1 INTRODUCTION

The fig tree is pollinated only by the insect Blastophaga grossorun. The larva of the insect lives in the ovary of the fig tree, and there it gets its food. The tree and the insect are thus heavily interdependent: the tree cannot reproduce without the insect; the insect cannot eat without the tree; together, they constitute not only a viable but a productive and thriving partnership. This cooperative “living together in intimate association, or even close union, of two dissimilar organisms” is called symbiosis. . . . “Man-computer symbiosis” is a subclass of man-machine systems. There are many man-machine systems. At present, however, there are no man-computer symbioses. . . . The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today. (Licklider, 1960, pp. 4–5)

Although this was written over 40 years ago, Licklider’s vision fully characterizes the current status of interactive computing and contemporary aspirations for its future. Historically, visionaries such as Licklider (1960) and Engelbart (1963) suggested that human–computer symbiosis should augment human intelligence and extend human cognitive abilities. Yet such intelligence augmentation has so far proved elusive for interactive system developers. There is a burgeoning paradigm shift in interactive computing that has the potential to realize these visionary projections; it is called augmented cognition.

Augmented cognition is a constellation of desires, concepts and goals aimed at maximizing human cognitive abilities through the unification of humans and computational systems (Schmorrow and McBride, 2004). As Licklider (1960) suggested, human brains and computing machines should be coupled together very tightly. The essence of augmented cognition is to achieve such coupling by leveraging the latest in powerful imaging techniques that enable mapping of...
distinct and detailed functions of the brain. Specifically, augmented cognition seeks to revolutionize the way that humans interact with computers by coupling traditional electromechanical interaction devices (e.g., mouse, joystick) with psychophysiological interaction (e.g., eye blinks, respiration, heart rate, electroencephalogram), such that subtle human physiological indicators can be used to direct human–system interaction. Fundamental research and related technology developments in this domain have centered on leveraging data from physiological indicators to alleviate and maximize throughput of human information-processing bottlenecks [e.g., sensory, working memory (WM), attention, executive function (EF)]. The basis for much of this work is grounded in the view that human information-processing capabilities are fundamentally the weak link in the symbiotic relationship between humans and computers. As computational prowess continues to increase, human and computer capabilities are ever more reliant on each other to achieve maximal performance. Demanding conditions, such as those associated with homeland security or military operations, call for expertise not from a specific human or computer system, but from a linked human–machine dyad. A dyad that is functionally a human and their computational system, which through shared experience and insight into how they both function, will jointly deliver solutions at a previously unimagined rate far surpassing that of a solitary entity.

Common within a majority of augmented cognition endeavors is the attempt to understand intrinsically how human information processing works so that augmentation schemes might be developed and effectively exploited to enhance human processing capacity. Thus, the central vision of augmented cognition is to extend human abilities substantially via computational technologies designed explicitly to address human information processing limitations.

### 1.1 Human Information Processing Limitations

Current understanding of human information processing suggests that information is perceived through multiple sensory processors. This information is then perceptually encoded (i.e., stimulus is identified and recognized), processed by a WM subsystem that is regulated and controlled by attention via the EF, which may be supported by long-term memory (LTM), to arrive at a decision, which in turn triggers a human response (Baddeley, 1986, 1990, 2000; Wickens, 1992). Within human information processing there are thus several “bottlenecks” or points of limited processing capacity, including sensory memory, WM, attention, and executive function.

#### 1.1.1 Sensory Memory Bottleneck

Sensory memory is responsible for encoding information and converting it to a usable mental form (Atkinson and Shiffrin, 1968, 1971). There is a different sensory memory system for each of the human senses, including visual, auditory, tactile (haptic), olfactory, and gustatory. Behavioral studies suggest that human information processing begins with information being perceived on average in about 100 ms (Cheatham and White, 1954; Harter, 1967) by one of the sensory processors. The visual **iconic** sensory memory modality has been suggested to have an average capacity of about 17 items, and this iconic percept is fleeting, decaying completely, on average, in about 200 ms if it does not transfer to WM (Sperling, 1960, 1963; Averbach and Coriell, 1961; Neisser, 1967). Audition, or **echoic** sensory memory, is suggested to have an average capacity of five items and is a bit more persistent, with the “internal echo" lasting an average of about 1.5 seconds (Neisser, 1967; Darwin et al., 1972). **Haptic** sensory memory is very limited in terms of capacity (Watkins and Watkins, 1974; Maher and Miles, 2002) and has a decay rate between 2 and 8 seconds (Bliss et al., 1966; Posner and Konick, 1966; Lachman et al., 1979). Little is known about olfactory and gustatory sensory memories. In general, a considerable amount of information can be perceived if it is allocated across multiple sensory systems. Thus, **given the limited capacity of sensory memory, augmented cognition seeks to enhance sensory perception by exploiting multiple sensory channels for increased input capacity** (see Table 1). Sensory stimuli that have passed the sensory memory bottleneck and are

<table>
<thead>
<tr>
<th><strong>Table 1 Tenets of Augmented Cognition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Human Information-Processing Bottleneck</strong></td>
</tr>
<tr>
<td>Sensory memory</td>
</tr>
<tr>
<td>Working memory</td>
</tr>
<tr>
<td>Attention</td>
</tr>
<tr>
<td>Executive function</td>
</tr>
</tbody>
</table>
1.1.2 Working Memory Bottleneck

Working memory allows people to maintain and manipulate information that has been perceived by sensory memory and is currently available in a short-term memory store. In general, WM is described as a functional multiple component of cognition that allows humans to comprehend and mentally represent their immediate environment, to retain information about their immediate past experience, to support the acquisition of new knowledge, to solve problems, and to formulate, relate, and act on current goals (Baddeley and Logie, 1999, p. 29). It is considered a temporary active storage area where information is manipulated and maintained for executing simple and complex tasks (e.g., serial recall, problem solving). Working memory is divided into separate processes that are required for short-term storage (according to Baddeley and Logie’s (1999) model, these include the phonological loop and visuospatial sketchpad) and for allocating attention and coordinating maintained information (i.e., the executive function).

Working memory is still being defined, and recent research has suggested dissociations in both the phonological loop (i.e., phonological store vs. articulatory rehearsal mechanism) (Baddeley and Logie, 1999) and visuospatial sketchpad (visual form and color recognition vs. localization) (Carlesimo et al., 2001; Mendez, 2001; Pickering, 2001). In general, WM is said to have a limited capacity of about seven chunks, a rapid decay rate of about 200 ms, and a recognize-act processing time of 70 ms, on average (Miller, 1956; Card et al., 1983). Recent research suggests, however, that presenting information multimodally can in fact enhance human information processing via an increase in WM capacity, with gains on the order of three times Miller’s (1956) “magic number” of seven being realized in one recent study (Samman et al., 2004). These gains could be tempered if the costs for modality switching are high; this is discussed in the next two sections.

Given separable WM components and WM capacity enhancements based on modality, Wickens’s (1984) Multiple Resource Theory (MRT) can be expanded to suggest that modality-based resources can be utilized strategically at different points in user interaction to streamline a user’s cognitive load (Stanney et al., 2004). In such a case, total WM capacity will depend on how dissimilar streams of information are in terms of modality. An expanded MRT would address how to allocate multimodal WM resources, particularly during multitasking, in such a way as to allow attention to be time-shared among various tasks. Thus, augmented cognition seeks to support simultaneous processing of competing tasks by strategically allocating data streams to various multimodal sensory systems while maintaining multimodal information demands within WM capacity (see Table 1).

1.1.3 Attention Bottleneck

Three general categories of attention theories can be found in the literature: (1) “cause” theories, in which attention is suggested to modulate information processing (e.g., via a spotlight that functions as a serial scanning mechanism or via limited resource pools); (2) “effect” theories, in which attention is suggested to be a by-product of information processing among multiple systems (e.g., stimulus representations compete for neuronal activation); and (3) hybrids that combine cause-and-effect theories (Fernandez-Duque and Johnson, 2002). In general, attention is suggested to be a selective process via which stimulus representations are transferred between sensory memory and WM and then contributes to the processing of information once in working memory. Attention improves human performance on a wide range of tasks, minimizes distractions, and facilitates access to awareness (i.e., focused attention). In the best case, attention helps to filter out irrelevant multimodal stimuli. In the worst case, critical information is lost due to overload of incoming information, stimulus competition, or distractions. Thus, if one were to try to enhance WM via multimodal interaction, such stimulation would impose a trade-off between the benefits of incorporating additional sensory systems and the costs associated with dividing attention between various sensory modalities. Attention must thus be moderated judiciously to enhance human–computer symbiosis. Augmented cognition seeks to “build systems that sense, and share with users, natural signals about attention to support fluid mixed-initiative collaboration with computers ... an assessment of a user’s current and future attention (could thus) be employed to triage computational resources” (Horvitz et al., 2003, p. 52). Thus, with augmented cognition, computers will become aware of subtle cues emanating from humans indicating how they are prioritizing incoming information (i.e., directing attention) and will capitalize on these cues to enhance human information processing (see Table 1).

1.1.4 Executive Function Bottleneck

The EF system is suggested to be responsible for selection, initiation, and termination of human information processing routines (e.g., encoding, storing, and retrieving) (Matlin, 1998; Baddeley, 2003). It controls (i.e., focuses, divides, and switches) attention, integrates information from WM subcomponents, and connects WM with contextually triggered information from LTM. The EF is thus associated with regulatory processes underlying the control of human information processing and sheds light on operational costs associated with these control activities (Zakay and Block, 2004). The EF is thought to be especially active in handling novel situations (i.e., those with contextual ambiguity), such as those involving planning or decision making, error correction or troubleshooting, novel sequences of actions or responses, danger or technical difficulty, or the need to overcome habitual responses (Norman and Shallice, 1980; Shallice, 1982). When a person faces such contextual ambiguity during human information processing, high-level
control functions of the EF become engaged. During such processing, a person will retrieve the multiple interpretations associated with a given uncertain situation, choose the more likely interpretation based on context and frequency of occurrence, discard alternative interpretations, and mark that point in their information representation as a choice point (Zakay and Block, 2004). Reducing contextual ambiguity, and thus effortful EF processing, would involve easing selection among multiple interpretations by increasing the number of contextual cues associated with any given alternative.

As indicated previously, frequent switching between one modality or task and another will incure a cost of switching that will be associated with inhibitions of responses to the previous modality stimuli or task, selection and activation of the response best associated with the new modality or task context, and resequencing of these stimuli. Since more frequent switching may entail greater contextual changes, it is expected to engage effortful EF processing. Thus, it is important during modality switching to consider the cost of such contextual changes. Augmented cognition seeks to enhance information processing by directing recall of contextual information that cues optimal interpretation of incoming information and moderates the effects of modality switching (see Table 1).

2 COGNITIVE STATE ASSESSORS

Augmented cognition seeks to enhance human–system interaction substantially by adopting a paradigm shift from primarily passive systems dependent on user input to proactive systems that gauge and detect, via diagnostic psychophysiological sensors, human information processing bottlenecks and then employing augmentation strategies to overcome these limitations. To realize this paradigm shift, one must first be able to characterize cognitive state such that the noted bottlenecks can be monitored and regulated appropriately. Research in psychophysiology, principally through brain-imaging techniques, has established a correspondence between cognitive processors and particular brain structures that have an identifiable locus in the brain. This allows use of neural signals from those structures as a diagnostic tool of cognitive load, which can be measured in real time while a person is engaged with an interactive system. Such psychophysiological data streams can be used to characterize cognitive state, specifically current load on information processing bottlenecks.

2.1 Psychophysiological Techniques for Capturing a Cognitive State

Many human–system interactive situations do not provide sufficient human performance information that can be used to infer cognitive state or what shall herein be called an operator’s functional state (OFS). This is especially true of highly automated systems, which for the most part put the human in a monitoring role (Byrne and Parasuraman, 1996). Because system monitoring does not require overt behavioral responses, it is difficult to assess user state. Thus, a user may not be in an optimal state at all times, and system corrections or malfunctions may not be detected and responded to correctly. A methodology is needed that provides accurate assessment of OFS in the absence of overt performance data and to provide additional information when performance data are available. Psychophysiological measures have been suggested to fill this role.

Psychophysiological signals are always present and can often be collected unobtrusively, thereby providing a source of uninterrupted information about user state (Kramer, 1991; Wilson and Eggemeier, 1991; Scerbo et al., 2001; Wilson, 2002a). Correlations between psychophysiological measures and OFS have been described (Wilson and Schlegel, 2003). Although these correlations do not prove causality, they do suggest that psychophysiological measures can be used to assess OFS and further, that this information can be used to modify system parameters to meet the momentary needs of users (i.e., cognitive augmentation via adaptive aiding). Of the several criteria for implementation of OFS driven adaptive aiding, three crucial ones are that (1) significant and meaningful system performance improvements must be demonstrated; (2) the sensors used must be nonintrusive to a user’s primary task, as this would hinder human–system performance; and (3) their use must be acceptable to users.

For widespread adoption, it must be demonstrated that OFS assessment and aiding either (1) improve human performance and enhance job success for work-related applications, or (2) enhance the interactive experience for entertainment-based applications. An example of a successful application of adaptive aiding is the use of antigravity (anti-g) suits, which require wearing additional gear that inflates at predetermined g-levels. These suits have been proven to save lives because they can prevent g-induced loss of consciousness in jet pilots and have therefore met with wide acceptance.

2.1.1 Current Status

In the past, the typical approach when using psychophysiological measures to assess OFS was to collect one or more measures and demonstrate that statistically significant differences exist between at least two levels of task demand or human state such as fatigue. Most of this research has been conducted in the laboratory. However, a growing body of research is expanding into operational environments. Psychophysiological measures have been applied successfully in driving, flight, and other test and evaluation environments (Wilson, 2002a). For example, heart rate has been shown to be increased significantly under high mental workload conditions compared to low mental workload conditions during flight (Hankins and Wilson, 1988; Wilson, 2002b). Electroencephalography (EEG), a physiological measure of the momentary functional state of cerebral structures, provides useful information about both high cognitive workload and inattention (Kramer, 1991; Wilson and Eggemeier, 1991; Gundel and Wilson, 1992; Sterman and...
can be directed through the scalp and skull and
Using near-infrared light emitters, near-infrared energy

2.1.4 Functional Near-Infrared Sensors

cortical blood volume, and neuronal activity.

sensors provide information about brain oxygen levels,
new sensor technology has been developed that pro-
video cameras that image the face from a distance,
eye activity can be recorded using
to be able to record EEG from anywhere on the scalp
and EOG. Currently, EEG can be recorded from non-
sensors are capacitive coupled and optical sensors. These
typically do not like to wear such sensors and associated equipment. Further, the sensors
are usually attached to the skin with some type of adhe-
repeated application in a day-to-day operational environment may cause skin irritation. There
are less invasive options, such as pupillometry and
eye point of regard, which are typically recorded with
are less invasive options, such as pupillometry and
are off-head eye point of regard devices are available but they restrict head movement,
which limits their applicability in real-world environments.

2.1.3 New Sensor Technologies

New sensor technologies promise to provide users with
psychophysiological methods cannot be used. For example, positron emission tomography (PET), func-
tional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG) are not practical OFS
gauges because the associated recording equipment is
requirement, among other prohibiting conditions. Even those
measures that are less prohibitive have drawbacks. Almost all currently available, operationally useful
psychophysiological sensors require contact with a
user’s body and use some form of electrolyte sen-
tors. This is the case for EEG, electrocardiography
(ECG), electrooculography (EOG), and electromyog-
graphy (EMG). Users typically do not like to wear such
sensors and associated equipment. Further, the sensors
are usually attached to the skin with some type of adhe-
sive, and repeated application in a day-to-day operational environment may cause skin irritation. There
are less invasive options, such as pupillometry and
are head-mounted sensors. Off-head eye point of regard
devices are available but they restrict head movement,
which limits their applicability in real-world environ-
ments.

2.1.2 Current Technology for Recording
Psychophysiological Data

Numerous psychophysiological measures have been
shown to provide valuable information concerning
OFs in real-world operational environments (Wilson,
2002a; Wilson and Schlegel, 2003). Because of the
restrictions of the operational environment, some
psychophysiological methods cannot be used. For example, positron emission tomography (PET), func-
tional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG) are not practical OFS
gauges because the associated recording equipment is
too restrictive, too large, and requires special shield-
ing, among other prohibiting conditions. Even those
measures that are less prohibitive have drawbacks. Almost all currently available, operationally useful
psychophysiological sensors require contact with a
user’s body and use some form of electrolyte sen-
tors. This is the case for EEG, electrocardiography
(ECG), electrooculography (EOG), and electromyog-
graphy (EMG). Users typically do not like to wear such
sensors and associated equipment. Further, the sensors
are usually attached to the skin with some type of adhe-
sive, and repeated application in a day-to-day operational environment may cause skin irritation. There
are less invasive options, such as pupillometry and
eye point of regard, which are typically recorded with
are head-mounted sensors. Off-head eye point of regard
devices are available but they restrict head movement,
which limits their applicability in real-world environ-
ments.

2.1.2 Current Technology for Recording
Psychophysiological Data

Numerous psychophysiological measures have been
shown to provide valuable information concerning
OFs in real-world operational environments (Wilson,
2002a; Wilson and Schlegel, 2003). Because of the
restrictions of the operational environment, some
psychophysiological methods cannot be used. For example, positron emission tomography (PET), func-
tional magnetic resonance imaging (fMRI), and magnetoencephalography (MEG) are not practical OFS
gauges because the associated recording equipment is
too restrictive, too large, and requires special shield-
ing, among other prohibiting conditions. Even those
measures that are less prohibitive have drawbacks. Almost all currently available, operationally useful
psychophysiological sensors require contact with a
user’s body and use some form of electrolyte sen-
tors. This is the case for EEG, electrocardiography
(ECG), electrooculography (EOG), and electromyog-
graphy (EMG). Users typically do not like to wear such
sensors and associated equipment. Further, the sensors
are usually attached to the skin with some type of adhe-
sive, and repeated application in a day-to-day operational environment may cause skin irritation. There
are less invasive options, such as pupillometry and
eye point of regard, which are typically recorded with
are head-mounted sensors. Off-head eye point of regard
devices are available but they restrict head movement,
which limits their applicability in real-world environ-
ments.

2.1.3 New Sensor Technologies

New sensor technologies promise to provide users with
more acceptable recording methods and valuable OFS
data. Sensors that require only “dry” (no electrolyte
or adhesive) contact with the skin have been devel-
oped (Kingsley et al., 2002; Trejo et al., 2003). Two
approaches that are being explored for dry EEG sen-
sors are capacitive coupled and optical sensors. These
technologies can also be used to record ECG, EMG,
and EOG. Currently, EEG can be recorded from non-
haired skin areas such as the forehead, but the goal is
to be able to record EEG from anywhere on the scalp
using these sensors. Eye activity can be recorded using
video cameras that image the face from a distance,
requiring no actual contact with users. Additionally,
new sensor technology has been developed that pro-
vides measures of brain activity using blood flow tech-
nology. For example, functional near-infrared (fNIR)
sensors provide information about brain oxygen levels,
cortical blood volume, and neuronal activity.

2.1.4 Functional Near-Infrared Sensors

Using near-infrared light emitters, near-infrared energy
can be directed through the scalp and skull and
reflected from underlying cerebral tissue. Two types of
cerebral information can be obtained from fNIR. The
first type is hemodynamic response, reflecting oxy-
hemoglobin and deoxyhemoglobin concentrations in
the brain. The consensus is that increased brain activ-
ity results in increased levels of local oxyhemoglobin
and decreased levels of deoxyhemoglobin (Gratton and
Fabiani, 2001). These responses have been used to
investigate cognitive activity (Hock et al., 1997; Vill-
ringer and Chance, 1997; Takeuchi, 1999). The second
type of information that can be obtained from fNIR is
to detect changes in the optical characteristics in brain
tissue that are related to neuronal activity (Gratton and
Fabiani, 2001). The exact cause of these optical
changes is not totally understood. This latter method
is said to provide millisecond temporal resolution; the
first method is much slower. For either procedure the
infrared emitters and sensors have only to touch the
scalp rather than being affixed to it (see Figure 1). The emitter–sensor unit can be held in place using
a strap or cap arrangement. fNIR systems have been
developed that function on hairy areas of the scalp
and so are not restricted to the forehead region. This
developing technology holds a great deal of promise
for advancing our understanding of cognition and may
be used more readily in operational environments than
sensor technologies that require adhesives.

2.2 Transforming Sensors into
Cognitive-State Gauges

To be useful, real-time assessment of cognitive activ-
ity using psychophysiological measures must be trans-
formed from individual measures to cognitive gauges.
Whereas consideration of individual measures provides
valuable information, augmented cognition requires
gauges that are composite estimates characterizing the
functional state of a user (such as those to gauge
load on the human information-processing bottlenecks,
as well as others, such as Kolmogorov entropy of
EEG signals and task load, which are mentioned in
Sections 3.1.3 and 3.1.4). Given the complexity inher-
ent to most operational environments, it is not suf-
ficient simply to be aware that statistical changes
exist in several measures. Measures or gauges must
be able to characterize the functional state of a user
such that this information can be used to implem
adaptive aiding (i.e., triggering of augmentation strategies) in real time in real-world situations. In 2003, the U.S. Defense Department Defense Advanced Research Project Agency (DARPA) conducted a technology integration experiment (TIE) with various psychophysiological sensors (i.e., EEG, event-related potential, fNIR, pupil dilation, heart rate variability, arousal, galvanic skin response) to demonstrate the feasibility of simultaneous data collection (Morrison et al., 2003). The TIE demonstrated that real-time computation of sensor data to produce online gauge information was feasible, and further confirmed that several sensor technologies could be combined with minimal interference. However, substantial variability between human participants in gauge sensitivity suggested the need for additional research. Additional research also needs to be focused on how to transform sensors to specific OFS gauges, such as gauges to measure the load on the human information processing bottlenecks. Thus, augmented cognition seeks to leverage a set of psychophysiological gauges that allow for real-time assessment of cognitive state, particularly current load on information processing bottlenecks, which can then be transformed directly into computer control commands for triggering implementation of augmentation strategies.

3 HUMAN–SYSTEM AUGMENTATION

In Section 1, various human information processing bottlenecks were discussed (i.e., sensory memory, WM, EF, attention). In Section 2, means of gauging the current cognitive load on a person were considered. Augmented cognition seeks to overcome the noted points of limited capacity processing through the utilization of human–system augmentation strategies, which will be triggered by cognitive state gauges. It is suggested that through augmentation strategies, the cost of these bottlenecks (e.g., degraded human performance due to overload, underload, stress, losses in situational awareness, or emotional state) can be overcome.

3.1 Augmentation Strategies

In conventional human–system interaction, an excessive amount of cognitively demanding tasks can be imposed on a user. In such situations, human information processing can break down at any of the bottlenecks. Instead of overloading users, interactive systems should seek to achieve cognitive congeniality (Kirsh, 1996) by (1) presenting an optimal level of task-relevant information and ensuring that it is readily perceived, (2) minimizing cognitive load on WM by sequencing and pacing tasks appropriately, and (3) reducing the number and cost of mental computations required for task success by delegating tasks when appropriate. Taken together, these strategies should increase the speed, accuracy, and robustness of human–system interaction. Each of these augmentation strategies (i.e., task presentation, sequencing, pacing, and delegation) is discussed below. It should be noted that other such strategies can and should be identified. Additional augmentation strategies to consider include but are not limited to techniques for supporting information filtering and triage, multitasking, mixed-initiative interaction, and context-sensitive interaction (Horvitz et al., 2003).

3.1.1 Task Presentation

When designing interactive systems, a central question is which information should be conveyed via which modality. Conventional interactive systems present information to users primarily via visual cues, sometimes offering auditory accessories. Yet to optimize sensory processing, thereby relieving the sensory memory bottleneck, one should consider the types of information each modality is particularly suited to display. Table 2 presents theorized suitability of sensory modalities for conveying various information sources. In addition to suitability, one must consider capacity. As aforementioned, Samman et al. (2004) demonstrated that multimodal WM capacity can reach levels nearly three times that of Miller’s (1956) magic number seven. Thus, rather than overloading a single modality, by distributing information across multiple modalities the WM bottleneck can be relieved. Table 3 represents the WM capacity of various modalities based on several studies (Bliss et al., 1966; Sullivan and Turvey, 1974; Smyth and Pendleton, 1990; Keller et al., 1995; Livermore and Laing, 1996; Woodin and Heil, 1996; Feyerisen and Van der Linden, 1997; Matsuda, 1998; Jinks and Laing, 1999; Laska and Teubner, 1999; Frenchman et al., 2003). The numbers in Table 3 suggest the upper limit on the number of items that should be presented via each modality, as individual modality capacity tends to decline during multimodal multitasking even though overall capacity increases (Samman et al., 2004). Thus, with knowledge of the information sources constituting a given application, a determination of optimal modalities can be made to direct multimodal task presentation. More specifically, after characterizing a given application’s information sources via a task analysis, first a matching to the optimal modality can be determined using Table 2. Then, given the outcome of the related OFS gauges (i.e., current load on sensory and WM bottlenecks), a determination of reserve capacity can be estimated using Table 3 and a selection of the optimal modality made (i.e., the one with the best match from Table 2 and adequate reserve capacity). The applied implication is that in cognitively demanding task environments, not only should information be presented in a modality that is most suitable but also in one that is not currently fully loaded, thereby easing the sensory memory and WM bottlenecks. Thus, the first augmentation strategy is to identify the optimal modality by which to present information based on consideration of suitability principles as well as current psychophysiological measures of cognitive load (see Table 4).

3.1.2 Task Sequencing

Once the modality by which to present an information source is determined, the information event can be scheduled. The MRT (Wickens, 1984) suggests that people are more efficient in time-sharing tasks when different resources are utilized in terms of
Table 2  Theorized Suitability of Modalities for Conveying Various Information Sources\textsuperscript{a}

<table>
<thead>
<tr>
<th>Information Source</th>
<th>Sensory Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Visual</td>
</tr>
<tr>
<td>Spatial acuity (size, distance, position)</td>
<td>++</td>
</tr>
<tr>
<td>2D localization (absolute/relative location in 2D)</td>
<td>++</td>
</tr>
<tr>
<td>3D localization (absolute/relative location in 3D)</td>
<td>□</td>
</tr>
<tr>
<td>Change over time</td>
<td>++</td>
</tr>
<tr>
<td>Persistent attention</td>
<td>++</td>
</tr>
<tr>
<td>Absolute quantitative parameters</td>
<td>++</td>
</tr>
<tr>
<td>Temporal (e.g., duration, interval, rhythm)</td>
<td>□</td>
</tr>
<tr>
<td>Instructions</td>
<td>+</td>
</tr>
<tr>
<td>Rapid cuing (e.g., alerts, warning)</td>
<td>+</td>
</tr>
<tr>
<td>Surface characteristics (e.g., roughness, texture)</td>
<td>+</td>
</tr>
<tr>
<td>Hand–eye coordination (e.g., object manipulation)</td>
<td></td>
</tr>
<tr>
<td>Memory aid (e.g., recognition of a formerly perceived object)</td>
<td>+</td>
</tr>
<tr>
<td>Affective or ambient information</td>
<td>□</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Key: ++, best modality; +, next best; □, neutral; −, not well suited, but possible; −−, unsuitable.

Source: Adapted from ETSI (2002).

Table 3  WM Capacity of Various Sensory Modalities

<table>
<thead>
<tr>
<th>WM Subsystem</th>
<th>Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>2–5</td>
</tr>
<tr>
<td>Verbal</td>
<td>4–7</td>
</tr>
<tr>
<td>Spatial</td>
<td>5–7</td>
</tr>
<tr>
<td>Tactile</td>
<td>3–5</td>
</tr>
<tr>
<td>Kinesthetic</td>
<td>3–5</td>
</tr>
<tr>
<td>Tonal</td>
<td>4–6</td>
</tr>
<tr>
<td>Olfactory</td>
<td>3–4</td>
</tr>
</tbody>
</table>

cognition, response), perceptual (sensory), modality (visual, verbal, spatial, tactile, kinesthetic, tonal, olfactory), visual processing channels (focal, ambient), and WM processing codes (spatial, verbal) (Wickens, 2002). An applied implication of this theory is that time sharing of tasks should be more effective with cross-modal as compared to intramodal information displays. Thus, through systematic sequencing of tasks, simultaneous processing of competing tasks can be allocated strategically across various multimodal sensory systems in an effort to maintain multimodal information demands within WM capacity. Beyond addressing the WM bottleneck, this augmentation strategy can assist in prioritizing incoming information by sequencing cues according to priority, thereby directing attention. When applying this strategy, it is essential to ensure that there is a means to avoid the adaptive state from oscillating too frequently. This can be done through the application of robust controllers (see Section 4). Through systematic control of the adaptive state, this strategy also addresses the EF bottleneck by moderating the effects of modality switching.

To determine task sequencing (i.e., ordering and combining of tasks), a conflict matrix could be calculated following Wickens’s (2002) approach, in which the amount of conflict between resource pairs for
task couplings is determined. This calculation factors in both conflict and task difficulty (i.e., resource demands), resulting in a task interference value. This could be done in conjunction with a time-line analysis (Sarno and Wickens, 1995), which calculates resource demand levels of time-shared tasks over the time during which the tasks are to be performed. In allocating resources, these principles could be coupled with a scheme of task priorities (as derived through an a priori task analysis), which taken together could guide task ordering and combining given current resource constraints (i.e., task interference values and OFS gauge outputs from all four bottlenecks).

The second augmentation strategy is thus to assign modalities to information sources and then schedule them, considering priority, such that they minimize interference over the performance period while leveraging robust controllers to moderate the effects of modality switching (see Table 4). This should help relieve the sensory, WM, attention, and EF bottlenecks.

### 3.1.3 Task Pacing

Time management is an essential component of many dynamic task situations (and is also critical to feedback stability of closed-loop systems, see Section 4). Yet, in cognitively demanding task environments, pacing skills can decline rapidly, as temporal judgments depend on the amount of attentional resources allocated to a temporal processor (Casini and Macar, 1999). Further, internal (self) pacing has been shown via EEG signals to impose higher human information processing demands compared to externally (e.g., via metronome) paced tasks (Gerloff et al., 1998). Disruption of an orderly rhythm is thought to increase the entropy of the human information-processing system, thereby increasing information content due purely to asynchronous pacing of a task. Such disruption can occur when a person becomes overloaded with information, as this often results in delayed event detection and more corrective responses (Boer, 2001). Interestingly, Boer (2001) developed a simple but highly predictive linear model based on Wickens and Hollands’s (2000) MRT, which predicted the effect of various tasks on steering entropy and driver performance. The model demonstrated that steering entropy was affected primarily by loading of spatial tasks, as would be predicted by MRT because driving is a highly spatial task. Thus, to achieve effective time management, a potential augmentation strategy would be to provide external pacing of tasks, which could be achieved by monitoring behavioral entropy (see Table 4). Specifically, the Kolmogorov entropy (K-entropy) of EEG signals can be used to assess information flow (Pravitha et al., 2003). K-entropy is proportional to the rate at which information about the state of a dynamical system is lost in the course of time. This entropy index has been shown to fluctuate with changes in the complexity of human information processing, such as that imposed by fatigue (leading to a lesser extent of information flow through particular brain regions) (Rekha et al., 2003) or information overload (King, 1991) while remaining quite stable during performance of demanding cognitive tasks (Pravitha et al., 2003). Thus, using K-entropy of EEG signals to direct task pacing should help relieve the WM, attention, and EF bottlenecks, as it could help minimize pace the processing of incoming information and minimizes disruptions.

### 3.1.4 Task Delegation

In the context of augmented cognition, the purpose of dynamic task delegation would be to increase information throughput by balancing the utilization of human resources across a network of users. Task delegation allows for distribution of task demands across individuals as well as coordination between humans and automated systems. In task delegation, certain actions required by a particular task performer are delegated to another performer or back to the system itself once task load gets above some threshold (Dearden et al., 2000; Hoc, 2001; Debernard et al., 2002). Such handing off can be implicit (i.e., imposing an allocation based on current OFS load predictions) or explicit, in that it requires an action from the task performer prior to allocation. Although implicit allocation has been shown to lead to better performance than explicit, implicit
allocation does not always meet with user acceptance, as humans like to maintain control of dynamic task situations and become anxious when they lose control (Hock et al., 2002). This, in turn, could affect behavioral entropy, thereby affecting system pacing. Taken together, this could affect system stability properties negatively (see Section 4). Assisted explicit allocation is a compromise, where after detecting an over-load using an OFS gauge of task load, such as the task engagement index used by Prinz et al. (2000, 2003), the interactive system would make an allocation proposal which the human would be able to veto but would not be in charge of allocating. This cooperative task allocation strategy generally leads to effective performance while avoiding complacency by requiring the human to cooperate in the allocation process. Thus, a fourth potential augmentation strategy would be to direct assisted explicit task delegation based on psychophysiological indexes of task load (see Table 4). This should help relieve the attention and EF bottlenecks, as it eases the need to determine what to attend to.

4 ROBUST CONTROLLERS

Although augmentation strategies have the potential to enhance human performance through reducing the load on human information processing bottlenecks, they could also lead to an adaptive state that oscillates too frequently, thereby destabilizing human–system interaction over time. Thus, there is a need to identify techniques for ensuring that changes requested through the augmentation strategies are implemented so as to maintain system stability and enhance human performance. Mathematical system theory deals with the modeling, analysis, and design of complex dynamic systems. Robust control theory is a discipline of mathematical system theory that is concerned with the analysis and design of feedback controllers for situations where there is only partial or incomplete knowledge of the underlying system dynamics. In the work discussed in this chapter, whereby a user’s display/input is adapted based on his or her measured cognitive load, it is important to note that a feedback loop is being closed around the human. Moreover, since the underlying system dynamics involve the human, it is certainly true that only partial knowledge concerning a user’s state will be available, hence the need for this section on robust control.

4.1 Control System Models

Recent developments in the field of cognitive neuroscience have heralded a great deal of change in what is known about human mental operations (Posner and DiGirolamo, 2000). As has been discussed, these advances have the potential to allow psychophysiological indicators to direct human–system interaction (Farwell and Donchin, 1988). The ability to use sensors to measure the cognitive performance of a user immediately through psychophysiological characteristics, and virtually instantly adapt a system to meet user needs, presents an exciting new paradigm in interactive systems. The introduction of such real-time adaptive aiding offers the prospect of radically altering how humans interact with computer technology. However, one important aspect of such a potential change in the nature of human–system interaction is the inherent difference between open- and closed-loop systems.

Even well-understood, stable open-loop systems will show very different performance under closed-loop operation. A simple example of this effect can be seen when bringing a speaker and a microphone (connected to each other) too close together. A well-known audio feedback effect occurs as the signal from the speaker runs through the microphone, back out of the speaker, back into the microphone, and so on. The resulting feedback loop is (typically) unstable and produces a familiar (and unpleasant) sound. The volume of this sound may grow or decay (corresponding to unstable and stable feedback systems, respectively), depending on whether the microphone to the speaker (which implicitly sets the loop gain in the feedback system). Thus, two perfectly well-behaved open-loop systems (speaker and microphone) may or may not be closed-loop stable, depending on how feedback is applied. A more precise quantitative example of such behavior for an augmented cognition system will be provided later, where it is shown that a stable open-loop system may generate a stable or unstable closed-loop system, depending on how feedback is designed.

Although a great deal about human performance may be understood, the nature of the shift from an open- to a closed-loop system is a unique type of change. As a result, many standard predictable aspects of cognitive and motor performance may operate in drastically different ways in closed-loop systems. A prime candidate for understanding such closed-loop circumstances is through the use of engineering control systems theory. [For a discussion of the pros and cons of various types of models, see Baron et al. (1990).] Control systems theory deals with fundamental properties of systems as described (typically) by mathematical models. It provides a framework and tools for analyzing fundamental system properties, such as performance, noise rejection, and stability, and offers systematic approaches for designing systems with these desired properties.

The idea of applying control theory to humans has some history, with Wiener (1948) widely considered to be the first person to draw parallels between control systems in machines and the organization present within some living systems. However, few attempts have been made to apply control systems theory to human–system interaction (Flach, 1999; Jagacinski and Flach, 2003; Young et al., 2004), and thus this is an exciting area of research where much remains to be done. One notable exception that the current effort draws from is Card et al.’s (1983) Model Human Processor (MHP). The MHP is a human information-processing model consisting of a basic block diagram interconnect model of a human, with an associated estimate of the time taken by each processing stage to process relevant data. For augmented cognition purposes, the three most relevant stages (i.e., blocks)
are probably the perceptual, cognitive, and motor processors. This is illustrated in Figure 2, which shows a human operator piloting a vehicle. In this example, information from the operator’s system display would first pass through the operator’s perceptual (i.e., sensory) processor, being perceived, on average, in about 100 ms (Cheatham and White, 1954; Harter, 1967). Perceived information would then be available to the cognitive processor, which has an average cycle time of 70 ms. The cognitive processor would then make a decision, and that decision would be implemented by the motor processor, which has an average cycle time of 70 ms, with a resulting action on the vehicle controls. Note that these three blocks provide an internal model of the operator’s interaction with the external vehicle displays and controls. This block diagram model not only characterizes the flow of information and commands between the vehicle and operator but also enables us to access the internal state of the operator at various stages in the process. This allows modeling of what an augmented cognition system might have access to (internal to the human; e.g., load on human information processing bottlenecks) and how those data might be used to direct closed-loop human–system interaction.

If one considers a control systems model incorporating the flow of human information processing, the time taken by each block adds time delay to the model. However, it does much more than that. As indicated in the early discussion on bottlenecks, it also implies a certain bandwidth for the system, both in terms of channel capacity and because signals that vary more rapidly than the time constant of the system (i.e., high-frequency signals) do not pass through it. Hence the processing blocks act as low-pass filters, only allowing through signals that are below the system bandwidth. For example, humans do not generally perceive the flicker on a computer monitor because it typically occurs at a frequency (100 Hz) higher than that of the perceptual processor’s bandwidth of only about 10 Hz. As a first attempt at modeling such phenomenon, the effects of time lags in human perceptual, cognitive, and motor processing blocks are considered. This results in a dynamic model of the form shown in Figure 3.

Note that the setup depicted in Figure 3 is a generic dynamic model of any one of the MHP components (perceptual, cognitive, motor) shown in Figure 2 (although the model parameters will be different for each). The dynamic models associated with each MHP component (“first-order lag” and “time delay”) of the block model are given, respectively, in the time domain (i.e., convolution representation) as

\[
y(t) = \frac{1}{\tau} \int_0^t e^{-(t-\gamma)/\tau} u(\gamma) d\gamma
\]

\[
z(t) = y(t - \tau)
\]

Figure 2 Human information-processing model.
for each processing block [with overall input $u(t)$ and output $z(t)$], with the time constant $\tau$ taken from the relevant processing time in the MHP model. The first-order lag models the dynamic relationship between input and output signals, which captures the bandwidth effect described earlier. This is most easily seen using the Laplace transform to transform this model from the time domain to an equivalent frequency-domain representation:

$$Y(s) = G(s)U(s)$$

where the function $G(s)$ is given as

$$G(s) = \frac{1}{1 + s \tau}$$

This is known as the transfer function of the system. [See Phillips and Parr (1999) for an overview of transform methods for signals and systems; see Ogata (2002) for an overview of the application of these techniques to dynamic systems and feedback control.] A key point is that the time-domain convolution operator has been transformed into a simple multiplication operator in the frequency domain. That multiplication operator, $G(s)$, is both complex-valued and frequency-varying. The function $G(s)$ captures the frequency response of the system in both magnitude and phase.

To see this, one can evaluate the transfer function along the imaginary axis, that is, substitute $s = j\omega$ into the model (equivalent to specializing the Laplace transform to a Fourier transform) to yield

$$G(j\omega) = \frac{1}{1 + j\omega \tau} = \frac{1}{\sqrt{1 + (\omega \tau)^2}} e^{-\tan^{-1} \omega \tau}$$

which is the frequency response of the system (with $\omega$ the real-valued frequency). This has the desired low-pass frequency response. Low-frequency (slowly varying) signals pass through almost unattenuated, but higher-frequency (rapidly varying) signals are more and more attenuated until hardly any of the signal passes through the system at all. This variation of the magnitude response with frequency in the first-order lag block is what accounts for the computer monitor effect (i.e., lack of perceiving flicker) described earlier (one could not account for this effect with a time-delay block alone because the frequency response of a pure time delay is flat, i.e., no variation of magnitude with frequency).

Note that this magnitude response comes with an associated phase response. Low-frequency signals pass though this system with almost undistorted phase. However, as frequency increases, the signals start to incur phase lag, which ultimately reaches 90° at high frequency. Phase lag has a destabilizing effect on closed-loop feedback systems, so understanding the relationship between magnitude and phase of different frequency signals as they pass through the system is of crucial importance in designing any feedback control system.

These various steps have provided the separate pieces necessary to build a model of an entire open-loop system. Since transfer functions operate by multiplication, models for the individual blocks can be cascaded. These are linear models and therefore they commute, so the order of cascade can be changed, and hence time delays can be accumulated into a single block if desired. This now provides a quantitative dynamic model for the human as illustrated in Figure 4.

**Figure 3** Block model for each component.

**Figure 4** Dynamic control system model of the human.
Note that, as discussed above, this model captures the gain–phase relationship with frequency, which is crucial if the model is to be used in a feedback control loop.

This model should allow accurate predictions of open-loop performance and other properties of the system to be made. However, it is important to note that this control theory–based model is in a form that will also allow for prediction of how performance and properties are modified when transforming to a closed-loop setup, which is described in the following section.

### 4.1.1 Augmented Cognition Closed-Loop Models

Augmented cognition aims to provide display and information systems that take measurements from OFS gauges, such as those described in Section 2, and use these data to dynamically adapt human–system interaction. The sensor dynamics of any future OFS gauges are still to be determined, so as a starting point such sensors are modeled here as simple first-order lags with a time constant of $\tau = 1$ second. The sensor data would be used to dynamically change inputs to a user by directing instantiation of augmentation strategies, such as those described in Section 3. As an example, consider an application where workload is reduced via the *task delegation* augmentation strategy (Wickens et al., 1998). In such an application, using OFS gauges to detect cognitive overload (e.g., through a EEG-derived index of task engagement) (Prinz et al., 2003), lower-priority tasks would be offloaded to automated agents, with the goal of maintaining users working at their maximum capacity. Such a closed-loop human–system interaction model was implemented in the Matlab/Simulink simulation environment, which is illustrated in Figure 5.

Various pieces of an augmented cognition system can be seen in the model in Figure 5, including the human perceptual, cognitive, and motor processors. Note both the OFS gauge that detects the state of the human user (i.e., cognitive work overload measurement) and the augmentation strategies (i.e., within the PID controller) that will alter the input to the human. The rest of the model contains task inputs to the system, displayed outputs at various points (e.g., actual vs. measured cognitive workload), and a simple model of performance errors resulting from cognitive overload. The feedback loop being closed is now apparent in this simulation model, which drives the need for a systematic control theory approach.

### 4.2 Controller Analysis and Design

Even this simple model has already produced some important findings. In particular, one major finding from initial efforts with the model is to show how dynamic instability can result from introducing feedback within the system. That is to say that rapid detection of cognitive state under high workload might result in input being removed, which would reduce workload and hence information would be added, which would once again result in high workload, and the cycle repeats. This simple illustration indicates how users might find their display cycling rapidly through cluttered and decluttered states as a result of changes detected in workload. Control theory offers a means to remove such instability and optimize performance.

Figure 6 shows results from three simulations of a task overload situation. The input to each of these simulations is the same: Initially, the user is fully loaded (and making no errors), and then a step increase in workload is introduced 1 second into the simulation. This results in task overload from that point on, with subsequent performance errors. Note that each of these simulations uses the same system model, so the only difference is how (or if) the feedback control (i.e., augmentation strategy) is applied.

![Figure 5](attachment:figure5.png)  
**Figure 5** Matlab/Simulink model of closed-loop augmented cognition human–system interaction.
Figure 6  Simulation results for an augmented cognition closed-loop dynamic model.

Starting from the left, the first plot of Figure 6 shows the resulting performance errors for an open-loop simulation (i.e., with the augmented cognition system disabled). As the workload of the task increases, the plot shows how the number of errors quickly rises to a certain level and stays there. The next panel shows a poorly designed augmented cognition system. This system utilizes simple proportional control; that is, the control action $c(t)$ that reduces task workload to the user is just directly proportional to measured overload $m(t)$. Thus, the controller is of the form

$$c(t) = K m(t)$$

and the control designer simply chooses the proportionality constant or controller gain $K$, which determines when an augmentation strategy (i.e., task delegation in this case) is to be implemented. High-gain controllers, with large $K$ values, use a high-magnitude feedback signal that tries aggressively to drive the control loop to the desired point (for fast or high performance). If $K$ is chosen too aggressively, however, the closed-loop system will approach (or even exceed) stability margins. In this example, the gain $K$ is chosen poorly, resulting in instability of the type described above, with the input being reduced rapidly and then increased, resulting in highly fluctuating performance from the user. Note that the precise values of $K$ that drive the system into instability depend on the specific problem (and can be predicted accurately with control theory methods), but they can certainly occur at plausible real-world values (in this example, $K = 2.8$).

Proportional control is what people often think of when they consider feedback. A simple version is the cruise control in a car, which moves the gas pedal in a manner proportional to the difference between the desired and actual speeds. However, this simple control strategy can deliver only limited performance improvements, even when designed correctly. For instance, one could never get steady-state errors down to zero with this type of control. This approach is limited because it utilizes the same gain for all frequencies (and hence all signals), so one does not have sufficient degrees of freedom to exploit any trade-offs in the design. A very common type of controller used in engineering applications is the proportional–integral–derivative (PID) controller. This generates a corrective action from a measurement of the form

$$c(t) = K_p m(t) + K_I \int_0^t m(\tau) d\tau + K_D \frac{dm(t)}{dt}$$

There are now three constants to be chosen (designed), $K_p$, $K_I$, and $K_D$, which correspond, respectively, to the amounts of proportional, integral, and derivative feedback used in the closed loop. Note that the integral action effectively includes memory and thus allows better compensation at low frequency and hence improved steady-state performance. The derivative action essentially includes anticipation, which allows for improved high-frequency performance, resulting in better transient response and improved stability properties. The overall controller has frequency-varying gain, which allows design trade-offs to be exploited more properly. The right panel of Figure 6 shows a functional closed-loop system using a well-designed PID controller to deliver closed-loop stability and good performance. It is clear that even maximum errors never reach the level of the open-loop (automation-free) system and that they quickly drop to minimal levels (asymptotically approaching zero) without any undesirable oscillatory transient response.

4.2.1 Human Dynamics and Achievable Performance

The first benefit of the modeling approach described above is that it provides some proof of concept for the augmented cognition concept: namely, to show precisely how an integrated system of OFS gauges, augmentation strategies, and robust controllers can combine to augment performance. The caveat from this work, however, is to note that such systems need to be designed carefully, with a systematic control theory
approach rather than simple heuristic tuning, else augmented cognition may fail to fulfill its potential. Fortunately, systematic modeling can offer assistance in terms of determining the nature of information required and parameters necessary for driving specific OFS gauges. The types of questions that could be addressed by this type of analysis include:

- What time constant/bandwidth is necessary for a particular OFS gauge to have a significant useful effect (i.e., how fast)?
- What resolution is required of the OFS gauge (i.e., how accurate)?
- How much noise can reasonably be tolerated on any given measurement?
- What would additional measurements/gauges offer?
- What performance level could be achieved (given the above)?

These questions should be addressed in future work in the area of modeling and analyses. Note that both qualitative and quantitative analyses can be carried out, and both have their uses (e.g., qualitative analysis might steer one toward a particular technology, whereas quantitative analysis might allow one to design and implement it accurately). Note also that specific scenarios can be carried out in a simulation, which would allow one to test out certain strategies repeatedly and reliably before going to the expense of constructing an experimental setup, including low-probability events that might not occur in an experimental setting. Furthermore, control theory includes powerful analysis tools that go well beyond simple simulation to address fundamental trade-offs and limitations inherent in any feedback loop (Doyle et al., 1992).

Ultimately, the modeling strategies described in this section would aim to predict the impact on human–system performance of various augmentation strategies for changing how information is provided to a user. In addition, they have the potential to highlight areas that would receive particular benefit from such augmentation. Thus, overall, this work can provide the basis for future systematic closed-loop analysis and controller design, bringing to bear powerful tools from engineering control theory. The power of such analysis tools is demonstrated in the next section.

### 4.2.2 Individual Differences and Robustness Analysis

Preliminary robustness analysis was conducted on the closed-loop system with PID controller modeled in Figure 5. The theory of robust control deals with systems subject to uncertainty such as any closed-loop augmented cognition system would be subject to due to individual differences (among other reasons, as noted above). Control theory provides a means of examining what performance on such a system will be, rather than just an idealized simulation. It also allows for examination of variation between users, since relevant parameters in the model can be varied (e.g., speed of the MHP processors) to determine to what extent a given control scheme is robust against such variations.

The theoretical tools used here to model individual differences were based on the structured singular value (SSV), or $\mu$, and its extensions to handle real parametric uncertainty (Young, 2001). The idea is that one first has to use linear fractional transformations (LFTs) to rearrange the problem into canonical $M - \Delta$ form, as illustrated in Figure 7. Here $M(s)$ collects all the known dynamics of the (closed-loop) system, and $\Delta$ is a (block) diagonal structured perturbation, which in the case of individual differences analysis will consist of real parametric uncertainty representing variation in the parameters of the model. Thus, this approach handles LFT (block diagram) perturbations rather than handing perturbed coefficients directly in a (transfer function) model, but this apparent limitation is readily overcome, as we illustrate below.

The individual differences analysis considered variations in two time constants (i.e., speed of the perceptual and cognitive processors). These could arise due to variations among users, but could also be introduced through inaccuracies in the modeling approach. To realize this analysis, these variations were cast as a block diagram perturbation. This can be done by noting the interconnect in Figure 8, which shows an example of rearranging parametric uncertainty as an LFT (block diagram).

Mathematically straightforward block diagram calculations now reveal that the transfer function in Figure 8 is represented by

$$\frac{1}{1 + s(\tau + \Delta \tau)}$$

![Figure 7](https://example.com/fig7.png)  
**Figure 7**  
Canonical form for SSV analysis.

![Figure 8](https://example.com/fig8.png)  
**Figure 8**  
Variation in time constant as an LFT perturbation.
so that the block diagram perturbation in Figure 8 actually becomes a perturbed coefficient in the transfer function model, and to be specific it represents a perturbation in the time constant of the first-order lag model.

This approach was applied to the closed-loop augmented cognition system model represented in Figure 5, considering a parametric variation in the time constants of the first-order lag models of the perceptual and cognitive processor blocks. Note that the motor processor block is not in the feedback loop in this scenario, so it does not affect stability and hence was not included in the robustness analysis presented here. The LFT and \( \mu \) analysis machinery could then be applied to this block diagram. The mathematics of this approach is quite involved and uses computational complexity theory, complex analysis, and linear algebra among others (Young, 2001). Space constraints prevent going into any kind of detailed explanation here, but the end result of this analysis was to give a parameter range over which (robust) stability is guaranteed. This means that no parameter combination in the allowed range can cause instability. For example, in this case one could guarantee that no person with a combination of processing time constants for the perceptual and cognitive processors in the ranges specified would cause the closed-loop augmented cognition system model represented in Figure 5 to go unstable. It is important to note the power of this guarantee, because one cannot get such guarantees from any amount of exhaustive simulation or testing (it is always possible that a parameter combination is missed, which causes a problem no matter how many variations are tried).

The results of this analysis showed that both the perceptual and cognitive processor time constants could be reduced to very small numbers (practically all the way down to zero), indicating that faster user response than predicted was no problem. The upper limits for the perceptual and cognitive processor time constants were found to be 3.7 and 2.6 seconds, respectively. Thus, it is possible for the system to go unstable with slower users. However, the degree of robustness afforded by a PID controller is huge. Specifically, in this example the time constants are a factor of more than 37 times greater than those nominally assumed (e.g., the MHP’s “slowman” to “fastman” range for the perceptual processor is 150 ms and for the cognitive processor is 145 ms) (Card et al., 1983), meaning that tremendous variability in the perceptual and cognitive processor time constants can be tolerated between users and tasks. These robustness analysis results were also confirmed by a simulation model, which showed stable behavior for all parameter variations in the range allowed but which could be driven unstable by parameter combinations outside these ranges (Young et al., 2004). This individual-differences analysis serves to illustrate what could be done when an augmented cognition system is coupled with a systematic control theoretic approach.

### 4.2.3 Robust Controller Synthesis

The control theory methods reviewed above can facilitate the design of high-performance closed-loop systems, even for systems whose dynamics are only partially known (Packard and Doyle, 1993; Zhou et al., 1996; Young, 2001). This is not achieved by optimizing nominal performance measures as in classical optimal control techniques such as linear quadratic Gaussian/linear quadratic regulator control (Ogata, 2002). Rather, these new approaches attempt to optimize robust performance measures utilizing techniques such as \( \mu \)-synthesis (Packard and Doyle, 1993). In this way, systems can be designed which are insensitive (or robust) to variations in the system that are naturally occurring but hard to predict a priori (e.g., differences between users). The mathematical machinery underlying such techniques is quite involved, and the associated optimization problems can be nonconvex and even NP-hard. At first sight, such problems may appear to be intractable, and indeed, global minima usually cannot be guaranteed. Nevertheless, practical computation schemes have been developed using approximation schemes such as upper and lower bounds. These schemes are capable of finding very good approximate solutions in a reasonable amount of time. Moreover, there are numerically efficient implementations available of the associated algorithms, usually in convenient Matlab form (Balas et al., 1991), so such designs can readily be carried out (with the appropriate software) in a reasonable time using current computer hardware.

All this adds up to the fact that developers of augmented cognition systems have at their disposal a number of powerful tools for robust controller analysis and synthesis. These theoretical techniques offer the potential of safely optimizing performance in an augmented cognition system while maintaining guaranteed closed-loop stability.

### 5 APPLICATION DOMAINS

As with any scientific discovery or technical innovation, there are multiple paths upon which technology components will advance. Identifying specific applications at the dawn of a significant advance in our understanding of a field of study is problematic in that the assumptions on which hypothesized applications are based are very likely to be flawed. The assumptions that must be made (and that are likely to be wrong) include:

- Who will take advantage of emergent technology components?
- What components of the emergent technology will ultimately prove most useful and robust?
- When will the various components of the emergent technology be validated sufficiently for incorporation into real-world systems?
- Where will the emergent technology components be found to be most useful?
- Why will the emergent technology components be seen as beneficial?
• How will the emergent technology components be used?

Given the challenges inherent in answering these questions, a good starting point is to identify potential general application domains and then extract examples from these domains in hopes of describing potential uses of emergent technology components. The general application domains likely to be affected most by augmented cognition technology components include operational domains such as truck driving and power plant operation that would benefit from real-time cognitive readiness and assessment capabilities; educational domains, such as a scenario-based training system that can adapt in real time to trainee performance, as assessed by both overt behavior and cognitive state (as captured by OFS gauges); and clinical domains such as medical applications, where for example, the real-time attention processes of children with attention-deficit hyperactivity disorder (ADHD) are monitored and reinforcement interventions for “paying attention” are employed. The value of augmented cognition to the operational, educational, and clinical application domains should be noted; but the specific examples reviewed below should not be assumed to be predictions of actual application areas.

Another way to attempt to glimpse the future of applications is to examine early prototypes that incorporate the underlying science and technology of interest. The practice of postulating potential futures is common when one has demonstrated technology; fortunately, it is not unknown at the beginnings of basic science or the technology development process either. When he was president and CEO of Bellcore, George Heilmeier (1999), insisted that before starting off on any scientific endeavor or technology development project, the following questions be addressed:

• What are you trying to do? Articulate your objectives using absolutely no jargon.
• How is it done today, and what are the limits of current practice?
• What’s new in your approach, and why do you think it will be successful?
• Who cares? If you’re successful, what difference will it make?
• What are the risks and payoffs?
• How much will it cost? How long will it take?
• What are the midterm and final “exams” to check for success?

Conveniently, augmented cognition as a field of study was born in part through significant investment by DARPA in the Improving Warfighter Information Intake Under Stress Program, where advanced thought toward eventual application was mandatory. One can look to the applications and prototypes generated from this technology initiative to understand potential applications more clearly and as a guide for attempting to postulate potential futures and to answer the two sets of questions presented at the beginning of this section.

DARPA’s Improving Warfighter Information Intake Under Stress Program focused initially on challenges and opportunities associated with real-time monitoring of cognitive states with psychophysiological sensors. The ultimate goal was to demonstrate the use of this underlying technology to increase human information processing substantially in four operational military applications. The applications included applying augmented cognition component technologies to a military driving platform, an unmanned vehicle interface platform, command and control platforms, and a dismounted soldier platform. These operationally focused applications provided proof of concepts of benefits of the emergent component technologies; they also provided testbeds for refinement and testing of new methods of incorporating these technologies (the results of many of these efforts are available in the Proceedings of the First International Conference on Augmented Cognition, 2005).

There are also many potential nonmilitary operational applications of augmented cognition. The commercial sector is beginning to consider augmented cognition technologies for incorporation into operational systems. For instance, in 2002, IEEE held their seventh Conference on Human Factors and Power Plants, which focuses on new trends, and dedicated an entire session to reviewing state-of-the-art augmented cognition technologies and assessing the maturity level of these technologies. Components of interest included methods to detect and measure a power plant operator’s workload, strategies to facilitate multitasking in multimodal environments, and support for intelligent interruption and context recovery. Furthermore, NASA is supporting the development of personal monitoring capabilities to support both intelligent system automation and human performance aids (Prinz et al., 2000, 2003). The demanding environment of complex missions and associated dynamic information processing demands require NASA to seek enhanced system capabilities and maximized human performance. Leveraging augmented cognition, NASA hopes to increase the ability of a single human to perform numerous tasks while maintaining the strictest margins of safety.

To date, nonoperational domains such as training and clinical domains have been the least affected by augmented cognition technology. However, it is in the training area that the likelihood of successful application development would seem most promising. Joseph Cohn and Amy Kruse of DARPA’s Improving Warfighter Information Intake Under Stress Program have suggested that the development of an augmented cognition system that could turn novices rapidly into experts would revolutionize the training community. Such a system would identify a person’s current level of expertise and would allow the person to be guided rapidly to heightened levels of sustained performance in a context-independent fashion. Additionally, a person’s cognitive performance during training could be periodically or continuously assessed to ensure that their training was proceeding appropriately. Cohn and Kruse’s research seeks to develop a multilevel approach to training that capitalizes on being able to
observe patterns at both the overt behavioral level and at a deeper structure neuro-imaging level. They point to research results from the neurosciences which indicate that activation of specific brain regions is correlated to novice and expert behaviors, giving evidence for a neural correlate to observed expert behavior. Additionally, they suggest that changes in these structures can be assessed over time to track progression toward “expert” neural activation. An application that could characterize expert performance, identify where in the novice–expert continuum a trainee’s performance lies, and then mold the trainee’s patterns to more closely reflect an expert’s would revolutionize training.

Finally, within the clinical domain one can imagine that by leveraging augmented cognition technology, clinicians would be better able to diagnose, evaluate, and mitigate fatigue and learning decrements. Developments of emergent augmented cognition technology components have been least aligned with such potential clinical applications. However, successes in the operational and training domains will probably accelerate application developments in the clinical domain. Additionally, continued investment in real-time diagnostic tools by the National Institutes of Health will probably create a marketplace for new medical tools and associated applications that leverage augmented cognition technology.

In summary, applications of augmented cognition are in their infancy. Examples of possible applications of underlying technology components can readily be imagined; however, successful instantiation and usefulness of system applications can only be guessed at. There is significant evidence to suggest that the technology components are ready for insertion into mature applications, and that the operational, educational, and clinical domains have capability gaps that call for technology solutions offered by the field of augmented cognition. Thus, although mature augmented cognition science and technology components have now embarked on the path toward application, the only certainty along this journey is that the applications developed will be like no others that have come before them.

6 CONCLUSIONS

Augmented cognition seeks to achieve Licklider’s (1960) vision of human–computer symbiosis, where human brains and computing machines are tightly coupled, thereby achieving a partnership that surpasses the information-handling capacity of either entity alone. Such improvement in human–system capability is clearly a worthy goal, whether the context is clinical restoration of function, educational applications, market-based improvements in worker efficiency, or warfighting superiority. Augmented cognition is an attempt to realize a revolutionary paradigm shift in interactive computing, not by optimizing the friendliness of connections between human and computer but by achieving a symbiotic dyad of silicon- and carbon-based enterprises.

ACKNOWLEDGMENTS

This material is based on work supported in part by DARPA’s Improving Warfighter Information Intake Under Stress Program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views or the endorsement of DARPA.

REFERENCES


