Augmenting Human Cognition with Adaptive Augmented Reality

Jayfus T. Doswell[1] and Anna Skinner[2]

[1] Juxtopia, Baltimore, MD, United States of America
[2] AnthroTronix. Silver Spring, MD, Unites States of America
jayfus@juxtopia.com, anna.skinner@atinc.com

Abstract. Wearable Augmented Reality (AR) combines research in AR, mobile/ubiquitous computing, and human ergonomics in which a video or optical see-through head mounted display (HMD) facilitates multi-modal delivery of contextually relevant and computer generated visual and auditory data over a physical, real-world environment. Wearable AR has the capability of delivering on-demand assistance and training across a variety of domains. A primary challenge presented by such advanced HCI technologies is the development of scientifically-grounded methods for identifying appropriate information presentation, user input, and feedback modalities in order to optimize performance and mitigate cognitive overload. A proposed framework and research methodology are described to support instantiation of physiologically-driven, adaptive AR to assess and contextually adapt to an individual's environmental and cognitive state in real time. Additionally a use case within the medical domain is presented, and future research is discussed.

Keywords: augmented cognition, augmented reality, context-awareness, cognitive display, mobile computing, mixed-reality, wearable computing.

1 Background

1.1 Wearable Augmented Reality

Augmented Reality (AR) falls within Milgram's mixed reality continuum as illustrated in Figure 1.

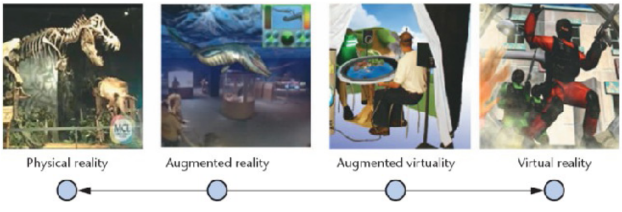

Fig. 1. Milgram's mixed reality continuum

D.D. Schmorrow and C.M. Fidopiastis (Eds.): AC 2014, LNAI 8534, pp. 104–113, 2014.

In AR, digital objects are added to the real environment. In augmented virtuality, real objects are added to virtual ones. In virtual environments (or virtual reality), the surrounding environment is completely digital [25].

While AR technologies include auditory, haptic, olfactory, and even gustatory displays, the vast majority of AR displays emphasize the visual modality. Within the visual modality, AR displays may be head-worn, handheld, or spatial (i.e., placed in the environment), including screen-based video see-through displays, optical see-through displays, and projective displays [37]. Wearable AR combines research in AR and mobile/pervasive computing in which a wearable displays, primarily head-mounted, and increasingly small computing devices facilitate wireless communication and context-aware digital information display [1][10][11][38]. Such displays facilitate multi-modal delivery of contextually relevant and computer generated information data over a physical, real-world environment.

The Sword of Damocles, developed in 1968 by Ivan Sutherland with the assistance of his student, Bob Sproull, is widely considered to be the first virtual reality (VR) and AR head-mounted display (HMD) system [37]. Since that time, AR technologies have become increasingly powerful and affordable, as evidenced by the recent release of Google's Glass consumer optical HMD technology. Such head-worn Human Computer Interaction (HCI) technologies have the potential for providing a natural method to augment human cognition while wearers interact with the natural environment. More specifically, wearable AR has capabilities to deliver context-aware assistance, combining and presenting multi-modal perceptual cues such as animation, graphics, text, video, voice, and tactile feedback via complementary wearable computing devices (e.g., digital gloves).

AR technologies are becoming prevalent across a wide variety of domains including military, commercial, education and training, entertainment, and medical applications. For example, through decades of empirical research, wearable AR has demonstrated the capability of delivering on-demand assistance and training to medical personnel on a variety of medical tasks ranging from emergency medical first response and surgery to combat casualty care and public health relief efforts. However, despite many advancements recently demonstrated by wearable AR systems, there are several research gaps that must be addressed in order for wearable AR to achieve an adaptive and wearable HCI model that augments human cognition, and consequently improves human performance. A primary limitation across current AR technologies is the potential for presented information to obscure critical environmental cues or to distract, disrupt, or overload the user [37].

1.2 Cognitive Overload and Information Display

A primary challenge presented by such advanced HCI technologies is the development of scientifically-grounded methods for identifying appropriate information presentation, user input, and feedback modalities in order to optimize performance and mitigate cognitive overload. Such modality selection methodologies must be dynamic, providing the capability to adapt interaction configurations to accommodate various operational and environmental conditions; as well as user cognitive states, which change over time in response to task demands and factors such as sleep, nutrition, stress, and even time of day. Ideally, interaction technologies should be capable of adapting interaction modalities in real-time in response to task, environmental, and user psychophysiological states such as cognitive load.

Cognitive overload may be best described by Cognitive Load Theory (CLT), which is an information processing theory used to explain the human limits of working memory based on current knowledge of human cognitive architecture. Cognitive architecture refers to the concept of our minds having structures such as working memory, long term memory, and schemas [34]. CLT may be summarized as follows:

1. Working Memory can only handle, on average, seven (plus or minus 2) disconnected items at once [24].
2. Overload occurs when Working Memory is forced to process a significant amount of information too rapidly.
3. Long Term Memory is virtually unlimited and assists Working Memory.
4. Schemas are memory templates coded into Long Term Memory by Working Memory.
5. Working Memory is overloaded when its ability to build a schema is compromised.
6. If Working Memory has capacity left over, it can access information from long term memory in powerful ways.
7. Automation (doing something without conscious thought) results from well developed Schemas due to Working Memory's interaction with Long Term Memory. Well developed schemas come with repeated effort and effective practice. [20].

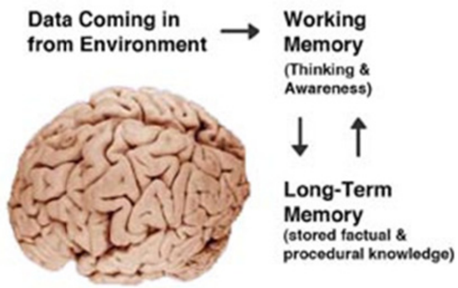

Fig. 2. Working Memory model

Furthermore, information retrieval from long term memory can be impacted based on external (i.e., fast moving or disruptive objects) or internal (i.e., physiological or emotional) stimuli, which may significantly impact task performance.

Based on CLT-based constraints of information storage and retrieval to and from long term memory, respectively, an effective contextually intelligent information display may be a useful intervention, especially with the use of pictorial mnemonic systems. Previous research demonstrates that the use of memorization techniques (i.e., mnemonic strategies) has resulted in improvements in humans' ability to recall learned information [7][9][18]. Mnemonic strategies are proven systematic procedures for enhancing memory recall [22], and are used to facilitate the acquisition of factual information because they assist in the memory encoding process, either by providing familiar connections or by creating new connections between to-be-remembered information and the learner [21].

According to Bellezza [3], memory experts learn to create mental pictures that endure in the mental space. A medical pictorial mnemonic system has the capability to assist the recall of procedural steps in a single pictorial form, especially if depicted as intuitively formed symbols that are easily and immediately recognizable to the user. A study by Estrada et al. [12], using a pictorial mnemonic system for recalling aviation emergency procedures found that the system facilitated the recall of uncommon, unfamiliar terms and phrases in a population to a level comparable to that of highly-experienced pilots in just one week. These findings highlight the potential for such a mnemonic strategy to aid in the encoding of information into long-term memory. This encoding and catalytic recall method involves "chunking" multiple pieces of information into a picture format, resulting in decreased human cognitive overload and accelerated human decision making; thereby augmenting human task completion performance.

1.3 Traditional Use of AR to Support Cognitive Load Reduction

AR technologies have traditionally been employed to reduce task switching and associated working memory load and attentional demands. Neumann & Majoros [28] proposed that AR provides a complement to human cognitive processes via integrated information access, error potential reduction, enhanced motivation, and the provision of concurrent training and performance. Bonanni, Lee, & Selker [6] provided evidence that multimodal augmented interactions can enhance procedural task performance, and suggested that providing visual cues decreases cognitive load because memory is a more complex process than pop-out in a visual search based on cueing and search principles from attention theory. However, metrics of cognitive load were not directly assessed.

Kim & Dey [19] demonstrated the effective use of context-sensitive information and a simulated AR representation combined to minimize the cognitive load, using user performance as a proxy for cognitive load assessment, in translating between virtual information spaces and the real world. Similarly, Tang, Owen, Biocca, & Mou [35] demonstrated improved task performance using AR and suggested that AR systems can reduce mental workload on assembly tasks. This study also addressed the issue of attention tunneling, indicating that AR cueing has the potential to overwhelm the user's attention, reducing performance by causing distraction from important relevant cues of the physical environment. This phenomenon has yet to be explored using objective measures of user attention and associated cognitive workload. Multiple AR studies have employed the NASA Task Load Index (NASA-TLX) to assess mental workload [4][5][23][36]. However, the TLX provides only subjective ratings of mental workload and is not assessing in real time during task performance, relying on the user's memory of perceived task demands.

Nagao [26] proposed that psychophysiological signals such as electroencephalogram (EEG), electrocardiogram (ECG), and galvanic skin response (GSR) could be used to support personalization of AR interaction. However, subsequent R&D has not fully explored this approach, with the majority of related research focusing on brain-computer interface (BCI) applications. For example, Navarro [27] outlined a framework in which physiological measures such as EEG, heart rate, and body temperature could be used to add intelligence to a wearable BCI predicting individualized, programmable user behaviors, both offline and online. This functionality was predicted to support faster user response rates, and to support increased freedom and flexibility. This framework also included the option to selectively vary the weights from the various physiological inputs according to contextual factors. Navarro also proposed that the incorporation of technologies such as AR and multimodal personalized BCI techniques could be used to increase accuracy in dynamic environments, and that this functionality, combined with the adoption of wireless technologies, would support instantiation of this paradigm within increasingly mobile and dynamic contexts. Scherer et al. [33] proposed the use of VR and AR in combination with hierarchical BCIs and learning models in order to increase BCI usability and interaction with physical and virtual worlds. Specifically, the proposed approach leverages the benefits of two paradigms of event related potential (ERP) stimuli: environmental stimuli and stimuli generated by mental imagery. The goal of this approach was to combine environmental and user-generated inputs within a hierarchical BCI system capable of adapting to individual users.

A primary multi-modal HCI gap that must be addressed in order for wearable AR to improve human working/long term memory involves real-time assessment of cognitive workload and real-time adaptive information presentation to mitigate cognitive overload.

2 An Augmented Cognition Framework for Adaptive AR

AugCog R&D is grounded in a multi-disciplinary scientific approach to addressing issues of human-technology interaction through a blending of cognitive science, human factors, and operational neuroscience. While much of the research in this field has focused on the use of physiological assessment metrics to drive real-time adaptive HCI, this paradigm has yet to be applied and validated within the context of AR.

A framework and research methodology is proposed to support instantiation of adaptive AR to reduce cognitive load and contextually adapt to an individual's environment or cognitive/physiological state (e.g., stress) using AugCog principles. Specifically, we propose real-time monitoring and assessment of neurophysiological measures capable of indicating user cognitive workload, and specifically differentiating between verbal working memory and spatial working memory. Within this framework, indices of cognitive workload would drive a closed-loop HCI adaptive AR interface, reducing information presentation to the user during periods of high workload and increasing information as appropriate during periods of low workload. The proposed system would further differentiate between verbal working memory overload and spatial working memory overload, adapting information presentation as necessary to avoid overtaxing one working memory system.

Additionally, a key component of the proposed framework is the integration of neurophysiological metrics with contextually-relevant environmental and interaction/performance-based measures to optimize modality selection in real time. The goal of such a methodology is to extend human cognitive capabilities while remotely interoperating with a network of federated computing services.

3 Medical Procedure Assistance Use Case

The following provides a use case scenario selected from the medical domain involving cognitive, psychomotor, and perceptual skills within a dynamic operational environment.

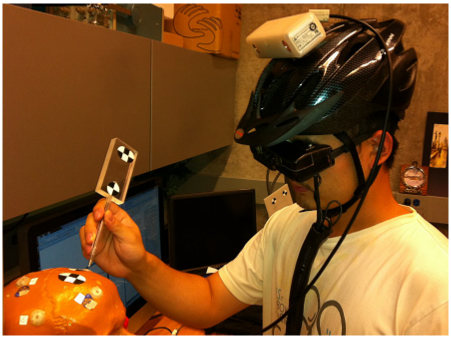

Fig. 3. Medical AR system

Recent wearable AR medical research from Azimi, Doswell, and Kazanzides [2] demonstrated the use of context-aware wearable AR for use in surgical assistance with future implications of augmenting human perceptual cues via multi-modal cognitive cues.

Surgical resection is one of the most common treatments for brain tumors. The treatment goal is to remove as much of the tumor as possible while sparing the healthy tissue. Image guidance using preoperative Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) is frequently used because it can more clearly differentiate diseased tissue from healthy tissue. Most image guidance devices contain special markers that can be easily detected by a computer tracking system. Registering the tracker coordinate system to the preoperative image coordinate system gives the surgeon “x-ray vision”. However, this modality can be challenging for the surgeon to effectively use a navigation system because the presented information is not physically co-located with the operative visual display field, requiring the surgeon to look at a computer monitor rather than at the patient. This is especially awkward when the surgeon wishes to move an instrument within the patient while observing the display. Such ergonomic issues may increase operating times, fatigue, and the risk of errors. Furthermore, most navigation systems employ optical tracking due to its high accuracy, but this requires line-of-sight between the cameras and the markers in the operative field, which can be difficult to maintain during the surgery.

After researchers observed surgeries, particularly neurosurgeries, and discussions with surgeons, they identified a need to overlay a tumor margin (boundary) on the surgeon's view of the anatomy, which served as a pictorial mnemonic; triggering memory recall and visual landmarks to cognitively assist the surgeon in accurate completion of a psychomotor procedure. It was also desired to correctly track and align the distal end of the surgical instruments with the preoperative medical images. The aim of the research was to investigate the feasibility of implementing a head-mounted tracking system with an AR environment to provide the surgeon with visualization of both the tumor margin and the surgical instrument in order to create a more accurate and natural overlay of the affected tissue versus healthy tissue, and hence, provide a more intuitive human-computer interface [2].

The resulting wearable AR allows the surgeon to see the precise boundaries of the tumor for neurosurgical procedures, while at the same time providing contextual overlay of the surgical tools intraoperatively, which are displayed on optical see-through goggles worn by the surgeon. This makes it feasible for AR, as the overlay provides the most pertinent information without unduly cluttering the visual field. It provides the benefits of navigation, visualization, and all of the capabilities of the existing modalities and is expected to be comfortable and intuitive for the surgeon.

The majority of related research has focused on AR visualization with HMDs, usually adopting video see-through designs. Many of these systems have integrated one or more on-board camera subsystems to help determine head pose [13][16][31] and some have added inertial sensing to improve this estimate via sensor fusion [8]. None of these systems, however, attempt to provide a complete tracking system and continue to rely on external trackers.

Combining the aforementioned wearable AR intervention with multi-modal sensors that track both dynamically changing internal and external stimuli, surgeons may visually monitor on the AR display feedback based on their cognitive fatigue and stress level as well as performance based multi-model cues including, but not limited to multimedia mnemonics that augment's the surgeon capability to better perform the surgery.

4 Conclusions and Future Work

Wearable AR has the capabilities to deliver context-aware assistance as multi-modal perceptual cues combining animation, graphics, text, video, and voice as well as tactile feedback to complementary wearable computer devices (e.g., digital gloves). A primary challenge presented by such advanced HCI technologies is the development of scientifically-grounded methods for identifying appropriate information presentation, user input, and feedback modalities in order to optimize performance and mitigate cognitive overload. Such modality selection methodologies must be dynamic, providing the capability to adapt interaction configurations to accommodate various operational and environmental conditions, as well as user cognitive states, which change over time in response to task demands and factors such as sleep, nutrition, stress, and even time of day. Ideally, interaction technologies should be capable of adapting interaction modalities in real-time in response to task, environmental, and user psychophysiological states. The field of Augmented Cognition (AugCog) has demonstrated the technical feasibility of utilizing psychophysiological measures to support real-time cognitive state assessment and HCI reconfiguration to mitigate cognitive bottlenecks within a wide variety of operational environments.

Researchers will continue to explore how to integrate a dynamically changing sensor/software framework into a context-aware AR platform [1][2][10][11] to augment human perceptual capabilities and improve human performance.

References

1. Adeboye, O.O., Ochieng, E., Gillis, D., Doswell, J., Sofolahan, A., Dickens, C.: Context-Aware Augmented Reality System: A Medical First Responder Situational Awareness Interface. In: Poster Session. Medicine Meets Virtual Reality (MMVR) Conference. Newport Beach, California (2010)
2. Azimi, E., Doswell, J., Kazanzides, P.: Augmented reality goggles with an integrated tracking system for navigation in neurosurgery. In: Conference: Virtual Reality, IEEE Annual International Symposium, pp. 123–124 (2012)
3. Bellezza, F.S.: The Mind s Eye in Expert Memorizers Descriptions of Remembering. Metaphor and Symbolic Activity 7(3&4), 119–133 (1992)
4. Biocca, F., Owen, C., Tang, A., Bohil, C.: Attention issues in spatial information systems: directing mobile users' visual attention using augmented reality. Journal of Management Information Systems 23(4), 163–184 (2007)
5. Biocca, F., Tang, A., Owen, C., Xiao, F.: Attention funnel: omnidirectional 3D cursor for mobile augmented reality platforms. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1115–1122. ACM (April 2006)
6. Bonanni, L., Lee, C.H., Selker, T.: Attention-based design of augmented reality interfaces. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems, pp. 1228–1231. ACM (April 2005)
7. Carney, R.N., Levin, J.R.: Promoting Higher-Order Learning Benefits by Building Lower-Order Mnemonic Connections. Applied Cognitive Psychology 17, 563–575 (2003)
8. Chai, L., et al.: Three-Dimensional Motion and Structure Estimation Using Inertial Sensors and Computer Vision for Augmented Reality. Presence 11(5), 474–492 (2002)
9. Cox, B.D.: Children s Use of Mnemonic Strategies: Variability in Response to Metamemory Training. The Journal of Genetic Psychology 155(4), 423–442 (2001)
10. Doswell, J., Sampson, O., Jeffrey, M.: WASA MIST Wearable Assistant for Real-Time Emergency Care Providers. In: 15th Annual International Meeting of the American Telemedicine Association (2010)
11. Doswell, J., Anderson, M., Wilson, K., Kazanzides, P., et al.: Wearable Augmented Reality for Medical First Response & Situational Awareness. In: Medicine Meets Virtual Reality (MMVR19) NextGen Conference. Newport Beach, California (2012)
12. Estrada, A., Keeley, J.A., Leduc, P.A., Bass, J.M., Rouse, T.N., Ramiccio, J.G., Rowe, T.L.: A novel approach in facilitating aviation emergency procedure learning and recall through an intuitive pictorial system. U.S. Army Aeromedical Research Laboratory Technical Report, No. 2007-07 (2007)

13. Fuchs, H., Livingston, M.A., Raskar, R., Keller, K., State, A., Crawford, J.R., Rademacher, P., Drake, S.H., Meyer, A.A.: Augmented reality visualization for laparoscopic surgery. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 934–943. Springer, Heidelberg (1998)
14. Höllerer, T., Feiner, S.: Mobile Augmented Reality. In: Karimi, H., Hammad, A. (eds.) Telegeoinformatics: Location-Based Computing and Services, Taylor and Francis Books Ltd., London (2004)
15. Hollerer, T., Feiner, S., Hallawa, D., et al.: User Interface Management Techniques for Collaborative Mobile Augmented Reality. Computers and Graphics 25(5), 799–810 (2001), `http://citeseer.ist.psu.edu/hollerer01user.html`
16. Hoff, W., Vincent, T.: Analysis of head pose accuracy in augmented reality. IEEE Trans. Visualization and Comp. Graphics 6 (2000)
17. Hughes, C., Stapleton, C.B., et al.: Mixed Reality in Education Entertainment, and Training. IEEE Computer Graphics and Applications, 24–30 (November/December 2005)
18. Kleinheksel, K.A., Summy, S.E.: Enhancing Student Learning and Social Behavior Through Mnemonic Strategies. TEACHING Exceptional Children 36(2), 30–35 (2003)
19. Kim, S., Dey, A.K.: Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 133–142. ACM (April 2009)
20. Learning and ID. Cognitive Load Theory, `http://im404504.wikidot.com/cognitive-learning-theory`
21. Levin, M.E., Levin, J.R.: Scientific mnemonomies: Methods for maximizing more than memory. American Educational Research Journal 27, 301–321 (1990)
22. Mastropieri, M.A., Scruggs, T.E.: Enhancing School Success with Mnemonic Strategies. Intervention in School & Clinic 33(4) (1998)
23. Medenica, Z., Kun, A.L., Paek, T., Palinko, O.: Augmented reality vs. street views: a driving simulator study comparing two emerging navigation aids. In: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, pp. 265–274. ACM (August 2011)
24. Miller, G.A.: The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review 63, 81–97 (1956)
25. Milgram, P., Kishino, A.F.: Taxonomy of Mixed RealityVisual Displays. IEICE Trans. Information and Systems E77-D(12), 1321–1329 (1994)
26. Nagao, K.: Agent augmented reality: Agents integrate the real world with cyberspace. In: Ishida, T. (ed.) Community Computing: Collaboration over Global Information Networks. John Wiley & Sons (1998)
27. Navarro, K.F.: Wearable, wireless brain computer interfaces in augmented reality environments. In: Proceedings of the International Conference on Information Technology: Coding and Computing, ITCC 2004, vol. 2, pp. 643–647. IEEE (October 2004)
28. Neumann, U., Majoros, A.: Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. In: Proceedings of the Virtual Reality Annual International Symposium, pp. 4–11. IEEE (March 1998)
29. Raley, C., Stripling, R., Kruse, A., Schmorrow, D., Patrey, J.: Augmented cognition overview: improving information intake under stress. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 48(10), pp. 1150–1154. SAGE Publications (2004)
30. Sadda, P., Azimi, E., Jallo, G., Kazinzides, P.: Surgical Navigation with a Head-Mounted Tracking System and Display. In: Medicine Meets Virtual Reality (MMVR20) NextGen Conference, San Diego, California (2013)

31. Sauer, F., Wenzel, F., Vogt, S., Tao, Y., Genc, Y., Bani-Hashemi, A., et al.: Augmented Workspace: designing an AR testbed. In: Proc. IEEE Intl. Symp. on Augmented Reality (ISAR) (2000)
32. Schmorrow, D.D., Kruse, A.A.: DARPA's Augmented Cognition Program-tomorrow's human computer interaction from vision to reality: building cognitively aware computational systems. In: Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants, pp. 7–1. IEEE (2002)
33. Scherer, R., Chung, M., Lyon, J., Cheung, W., Rao, R.P.: Interaction With Virtual And Augmented Reality Environments Using Non-Invasive Brain-Computer Interfacing. In: 1st International Conference on Applied Bionics and Biomechanics (October 2010)
34. Sweller, J.: Implications of cognitive load theory for multimedia learning. In: Mayer, R.E. (ed.) The Cambridge Handbook of Multimedia Learning, pp. 19–30. Cambridge University Press, Cambridge (2005)
35. Tang, A., Owen, C., Biocca, F., Mou, W.: Comparative effectiveness of augmented reality in object assembly. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 73–80. ACM (April 2003)
36. Tang, A., Owen, C.B., Biocca, F., Mou, W.: Experimental Evaluation of Augmented Reality in Object Assembly Task. In: ISMAR, vol. 2, p. 265 (2002)
37. Van Krevelen, D.W.F., Poelman, R.: A survey of augmented reality technologies, applications and limitations. International Journal of Virtual Reality 9(2), 1 (2010)
38. Wilson, J., Doswell, J.: Using Wearable Augmented Reality as a Clinical-Decision Support Tool. mHealth Summit. Session on Emergency Response & Mobile Technology, Washington, DC (2010)