants stressed the importance of the human element in flying an aircraft. Consequently, the second main theme was the cautious view pilots held of AI in the cockpit, despite all of them seeing the potential benefit. In the following, we present these themes and their subthemes in more detail. When referencing themes, we underline their names, and we denote "Pilot x" as Px when presenting quotes from the interviews.

A. Importance of human pilots

Modern aircraft may be highly automated, up to the point that pilots have little active work left to do during a usual flight. Nevertheless, human situation awareness and problem-solving skills are still indispensable to handle complex situations, like changes to the original plan or abnormal events. In handling such complex situations, both the human-human interaction (cooperative problem solving) and the human-machine interaction (handling the automation) are important.

1) Cooperative problem solving: The cooperation between the two pilots in today's aircraft facilitates safety by providing the redundancy of *"two sets of eyes and brains"* (P1). To ensure effective collaboration between crew members, work on the flight deck is governed by well-defined procedures on the one hand, while situation-dependent flexibility is encouraged on the other hand.

Well-defined procedures: Aviation is highly procedural in nature, and the flight deck is no exception. Roles are clearly defined, with the pilot-in-command having the final say and the full responsibility for the flight, while the tasks are clearly distributed among the pilot-flying and the pilot-monitoring. Checklists clearly prescribe who needs to do what in which situation, and pilots employ a *"separate language with fixed phrases and special expressions"* (P1).

Situation-dependent flexibility: While prescribed procedures help to clarify what to do most of the time, no procedure can cover every eventuality in complex domains like aviation. Pilots can therefore deviate from procedures when necessary:

> *"For every abnormality it is actually prescribed how it needs to be done. But this is where the human part comes in. If he now thinks this is such a complex thing, then the captain ultimately has the decision what is the right thing to do, no matter what the regulations say."* (P4)

For instance, pilots can fluidly hand over tasks to each other depending on the situation, and co-pilots are encouraged to speak up and to challenge the captain.

This kind of situation-dependent flexibility is currently an exclusively human capability. All interviewees perceived this human capability as the key differentiator between conventional automation and future AI applications, driving an anthropomorphic conception of AI:

> *"But what the AI does, (...) it's not just to support humans somehow, but you want to try to create a system in such a way that you can gradually, well,*

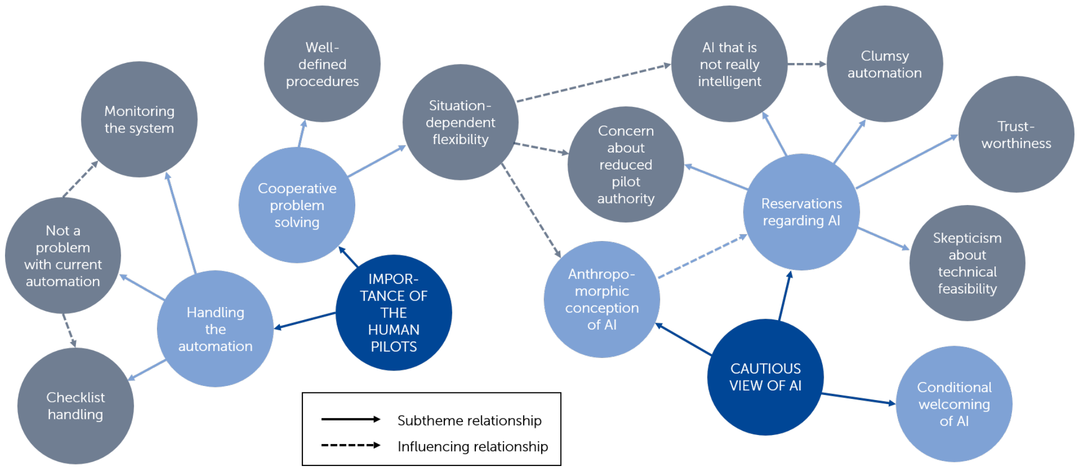

Fig. 1. Thematic map of the pilot interviews. *Dark blue*: main themes. *Light blue*: subthemes. *Gray*: sub-subthemes.

> *not replace humans, but complement them. That you can get the human traits into the system, so to speak."* (P4)

At the same time, this perception appears to be the underlying reason for most of the reservations regarding AI expressed during the interviews. For instance, interviewees discussed at length how AI that is not really intelligent would fall short of tasks that demand situation-dependent flexibility. The participants also expressed their concern about reduced pilot authority due to the introduction of AI, limiting pilots' ability to intervene flexibly.

2) Handling the automation: Between take-off and landing, the autopilot is usually engaged, i.e. for most of the flight, the pilots' task is to handle the automation. For the most part, this means monitoring the system for possible errors. In case something is wrong, the crew tries to address the issue with the help of the corresponding checklist (checklist handling).

Monitoring the system: Through the flight instruments, the pilots monitor the system for failures and other problems. In modern aircraft, the system additionally supports the pilots through alerts, often even displaying the appropriate checklist automatically. Still, human monitoring is required to deal with detection errors, e.g. due to sensor failures. However, the interviewees displayed high confidence in the reliability of the system alerts, suggesting that detection errors are rare:

> *"(...) which is also only psychological, because if any value deviated from normality, there would be a warning anyway."* (P3)

Checklist handling: For the vast majority of system failures or issues, appropriate checklists exist. Hence, pilots can usually simply work through the respective checklist in case of a system alert, seemingly suggesting that these tasks could be easily automated. However, as with system monitoring, the humans on the flight deck are important, as they can recognize the rare instances where the checklist does not properly address the issue at hand and react accordingly:

> *"I just wanted to express that the person behind it can also question the checklist. That it's not a bible. Well, it's the bible for fixing the mistakes, but it's not the panacea, so you're allowed in certain situations, it makes sense to deviate from it as well."* (P4)

Not a problem with current automation: When asked about whether difficulties exist with system monitoring or checklist handling, interviewees reported no issues or even reacted with irritation to the questions (*"Of course! That's what they're trained for"* (P4)). Given the vast body of research on human factors issues like complacency and automation bias [21], these reactions should be interpreted carefully. However, some pilots pointed out a helpful design feature of today's flight automation: While working with the system, pilots can simply focus on deviations of system states from their expectations instead of more demanding forms of information processing. For instance, the flight instruments indicate ranges of normal operation for the respective quantities. While monitoring the system, pilots therefore do not need to interpret every single value, but can simply check whether any value has fallen out of the range of normal operation:

> *"So I don't start reading and interpreting every number and thinking about whether my hydraulic pressure at 270 bar is correct or not, but I look at the display and see that it's in the green range. Whether it's 260 or 280 bar doesn't matter to me at that moment, but I see that it's green."* (P3)

Similarly, checklists allow pilots to form expectations of how the system should react to every single step. Hence, pilots can check the system feedback against their expectation while working through the individual steps and quickly notice when the checklist does not properly address the issue:

> *"You have to check every step you take in order to get feedback. Because every action is expected to have some counter reaction that is checked. Otherwise, you could automate it, simply."* (P4)

Furthermore, participants noted that highly critical incidences are extremely rare and none of the interviewees had personally experienced such situations.

B. Cautious view of AI

The interviewee's statements hinted at an anthropomorphic conception of AI. Against this backdrop, all interviewees were in principal open toward the introduction of AI to the flight deck (conditional welcoming of AI). However, all of them had reservations whether AI could live up to its promise (reservations regarding AI).

1) anthropomorphic conception of AI: Participants talked about AI on the flight deck as a single intelligent agent (as opposed to e.g. a number of separate AI-driven systems) and likened that agent to humans, co-pilots, or artificial teammates (*"The AI somehow must fear for own life, you know?"* (P2); *"the AI human, so to speak, which it could be called then"* (P4)). This conception shaped participants' significant reservations regarding AI: Given their experience as pilots, they focused on the complex requirements that AI would have to meet to behave human-like and to be of value as teammates.

2) Conditional welcoming of AI: Participants stated that if automation works, it is helpful and reduces workload; however, they put emphasis on that qualifier, stressing that automation can also become a burden when the situation surpasses the system capabilities:

> *"If everything goes as planned, flight automation is a dream. The problem starts when things go wrong, or you have to deviate from your plan. Because sometimes, the automation doesn't keep up."* (P3)

Yet, overall, the interviewees acknowledged that automation has increased flight safety and has made work easier for pilots. This differentiated view of current flight automation also translated to the pilots' attitude toward AI: Participants were in principal open toward AI and could all at least vaguely imagine ways it could be helpful to pilots, e.g. as assistance in error diagnosis or by suggesting options for action in complex situations. However, the interviewees all emphasized that AI would only be helpful if it could handle the complexities of aviation and does not become a burden frequently (*"In principle, yes. If it does not become a burden. (...) So this typical thinking along, the experience, so to speak."* (P3)).

3) Reservations regarding AI: The interviewees voiced a range of reservations regarding AI on the flight deck, mostly revolving around potential usability issues.

AI that is not really intelligent: The most prominently voiced reservation was the perception that many of today's AI systems are not really intelligent and therefore not able to exhibit the situation-dependent flexibility that distinguishes humans. Participants regarded this flexibility as mandatory to deal with the complexity of flying an aircraft and to avoid clumsy automation. One pilot gave a vivid account of his experience with a research prototype that did not meet his expectations of an intelligent assistance system:

> *"If I had such a co-pilot, I basically would throw him out my flight deck. The problem was it was just not anticipating and thinking through what might be now the correct action. It was not intelligent so to speak. It was just basically a regulator."* (P3)

Participants mentioned several ways AI falls short of truly intelligent behavior in their view: They missed the ability of AI to communicate like humans (*"Humans could ask: Did you also hear something?"* (P1)); to understand human intentions and to think ahead accordingly (*"I mean each time the pilot is doing something, 'uh, babab, what are you doing now?', this is not AI."* (P2)); or to react properly to complex and changing situations (*"when it gets like 'uuuaah', assessing the overall situation and then executing actions, there I still see room for improvement."* (P3)).

Clumsy automation: Coined by Wiener [22], this term describes automation that reduces workload in situations where it is low anyway, but further increases it in high-workload situations. Participants warned about the potential of AI systems to increase workload when they fail to understand pilot intentions or other context information. For instance, pilots might need to dismiss inappropriate alerts and suggestions or might have to feed information to the system through cumbersome inputs:

> *"The mission you will prepare, with all those alerts to help, those will not fit, maybe, when your mission will change. And then it becomes a pain."* (P2)

Concern about reduced pilot authority: As of today, pilots bear the full responsibility for the safety of the flight. Hence, the interviewees stressed that pilots must always have the final say over the automation and be able to take whatever action they deem necessary to keep the flight safe (situation-dependent flexibility):

> *"I, as the pilot in charge, would like to be involved at least before my aircraft does anything automatically, so that I could intervene at any time."* (P3)

One participant pointed toward the tendency of engineers to overlook this aspect and to suggest system designs that limit pilots' authority:

> *"For example, on some project I was working in (...), people said, 'OK, we'll prevent for example the aircraft to go in a no-fly area, in a forbidden area.' And I said, 'no, the system can alert the pilot that in the front, there is a forbidden area, but never prevent the aircraft to go in, because maybe that's the mission, it's inside the forbidden area.' "* (P2)

Trustworthiness The issue of trustworthiness is core to EASA's AI roadmap [4] and is also often at the center of human-AI interaction research [9], [10]. While trustworthiness was of concern for the interviewees, it appeared to be of secondary importance to them as it was brought up less frequently than the other reservations. Participants discussed trustworthiness form three angles. Most importantly, they expressed the need for the system to explain why it makes certain suggestions, especially when they deviate from the pilot's expectations:

> *"And then the problem is to be confident and to know why my system says 'do this like that and not like that', because me, I was not having the same point of view, let's say. So it's also to understand why, and*

then to say 'oh yes, you are right', or 'no, here, do like that.' (P2)

Another aspect raised was the difficulty of assessing the capabilities of the AI system (*"You can talk to a young co-pilot (...) to set expectations. How to do that with AI? How does it tell how good it is?"* (P1)). Lastly, interviewees felt a lack of public trust in AI, which they saw as an obstacle to AI deployment in aviation (*"I am sure that you could fly passengers from A to B without any pilot, but passengers would not get onto the plane."* (P1)).

Skepticism about technical feasibility: Participants voiced their doubts about whether it is technically feasible for AI systems to understand complex situations well enough to support pilots effectively (*"I have met several people who were working in the AI to make a 'cognitive co-pilot' and so on and so on, and I have never seen anything!"* (P2)). Hence, all interviewees thought that current AI technologies are still far from ready for deployment on flight decks. However, participants were careful not to dismiss the possibility of AI playing a significant role in future cockpits (*"I don't say it will not work, but it will take time."* (P2)).

VI. Discussion

A. Expected usability issues

Current efforts to bring AI to aviation focus on XAI and HAT, the need for which is also reflected in our interviews. Participants discussed how effective communication with the system is a necessity (AI that is not really intelligent) and that they need explanations to trust the system (trustworthiness). However, our analysis reveals that pilots are much more concerned about how AI would handle operational complexities. Permeating the entire data set, this is particularly apparent in the theme situation-dependent flexibility and how it relates to reservations regarding AI. The test pilots' personal accounts of research projects also echo designers' tendency to oversimplify these complexities, as described by Sarter et al. [1].

As a result of oversimplification, designers tend to focus on the benefits an AI application can provide when it works as intended. What is easily missed then are the adverse effects when the operational complexity exceeds system capabilities, which is what the pilots focused on during the interviews. Given that AI is expected to assist pilots in more complex tasks than traditional automation, such issues of clumsy automation are likely even more serious than with today's flight deck systems and warrant much more attention.

Recent work on designing interactions for *imperfect AI* appear promising to address our participants' concerns as it prompts designers to explicitly consider the adverse effects of inadequate system behavior. For instance, researchers have investigated users' tolerance for imperfect AI [23], interfaces allowing users to modify AI outputs [23], [24], or system designs that integrate AI models beneficially despite their imperfection [25]. The latter aspect is of particular interest in our view. Efforts to bring AI to the flight deck are often led by the notion of autonomous agents, which is also reflected in our interviewees' anthropomorphic conception of AI. But judging by the pilots' statements, such agents would need to master the operational complexities of aviation to be truly helpful. At least for the foreseeable future, it appears more promising to search for alternative ways to integrate AI into the overall system design. However, examples like [25] for such designs and human-AI interaction patterns are still rare.

B. Design guidelines

We derive five high-level design guidelines for deploying AI on the flight deck from the interviews. In response to pilots' concerns, the focus is on avoiding adverse effects in case operational complexities exceed the capability of the system. We distinguish between system- and interface-level guidelines.

System-level guidelines:

- **Imperfect AI must not increase workload:** System designers should not only focus on what benefits an AI model could provide, but also prevent that pilots' workload increases when the operational complexity exceeds the model capabilities. This has to inform what the intended functionality of the AI should be in the first place. In particular, it might be necessary to look beyond the inclination to replicate human skills with AI agents.
- **Support pilots both in forming expectations and getting feedback about system states:** Today's flight decks allow pilots to form expectations and to get feedback about system states which they can compare. Complex black-box AI models can potentially prevent both. XAI mostly addresses the feedback part by helping users to understand the model outputs. However, for the forming of expectations about system states, it is likely necessary to design the system around active engagement of pilots, given that situation awareness is lowered when operators are processing information passively [26].

Interface-level guidelines:

- **Provide explanations when system outputs deviate from pilot expectations:** The interviewees' statements suggest that explanations are especially important when system outputs deviate from pilot expectations. In such situations of surprise, explanations may help pilots to decide whether the system output is unreasonable or particularly insightful. This is in line with studies showing that medical professionals particularly seek explanations when they are surprised by the AI's suggestion [27].
- **Allow effortless expression of intentions:** Participants considered the inability of the system to understand pilots' intentions as one of the primary reasons for inappropriate automation behavior. The same issue can be expected with advanced AI-based automation. In many cases, effective correction of unhelpful or even counterproductive AI outputs thus comes down to enabling effortless communication of pilot intentions to the system.
- **Allow effortless bypassing, disengaging, or overwriting of AI functionality:** As pilots bear the full responsibility for the safety of their flight, they need to have the

authority to take whatever action they deem necessary. It must therefore be always possible for the pilot to bypass, disengage, or overwrite any AI functionality, e.g. when AI outputs are unhelpful and not easily correctable.

C. Limitations

Given the difficulty of recruiting pilots for extensive interviews, our sample size is small with only four participants. However, whether AI can handle the operational complexities of flying was the key concern for all participants. We are therefore confident that a larger sample size would have produced similar results.

Moreover, while our guidelines point out important design considerations, they require further empirical validation.

VII. Conclusion

Our interviews and their analysis point to a range of potential usability issues of AI-driven cockpit assistant systems—and as we believe, of AI systems more generally—that are frequently overlooked: System designers tend to focus on the benefits of AI systems and how they should communicate with users so that these benefits come to fruition. What is often missed are the adverse effects when the operational complexity exceeds the capabilities of the AI system. Our guidelines point toward these issues on a very high level, but much more work is necessary to understand how to address them under real operating conditions. We see the recent works on designing for imperfect AI as a promising path in this regard.

Acknowledgment

This work was supported by the German Federal Ministry for Economic Affairs and Energy (BMWi) under the LuFo VI-1 program, project KIEZ4-0.

We thank our participants for their time and valuable input.

References

[1] N. B. Sarter, D. D. Woods, and C. E. Billings, "Automation surprises," in *Handbook of Human Factors & Ergonomics*, 2nd ed. Wiley, 1997, pp. 1926–1943.

[2] S. W. A. Dekker and D. D. Woods, "MABA-MABA or abracadabra? Progress on human-automation co-ordination," *Cognition, Technology & Work*, vol. 4, no. 4, pp. 240–244, Nov. 2002.

[3] C. E. Billings, "Human-centered aviation automation: Principles and guidelines," NASA Ames Research Center, Technical Report NASA-TM-110381, A-961056, NAS 1.15:110381, Feb. 1996.

[4] EASA, "Artificial intelligence roadmap: A human-centric approach to AI in aviation," Tech. Rep., Feb. 2020.

[5] B. Shneiderman, "Human-centered artificial intelligence: Three fresh ideas," *AIS Transactions on Human-Computer Interaction*, vol. 12, no. 3, pp. 109–124, 2020.

[6] D. Schuler and A. Namioka, *Participatory design: Principles and practices*. CRC Press, 1993.

[7] EASA, "EASA concept paper: First usable guidance for level 1 machine learning applications," Tech. Rep., Apr. 2021.

[8] A. Abdul, J. Vermeulen, D. Wang, B. Y. Lim, and M. Kankanhalli, "Trends and trajectories for explainable, accountable and intelligible systems: An HCI research agenda," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. ACM, Apr. 2018, pp. 582:1–582:18.

[9] F. Yang, Z. Huang, J. Scholtz, and D. L. Arendt, "How do visual explanations foster end users' appropriate trust in machine learning?" in *Proceedings of the 25th International Conference on Intelligent User Interfaces*, ser. IUI '20. ACM, Mar. 2020, pp. 189–201.

[10] Y. Zhang, Q. V. Liao, and R. K. E. Bellamy, "Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ser. FAT* '20. ACM, Jan. 2020, pp. 295–305.

[11] C. J. Cai, J. Jongejan, and J. Holbrook, "The effects of example-based explanations in a machine learning interface," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, ser. IUI '19. ACM, Mar. 2019, pp. 258–262.

[12] H.-F. Cheng, R. Wang, Z. Zhang, F. O'Connell, T. Gray, F. M. Harper, and H. Zhu, "Explaining decision-making algorithms through UI: Strategies to help non-expert stakeholders," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. ACM, May 2019, pp. 559:1–559:12.

[13] G. Bansal, T. Wu, J. Zhou, R. Fok, B. Nushi, E. Kamar, M. T. Ribeiro, and D. S. Weld, "Does the whole exceed its parts? The effect of AI explanations on complementary team performance," *arXiv:2006.14779 [cs]*, Jun. 2020.

[14] V. Lai and C. Tan, "On human predictions with explanations and predictions of machine learning models: A case study on deception detection," in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, ser. FAT* '19. ACM, Jan. 2019, pp. 29–38.

[15] R. J. Shively, J. Lachter, S. L. Brandt, M. Matessa, V. Battiste, and W. W. Johnson, "Why human-autonomy teaming?" in *Advances in Neuroergonomics and Cognitive Engineering*, ser. AHFE 2017, vol. 586. Springer International Publishing, Jun. 2017, pp. 3–11.

[16] J. Y. C. Chen, S. G. Lakhmani, K. Stowers, A. R. Selkowitz, J. L. Wright, and M. Barnes, "Situation awareness-based agent transparency and human-autonomy teaming effectiveness," *Theoretical Issues in Ergonomics Science*, vol. 19, no. 3, pp. 259–282, May 2018.

[17] S. L. Brandt, J. Lachter, R. Russell, and R. J. Shively, "A human-autonomy teaming approach for a flight-following task," in *Advances in Neuroergonomics and Cognitive Engineering*, ser. AHFE 2017, vol. 586. Springer International Publishing, 2017, pp. 12–22.

[18] M. Matessa, "Using a crew resource management framework to develop human-autonomy teaming measures," in *Advances in Neuroergonomics and Cognitive Engineering*, ser. AHFE 2017, vol. 586. Springer International Publishing, Jun. 2017, pp. 46–57.

[19] R. J. Shively, S. L. Brandt, J. Lachter, M. Matessa, G. Sadler, and H. Battiste, "Application of human-autonomy teaming (HAT) patterns to reduced crew operations (RCO)," in *Engineering Psychology and Cognitive Ergonomics*, ser. EPCE 2016, vol. 9736. Springer International Publishing, Jun. 2016, pp. 244–255.

[20] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[21] R. Parasuraman and D. H. Manzey, "Complacency and bias in human use of automation: An attentional integration," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 52, no. 3, pp. 381–410, Jun. 2010.

[22] E. L. Wiener, "Human factors of advanced technology (glass cockpit) transport aircraft," NASA Ames Research Center, Contractor Report 177528, 1989.

[23] J. D. Weisz, M. Muller, S. Houde, J. Richards, S. I. Ross, F. Martinez, M. Agarwal, and K. Talamadupula, "Perfection not required? Human-AI partnerships in code translation," in *Proceedings of the 26th International Conference on Intelligent User Interfaces*, ser. IUI '21. ACM, Apr. 2021, pp. 402–412.

[24] C. J. Cai, M. C. Stumpe, M. Terry, E. Reif, N. Hegde, J. Hipp, B. Kim, D. Smilkov, M. Wattenberg, F. Viegas, and G. S. Corrado, "Human-centered tools for coping with imperfect algorithms during medical decision-making," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. ACM, May 2019, pp. 4:1–4:14.

[25] M. Lindvall, C. Lundström, and J. Löwgren, "Rapid assisted visual search: Supporting digital pathologists with imperfect AI," in *Proceedings of the 26th International Conference on Intelligent User Interfaces*, ser. IUI '21. ACM, Apr. 2021, pp. 504–513.

[26] M. R. Endsley, "From here to autonomy: Lessons learned from human–automation research," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 59, no. 1, pp. 5–27, Feb. 2017.

[27] M. Jacobs, J. He, M. F. Pradier, B. Lam, A. C. Ahn, T. H. McCoy, R. H. Perlis, F. Doshi-Velez, and K. Z. Gajos, "Designing AI for trust and collaboration in time-constrained medical decisions: A sociotechnical lens," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, May 2021, pp. 659:1–659:14.