Human factors and usability testing have gained rapid acceptance in healthcare during the last decade as a result of guidelines put in place by regulatory bodies and standard development organizations, such as the Food and Drug Administration, International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC), and Association for the Advancement of Medical Instrumentation. Although healthcare usability professionals and healthcare organizations should continue to adopt design and research practice from industry standards, we should also continue to seek out human-centered research (e.g., psychology, cognitive neuroscience) and integrate new research findings as a key part of our understanding of best practices. We should strive to identify basic (as opposed to applied) research that can be used to augment and improve regulatory guidance as we develop more effective methodologies and advance our field.

Academia has a history of struggling to disseminate findings throughout relevant industries, and healthcare is no exception. However, human factors practitioners are well positioned to serve as the bridge between basic, human-centered research findings and the medical device/healthcare industry. Without such a bridge, a delay occurs between scientific discoveries surrounding human cognitive capacity and their application in the real world (i.e., medical devices and medical environments). Such a delay stagnates progress in applied research and design techniques, including usability testing, root cause probing, user experience design, and risk mitigation.

This article seeks to address a specific area of human-centered research and its application in user experience, usability, and human factors practice: the evolving understanding of human attentional capacity and how usability engineering practices can leverage that understanding when designing and developing medical devices. The author asserts the following:

1. Scientific literature investigating attention and cognitive control provides valuable information that we, as an industry, can apply to usability research and device design.
2. Medical device users regularly use devices in low-cognitive resource states, decreasing the attention they’re able to pay to device interactions, increasing risk of error.
3. Attention-related risk in device interaction is identifiable and mitigable using techniques, such as the perception-cognition-action (PCA) analysis and usability testing.
4. To mitigate risk related to inattention, medical device development teams should use “low-demand design,” which seeks to facilitate safe and effective use for low-resource cognition, rather than demanding that users allocate more resources to a feature, warning, or task.
   a. Rather than just keeping interfaces simple, understanding and identifying intended user groups’ preconceptions (e.g., mental models, automatic thought patterns) and limitations (e.g., finite resources) also is important.
   b. Understanding these facets of user cognition can help in designing complex devices and environments that ask less of the user and allow them to function within their limitations, without increasing risk.

**Understanding Attention and Cognitive Control**

Attention has been defined as “the concentration of awareness on some phenomenon to the exclusion of other stimuli.” Researchers explain this colloquially as a cognitive spotlight, “illuminating” certain stimuli and leaving others unattended or “in the dark.” According to behavioral evidence that has
recently emerged, attention, in addition to being exclusive, also is a finite cognitive resource.\textsuperscript{6,7} That is, attentional capacity decreases as we use it and must be replenished once spent. In the case of human cognitive functioning, the dominant resource is glucose, which fluctuates within a given individual based on behaviors, as well as across individuals.\textsuperscript{8} Individual differences in attentional capacity can stem from any number of biophysiological factors, but fluctuations within an individual are fairly well understood. Inadequate diet, lack of sleep, high-stress scenarios, and complex tasks requiring great concentration deplete glucose levels and, consequently, attentional capacity.\textsuperscript{9}\textsuperscript{10} When resources must be allocated across multiple or highly demanding functions, a situation termed “high cognitive load” results. Subjecting an individual to a high cognitive load has been shown to negatively affect the individual’s task performance across multiple paradigms, increasing error rates and/or decreasing speed, accuracy, and efficiency.\textsuperscript{11,12}

The discussion of attentional resources as finite is particularly relevant for medical device developers, because device users are frequently forced to assume high cognitive loads.\textsuperscript{13–15} For example, physicians are on call at night, nurses work “double twelves,” technologists manage complex software, patients and lay users work with needles despite phobias, and so on. Nearly every user must concentrate to avoid making mistakes. The very nature of medical practice and the associated stakes demand substantial exertion of cognitive resources. Such exertion depletes cognitive resources, thereby requiring users to prioritize their resources and differentially allocate them across tasks (i.e., attend more to some tasks than others). This is especially true for instances in which an individual lacks the means to replenish his/her attentional resources but is asked to continue operating at a high level following the depletion of resources. Cognitive resources can be replenished acutely (i.e., food) or over time (e.g., cognitive restoration, sleep, “mental breaks”).\textsuperscript{16,17}

The good news is that we are not constantly stuck exerting attentional resources (i.e., controlled processing) on every task. The brain can practice a certain task long enough—or the task may be intuitive enough at the outset—that its execution becomes automatic, requiring fewer resources to perform successfully.\textsuperscript{18,19} Flexibly transitioning between controlled and automatic processing is one strategic method by which our brains preserve their cognitive resources. However, automaticity can also lead to increased errors when applied to tasks with which people are not familiar or that are not intuitive.\textsuperscript{20}

Throughout the healthcare field, myriad situations exist for which automatic cognition is ineffective or inappropriate (e.g., navigating complex patient anatomy, interpreting graphics or notifications pertaining to insulin levels). When automatic processes are used to complete a task that requires controlled thought, errors may occur. This can be illustrated using the example of powered contrast injection systems, in which contrast media are injected at high pressures (e.g., 1,000 PSI) to image vascular structures. Recently, thinner microcatheters have been engineered to allow for less invasive navigation, and if a user were to choose the same high pressure with the microcatheter as he/she did with the previous catheter of choice, the microcatheter could “whip” (i.e., move suddenly and erratically) under the pressure. This could cause physicians to lose their place in hard-to-reach areas and damage patient vessels. If a technologist or physician who has been working with the same 6-Fr catheter for 20 years switches to a microcatheter, he/she could inadvertently leverage automatic processing when choosing the pressure at which to inject the contrast. This would increase the likelihood of catheter whip and create undue risk to the patient.

In this example, the user did not possess or did not allocate the attention required to recognize or remember the type of catheter and process the different tolerance. Thus, the user failed to overcome the automatic process of protocol selection compatible with his/her previous device. This example demonstrates a combination of inattention and negative transfer (i.e., a learned, automatic process), but automaticity can intrude very differently across contexts and devices.

In light of the issues that inappropriate automaticity can cause, one important task for
The more complex the environment, the more resilient a device should be in the face of automatic thought.

Medical device development teams is to understand the cognitive demands inherent to the intended use environment. In particular, when the environment is found to regularly require large amounts of controlled attentional resources, development teams must identify elements of the device that demand large amounts of attentional resources from users, because users may or may not possess the needed resources at any given moment. In simple terms: The more complex the environment, the more resilient a device should be in the face of automatic thought.

To make a device more tolerant of, or compatible with, low attention and automaticity, development teams need to pinpoint sources of high demand in the workflow. Upon identifying sources of high demand within a device (using iterative risk analyses and usability testing), the teams can work to redesign the feature(s) in question, such that the device facilitates successful use, even during low-attentional-resource or automatic processing. These types of design modifications have the potential to facilitate safer device use and decrease the frequency of attention-related errors. That said, the process of identifying and mitigating inattention can be complex, and understanding best practice for that process starts with how we think about attention and its role(s) throughout the interaction process.

Attention in the PCA Analysis
ISO/IEC 62366-2 describes a method for error assessment known as perception-cognition-action (PCA) analysis. Recommended as a process to conduct alongside the human factors task analysis, the PCA analysis categorizes errors as occurring in one of the three, titular stages of interaction. The reason for including PCA analyses in the risk management process is that the stage in which an error occurs can aid product development teams in determining optimal risk mitigation strategies. For example, if the error were an incidental button press on a surgical robot as the user attempted to press the correct button (i.e., an action error), designers might change the spacing or location of the buttons. They also might limit button function, such that only relevant buttons are active during the appropriate state. This is in contrast to a how one might mitigate a button press due to a perception error (e.g., complementary colors, rather than different shades of a similar hue). Although the PCA analysis can be extremely effective in helping to choose mitigations, its interaction with attention is tricky.

Unlike muscular movement or visual acuity, attention (i.e., the selective allocation of cognitive resources) influences each of the three stages of interaction, as illustrated in Figure 1. For any controlled processes, users need to pay attention to a stimulus to sense and perceive it. Once perceived, attention is

Figure 1. Involvement of attention in a perception/cognition/action model for medical device interaction processes
Root Cause: Inattention

To understand the nuances of inattention, we should first make a clear distinction between high cognitive load and low cognitive resources, because although the former often leads to the latter, they need not occur in tandem. In usability research, practitioners frequently leverage high cognitive load as a root cause for use error, but the effect of a high cognitive load (i.e., draining a user’s resources) can last longer than the high load itself. That is, an individual’s attentional resources can be low, regardless of the cognitive load under which the individual is currently operating. This distinguishes the root cause of inattention from that of high cognitive load. During a medical device task that requires attentional resources, users may not be able to exert enough resources to succeed in completing the task. These are the scenarios during which researchers observe “inattention errors.”

When it comes to discovering an error’s root cause, most practitioners probably would agree that best practice includes focusing on system- and design-based root causes rather than human-centered root causes (e.g., avoiding root causes such as “human error”). After all, “don’t blame the user” has become the unofficial mantra of the field. However, human-centered root causes can’t always be avoided, and acknowledging the limitations of human cognitive capacity does not need to be synonymous with blaming the user. Accounting for limitations, such as finite cognitive resources, can greatly improve safety and user experience, as long as we propose mitigations that strive to create more considerate and intuitive devices.

We can determine a human-centered root cause and use the research findings of cognitive neuroscience to designate a change to device design as the resulting mitigation. Inattention is one human-centric root cause that possesses great potential to improve design. However, due in part to the complexity and ambiguity surrounding attention’s role across the interaction process, the author has observed practitioners treating inattention errors as if they fall under the “cognition” category. Although this is not inaccurate, it is problematic for two reasons:

1. Cognition is the most intricate and opaque of the PCA categories and, therefore, the most difficult to mitigate. Controlling the way in which users process the stimuli a device presents is unrealistic compared with our ability to control the stimuli themselves.
2. As shown in Figure 1, cognition is the only step in the process that does not interact directly with the user interface/device—the only elements over which reasonable control can be exerted. Thus, by classifying all inattention errors as stemming from cognition, we ostensibly assert that inattention-related errors cannot reasonably be mitigated. However, given the presence of attention throughout the interaction process (i.e., perception, cognition, and action), this is not necessarily the case.

How exactly does this affect usability testing itself? The way in which inattention should be analyzed as a root cause, even under the perception and action categories, deviates from the norm. Although typical root cause probing seeks to ask “why?,” “Why weren’t you paying attention?” is not the right question to ask in the face of inattention errors. This is because human beings, as a species, possess poor metacognitive abilities. (Metacognition can be thought of as the awareness and understanding of one’s own thought processes.) Expecting an individual to accurately identify
Even if a participant were to properly identify the precise point in their processing at which resources were incorrectly allocated, the root cause likely would not be one that development teams could reasonably mitigate.

The underlying reason for their own inattention during root cause probing is unrealistic, to the point that asking participants this question during usability testing may stymie the root cause analysis altogether.

The following example demonstrates the method by which practitioners might probe for root-cause when categorizing inattention errors under “cognition”:

- **Moderator:** “I noticed you moved the pointer here [indicates location] during the first task. Can you tell me what was happening there?”
- **Participant:** “Oh, I meant to line up [the pointer] with this [hash mark], but I guess I didn’t make it all the way there.”
- **Moderator:** “Can you think of a reason you didn’t make it there during the task today?”
- **Participant:** “I don’t know; I just wasn’t paying attention.”
- **Moderator:** “Why do you think you weren’t paying attention?”
- **Participant:** “I don’t know ... because my brain isn’t working? Maybe I just haven’t had enough coffee. I don’t know. I have no idea. I just wasn’t.”

Root cause probing that focuses on the individual’s failure to attend serves little purpose other than frustrating both the moderator and participant. Thus, even if a participant were to properly identify the precise point in their processing at which resources were incorrectly allocated, the root cause likely would not be one that development teams could reasonably mitigate.

Conversely, by approaching inattention error root cause analyses by seeking to understand what it was the participant neglected to pay attention to, as well as what they think might have helped them complete the task correctly, researchers can begin to understand device features and workflow tasks that fail to facilitate low-resource cognition. Consider the following example:

- **Moderator:** “I saw that you didn’t take a measurement on this screen during that second task. Can you think of a reason you might not have done that today?”
- **Participant:** “Oh wow, I didn’t even think about it. I guess I wasn’t really paying attention.”
- **Moderator:** “What weren’t you paying attention to?”
- **Participant:** “The instruction bar. Well, the whole left-hand side, really.”
- **Moderator:** “Do you remember what you were thinking about or doing at that point?”
- **Participant:** “I was definitely still looking at the screen, but I was focused on the image. I had no idea it wanted me to measure the [item] there.”
- **Moderator:** “Why were you focused on the image?”
- **Participant:** “It seemed more important. The image is always going to be the most important thing for us. Plus, looking at it now, everything on the left is grey and the image is in color.”
- **Moderator:** “Is there anything you think might help you notice the measurement instructions?”
- **Participant:** “More color, I guess ... just something I’m likely to see.”

In this example, rather than hitting a dead end in the root cause analysis, the moderator was able to understand what was—and was not—drawing the participant’s attention. With the information above, the researchers can now provide a root cause of “inattention, software design.” This dual approach to root cause determination tells designers that inattention caused the error while also tying the inattention error to a specific interaction point. The details of the analysis then will further clarify relevant device features that designers and engineers may wish to modify.

In the example above, it may be tempting to heed the participant and add more color to the instructions. However, the most important aspect of interaction was the phrase, “The image is always going to be the most important thing for us.” Rather than fighting a user’s natural response to focus on the image, software should be designed to facilitate proper use with that tendency in mind. For example, perhaps the cursor automatically becomes the measurement tool when hovering over the image, as measurements are needed. Alternatively, maybe the software takes a measurement automatically and then asks the user to confirm or adjust the end points. The key is to consider solutions that facilitate automatic
processing or ask little of the user (e.g., color-coordinated plugs and ports; Figure 2), such that users can complete a task safely and effectively without expending large amounts of cognitive resources.

Of course, not every error will be found to require mitigation. Prioritizing mitigations based on the error severity and expected frequency remains a foundational aspect of risk management. However, when the development team determines a mitigation is necessary, design solutions for inattention errors should consider the relevant points of interaction on the user interface and use modifications that ask less (rather than more) of users. By designing devices with features that are noticeable, are usable, and involve minimal cognitive effort, we can improve device-centric task performance and decrease the risk of medical errors.

Here are a few methods we can employ to identify ways in which inattention poses risk to users and/or patients:

1. **Invest in ethnography.** Ethnographic research and contextual inquiry are the best ways to understand the demands of a device's ecosystem. Understanding environmental impact on users improves the development team's ability to design a device that fits its role and users' typical cognitive states. Spending time alongside users can yield a surprising amount of detail.

2. **Analyze early.** The human factors task analysis (typically a workflow diagram describing all of the tasks and subtasks performed by users during device interaction) and the analyses that follow (e.g., hazards analysis, PCA, use failure mode and effects analysis) give development teams a head start on mitigating errors, including those related to inattention. When completed thoroughly, this process can save a lot of time and effort by limiting the need for risk-related redesigns in the later stages of development.

3. **Iterate.** As the design of a device changes, potential errors change as well. Usability testing is the most effective means of identifying new, use-related risks throughout the development cycle. Conducting thoughtful root cause analyses during usability testing results in truly actionable data from which designers and engineers can work.

**Designing with Inattention in Mind**

On the occasion that attention-related risk is identified as mitigable by common root cause methods, the instinct of researchers and designers often is to update a given feature in an effort to evoke additional attention from users. For example, caution iconography may be added, the brightness of an LED may be increased, the volume of an alarm may be raised, or a graphical element may be programmed to flash. The logic here is that if users were not paying enough attention to the previous design, the device should demand that users pay more attention to relevant features in the updated design. Although this process is understandable, approaching the solution in this manner is far from optimal, as it fails to account for human cognitive capacity or users’ limited (and frequently low) resources. In fact, efforts to demand more attention compound across devices and features, leading to inefficient use and needless exhaustion of resources. This leaves the brain to optimize itself through desensitization.

A well-documented example of this is “alarm fatigue”—a condition in which clinicians create increasingly higher thresholds for signal detection of alarms due to an overabundance of auditory “caution” and “danger” cues in the hospital environment. This leaves the brain to optimize itself through desensitization.

Efforts to demand more attention compound across devices and features, leading to inefficient use and needless exhaustion of resources. This leaves the brain to optimize itself through desensitization.
are lower. Alarm fatigue demonstrates the limited capacity of the human brain, as well as its ability to adapt processes to increase efficiency. Device design should consider both of these cognitive principles throughout the development process.

Unfortunately, devices are not reasonably able to provide users with additional cognitive resources; therefore, the best approach is to design a device that asks users to exert as few cognitive resources as is safe, allowing users to leverage automatic processing without increasing the risk of error. The author refers to this practice as “low-demand design,” which is a subset of human-centered design that recognizes human cognitive limitations and asks as little as possible of the user while still maximizing usability for intended use cases. Based on current understanding of human cognitive functioning, low-demand design can take many forms. Below are a subset of principles that support this method:

1. **Call attention rather than demanding it.** Emphasize the important features or information by leveraging subtle processes (e.g., visual contrast). This will help users to naturally allocate attention where it’s needed without requiring undue resources (Figure 3).

2. **Keep your interface clean.** Lowering the number of interaction points or stimuli on the user interface allows users to more easily prioritize and differentiate among items (e.g., Figure 4). If a feature is not relevant or is unavailable during a task, it either should not be included (software) or its functions emphasized based on use frequency and importance (hardware and software).

3. **Keep cues multimodal.** Different sensory avenues (e.g., vision, hearing) are thought to access resources via exclusive pathways. This means that two tasks allocating attention to different sensory modalities are more effectively completed in parallel than two tasks allocating attention to the same sense (e.g., listening to one thing and watching another versus listening to two things). For example, a connection that clicks firmly into place, leaving no gap between the male and female ports, provides haptic (the feel of the click), auditory (the sound of the click), and visual (the lack of gap) cues to convey proper connection. Even when a user is distracted visually, for instance, the haptic and auditory feedback might still be perceived. Conversely, if the user were distracted auditorily, he/she might feel and/or see the confirmation of the connection, and so on.

Increasing the application of these and other low-demand design principles in the medical device industry not only improves the user experience but also has the potential to lower error rates and improve patient safety (Privitera M, MacNamee K, unpublished observations).

**Conclusion**

Cognitive science continues to have high relevance for human factors professionals, especially in healthcare, where human cognitive resources are precious, and precision is vital. In gaining a deeper understanding of human cognitive capabilities, we find that acknowledging human cognitive patterns and limitations, such as automatic thought and finite resources, does not necessarily mean we are blaming the user. Instead, acknowledging such limitations provides a theoretical scaffolding that supports development teams in identifying and mitigating risks associated with human limitations, ideally through thoughtful changes to design.

In taking an informed neurocognitive approach to product development tools, such as the PCA analysis, we can better understand inattention-related risk in medical device use. Moreover, as we work toward understanding inattention-related risk, we should guard
against dismissing inattention errors as user-centric risk (i.e., mitigating the cognition phase of interaction) and instead seek out opportunities to identify tasks that push the boundaries of human capabilities and use those data to improve the design of controllable system elements (i.e., making design changes that affect the action and perception stages).

In addition to being conducted frequently, ethnographic research and usability testing should be designed to collect data that are relevant to human neurocognitive capacity (e.g., mental models, environmental demands, system-based cognitive load). Such studies augment and evolve risk analyses throughout the development process. In usability testing, root cause analyses for inattention errors should focus on understanding whether the design is considerate of human attentional limitations, such that development teams can implement appropriate mitigations.

The most valuable mitigations for inattention errors include design updates that use low-demand design principles. Although sometimes more challenging or less obvious, low-demand design provides a deeply beneficial framework from which more considerate and resilient product design can emerge. By designing with an understanding of human limitations, product development teams can simultaneously lower risk, increase safety, and improve user experience. These improvements help drive success for the product and its manufacturer, while contributing to patient and/or user well-being.

References


