Perceptual momentum for design of multimodal displays

Patrik Lif, Peter Svenmarck & Per-Anders Oskarsson
Swedish Defence Research Agency
Sweden

Abstract

Many theories and studies show the benefits of multimodal displays for coding information in several modalities to increase information assimilation. However, the design principles are still unclear for coherent coding of information between modalities in multimodal displays. There is therefore a need for an overall design approach for multimodal displays, where we propose the design concept *perceptual momentum*. This concept is based on the idea of visual momentum for how correspondence between multiple displays and windows facilitates an integrated understanding. Similarly, perceptual momentum means corresponding information between modalities for redundant presentation in a bi- or multimodal display. The prediction is that this provides enhanced intuitive understanding of the coded information, which improves performance. Further, we propose four levels of perceptual momentum, high, partial, low, and conflicting information. High correspondence will likely enable the highest performance, a partial correspondence may enable high performance but with higher mental workload, whereas low correspondence and conflicting information will be detrimental for performance and mental workload. The use of perceptual momentum as a design principle for multimodal displays is discussed and exemplified with a number of studies and practical examples.

Introduction

People often have to process a tremendous amount of information in many civil and military systems. For example, driving a car includes operating the car, understanding the relations between the own car and other cars, and finally to have a good understanding of the environment and potential obstacles. While the amount of information both within the car and the environment were limited 50 years ago, the situation is different today with a more complex environment and additional information within the car. Furthermore, today an integrated computer commonly provides information about the car status, and additional information from navigation systems, radio, and cell phones may be handled. Some cars even have head-up displays (BMW, 2012; Chu et. al., 2008), tactile information in the steering wheel (Hwang & Rye, 2010) or seat (De Vries et. al., 2009), and auditory warnings for critical situations. This is partly a technical dilemma where new technological

In D. de Waard, K. Brookhuis, F. Dehais, C. Weikert, S. Röttger, D. Manzey, S. Biede, F. Reuzeau, and P. Terrier (Eds.) (2012). Human Factors: a view from an integrative perspective. Proceedings HFES Europe Chapter Conference Toulouse. ISBN 978-0-945289-44-9. Available from http://hfes-europe.org

inventions give us more and better information, but at the same time may cause accidents since we focus on the wrong things in some situations.

In more complicated system, such as an airplane or a helicopter, the pilot needs years of training, but may still get disorientated due to reduced or conflicting sensory information. Especially, the information provided by the vestibular system and the proprioceptors may be interpreted incorrectly when exposed to gravitational forces, which may lead to spatial disorientation (Albery, 2012). Further, when flying in reduced visibility the probability of spatial disorientation increases. Disorientation may also occur when landing a helicopter in a sparse environment, such as a desert when blowing sand occludes the pilot's view. This situation is known as brown-out and has caused numerous accidents for NATO forces in Afghanistan and Iraq (Albery, 2012). Both technical and human factors solution have been proposed to mitigate the risks of landing during brown-out. Some examples of human factors solution are visual displays with 2-D (Szoboszlay et. al., 2008) and 3-D symbology (Goff et. al., 2010), 3-D audio (Carlander & Eriksson, 2006), tactile displays (Eriksson et. al., 2006; Van. Erp, 2007), synthetic vision (Prinzel III et. al., 2006), and head-mounted displays (Martin-Emerson & Wickens, 1997).

The visual system is usually the dominating perceptual system for obtaining information. Consequently, the visual system is often overloaded and the user can thus not process all information. Alternative tactile displays have therefore been developed and proven to be useful in numerous situations for both civil and military applications. For example, a cane that gives tactile feedback to people that are deaf and blind (Castle & Dobbins, 2004), target acquisition in military airplanes (Eriksson et al., 2006), soldier navigation (Krausman & White, 2006), and helicopter landing aids during brown-out situations (Albery, 2012). In many of these situations, performance improves when multimodal information presentation is used to divide the presentation between visual, auditory, and tactile displays. The discussion here about multimodal displays is limited to these three modalities since they usually are the most beneficial for display presentation.

The idea that there are multiple systems for coding information is nothing new. Early on, Paivio and Clark proposed the dual coding theory for visual and auditory stimuli (Paivio, 1986; Clark & Paivio, 1991). Later, Wickens' (2002; 2008) multiple resource theory describes three dimensions: perceptual modalities, processing stages, and codes of processing. The amount of interference between information processing in these dimensions determines whether there is risk for mental overload when simultaneously performing several tasks. Proper design of multimodal displays can therefore reduce the mental workload by dividing the information in different channels (i.e., visual, tactile and auditory). Although the original theory only considers visual and auditory modalities, there are other versions of the model that also include tactile and olfactory input (Greenwood & Tsang, 2012). Finally, Van Erp and others describe a more complete model of human information processing called Prenav (Van Erp et. al., 2006). This model is based on Sheridan's model for supervisory control (1992), Wickens' information processing model (Wickens & Hollands, 2000), Veltman and Jansen's workload framework (2004), and Rasmussen's skill, rule and knowledge based classification (Rasmussen et. al., 1994). Prenav describes how information in many modalities can be processed in parallel either automatically or trough conscious processes. Intuitive displays facilitate automatic and parallel processing to enable direct action without further cognitive processing.

Problem statement

Multimodal displays that combine different displays in two or more modalities are beneficial for performance in many applications, such as cars, helicopters, and aircraft. However, the design principles are still unclear for how to combine modalities in multimodal displays into a coherent experience. For example, there are at least two different coding principles for tactile hover displays in helicopter. Van Erp et. al., (2002) coded a display so the pilots should follow a tactile stimuli to fly in the direction of the tactor indication. Raj et. al., (1998), on the other hand, used a coding principle where the pilots should steer away from the tactile stimuli. This illustrates the lack of clear and accepted design principles for tactile hover displays, and multimodal displays in general. Although it is not necessary to use the same coding of information, at least redundant information from multiple modalities for a task should be coded using the same design principle. Redundant information from visual and tactile modalities should therefore correspond in multimodal displays. This correspondence between modalities is not always sufficiently considered.

Display design

Human factors has a long research tradition for how to design displays that facilitate interaction with technology, which covers ergonomics, psychology, cognitive science, cognitive engineering, and human computer interaction (Preece et. al., 2002). For example, representation design (Bennett & Flach, 1997; Woods, 1991, 1995) considers the triad of the environment (including the task), user, and interface. Representation design describes how the interface must be designed with an understanding of what the user’s task is within the environment. Further, the interface is considered both in term of user’s visual acuity and colour limitations, as well as the cognitive requirements for processing the information.

When presenting information for a task on different displays, it is important that the information on the displays fit together and is placed in a way that the user easily can access. For example, the proximity compatibility principle (PCP) describes how information from multiple panels or display areas must be integrated, depending on the task (Haskell & Wickens, 1993; Wickens & Carswell, 1995). This means that pieces of information that need to be used to solve a specific task (mental proximity) should be displayed close to each other in perceptual space (spatial proximity). Spatial proximity is not directly applicable to multimodal displays where information is distributed between modalities. However, the spatial component is important when it comes to coding of spatial correspondence between modalities. For example, if a visual display shows left and right, a tactile display that provides the same information should provide corresponding information on left and right. The idea of mental proximity, on the other hand, is directly valid and important. Multimodal displays are often touted as facilitating parallel processing of modalities without sufficiently considering such design issues and restrictions.

Another fundamental design issue that must be considered is which modalities that should be included in a display configuration. To answer this questions a task analysis (e.g. Millitello & Hutton, 1998) should be performed to make sure that the information requirements in the present situation or task match the use of chosen modalities. In some situations multimodal displays will use the visual, tactile and 3-D audio channels redundantly, which then should be coded to correspond to each other. In other situations the best alternative will be to use different modalities sequentially, e.g. initially present a tactile stimulus to get the users attention, if that does not help present an audio signal, and finally provide visual information. Here, the focus is on multimodal displays where the redundant information from bi- or multimodal modalities can improve performance.

Perceptual Momentum

The concept visual momentum derives from cinematology, where filmmakers strive to make sure that the viewer can understand and follow the film between multiple scenes by providing a partial correspondence between the scenes. Woods (1984) describes how the same principle can be used for display design through users' ability to extract and integrate information from multiple displays, or between multiple objects within one display. In the visual system, information can only be processed in a serial fashion. The multimodal situation, on the other hand, enables parallel processing of information that multimodal displays can utilize to enhance information assimilation. We therefore introduce the design principle *perceptual momentum* that like visual momentum is intended to help the user or viewer to integrate information from multiple sources or sequences. Perceptual momentum can be defined as:

Corresponding information between modalities for redundant presentation in a bi- or multimodal display. The correspondence may be high, partial, or low, or the information may be conflicting, with a degrading scale from positive to negative (Figure 1).

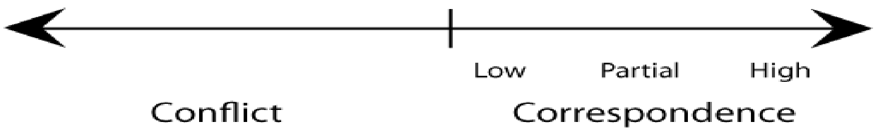

Figure 1. Continuous scale for perceptual momentum with high-, partial-, low correspondence, and on the negative side conflicting information.

The definition states the importance of considering multimodal display design as a means to facilitate integrated processing of information between different modalities. Designing multimodal displays using perceptual momentum assures that conflicting information between display modalities is avoided. Further, there is a distinction between whether the correspondence between redundant information in different modalities is high, partial, or low, or in the worst case an information conflict. A high correspondence will likely enable the highest performance, a partial correspondence may enable high performance but with higher mental workload or less spare capacity, whereas low correspondence or conflicting information will be detrimental for performance and mental workload.

Practical examples

We illustrate the benefits of using perceptual momentum for multimodal display design with a number of practical examples from both our own and others' research. The intention is to show the extent of the applicability of perceptual momentum as a design principle, not to question the research itself. The focus is on when different displays are used to present information for the same task, although multimodal displays may also be used for presentation of information for different tasks to increase performance and avoid information overload.

Firstly, we have performed three experiments with multimodal threat cuing in a simulated combat vehicle. The task was to localize threat direction and turn the combat vehicle in the direction of threat as accurately and quickly as possible. In the first experiment there was a perfect correspondence between the display modalities and a high perceptual momentum since both the visual, auditory, and tactile display presented the actual threat position. However, in this experiment a visual head down display (HDD) was used which conveyed certain disadvantages. Therefore, in the second experiment a visual head up display (HUD) was used. The HUD was, however, coded with a simplified symbology that presented initial threat direction but not initial position. Thus, concerning coding of initial position there was a lack of perceptual momentum between the visual display and the other modalities (Oskarsson et.al., 2012; Carlander et.al., 2007; Oskarsson et. Al., 2008). In a third experiment, the coding of the visual display was modified to also present information on initial threat position, i.e. information that corresponded to the tactile and 3-D audio modalities (Lif et. al., 2011).

Secondly, Van Erp and Van Veen (2004) showed that a visual and tactile display for car navigation can improve performance and reduce mental workload. The visual display showed whether the next turn should be to the left or right, as well as alphanumerically the distance (50, 150 and 250 meters) to the intersection. The tactile display showed the direction of the next turn with vibrations in the driver seat, under the left or right thigh, and the distance with different rhythms in the vibrations. Since distance in the visual display was presented alphanumerically, there was only a partial correspondence between the visual and tactile display. This illustrates how the highest perceptual momentum may not be achievable in all situations. Further, many navigation systems can provide distance estimates (that change over time), which may be difficult to present with a tactile display. A visual presentation of the continuous distance would therefore further reduce the correspondence between displays. One example of this problem is a study by Elliot et. al., (2007) where a multimodal display was used for land navigation. While there was a direct correspondence for the directional information between the visual, auditory, and tactile display, the visual display presented distance continuously, whereas the tactile display only indicated distance by a few different rhythms. The waypoints were not presented tactually.

Finally, in an on-going study with a simulated helicopter we investigate if multimodal display information can improve landing performance during brown-out and other landing situations with reduced visibility. The visual display shows drift through a dynamic vector that indicates lateral drift by a vector in a compass circle. When there is no drift the vector points upwards and when drift sideways increases the vector turns sideways in respective direction. The tactile display is a tactile belt with twelve tactors, but only two are used. The tactors indicate deviation sideways by tactile pulses under respective armpit, and the frequency of the pulses increases with speed of lateral drift. The level of perceptual momentum between the visual and tactile displays in this study is thus only partial.

Discussion and conclusions

Our goal has been to demonstrate how present design principles of visual momentum for visual displays by minor adaptation and modification are applicable to design of redundant multimodal displays. Further, when different displays are used to code information for the same task the joint multimodal display should provide coherent information to the user, which leads to a high level of perceptual momentum. Based on available theories and studies, we suggest that the usability of multimodal displays for a task depends on the level of perceptual momentum, which means the type of correspondence between the modalities. Multimodal displays that provide a direct correspondence have the highest perceptual momentum, such as showing direction with a visual compass as well as a tactile belt with roughly the same resolution. Consequently, it is also important to avoid using different design principles for each modality, since this may result in displays that conflict, and a low level of perceptual momentum. For example, with a visual display where an arrow indicates the suggested direction of movement, the tactile stimuli should be on the same side as the arrow. On the other hand, if a tactile stimulus is presented on the other side with the purpose to push the subject in the intended direction, it will be perceived as a conflict. Finally, only four levels of perceptual momentum are discussed here. This is probably sufficient as a starting point for a framework for design principles for intuitive coherent multimodal displays. Above all, multimodal displays with the highest possible perceptual momentum are more likely to be perceived as intuitive and effective.

References

Albery, W.B. (2012). *Rotary-Wing Brownout Mitigation: Technologies and Training* (No. ISBN 978-92-837-0149-1): NATO.

Bennett, N., & Flach, J. (1997). Visual Displays. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 659-696). New York: John Wiley & Sons, Inc.

BMW. (2012). *BMW Head-Up Display – Simulation Augmented Reality*, from http://www.bmwblog.com/2011/10/10/video-bmw-head-up-display-simulation-augmented-reality/

Carlander, O., & Eriksson, L. (2006). Uni- and bimodal threat cueing with vibrotactile and 3D audio technologies in a combat vehicle. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1552-1556). Santa Monica, CA, USA.

Carlander, O., Eriksson, L. & Oskarsson, P.-A. (2006). Handing uni- and multimodal threat cueing with simultaneous radio calls in a combat vehicle setting. In C. Stephanidis (Ed.), Proceedings of HCI International 2007 (HCI Part II, HCII 2007, LNCS 4555) (pp. 293–302). Berlin, Germany: Springer-Verlag.

Castle, H., & Dobbins, T. D. (2004). Tactile displays for enhanced performance and safety. In *Procedings at SPIE 5443, 269.*

Chu, K., Brewer, R., & Joseph, S. (2008). *Traffic and Navigation Support through an Automobile Heads Up Display (A-HUD)* (Technical Report No. ICS2008-05-02). Hawaii: University of Hawaii at Manoa, Information & Computer Sciences.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149-170.

De Vries, S. C., Van Erp, J., & Kiefer, R. J. (2009). Direction coding using a tactile chair. *Applied Ergonomics, 40*(3), 477-484.

Elliot, L. R., Duistermaat, M., Redden, E., & Van Erp, J. (2007). *Multimodal Guidance for Land Navigation* (No. ARL-TR-4295). Abredeen, MD: U.S. Army Research Laboratory.

Eriksson, L., van Erp, J., Carlander, O., Levin, B., van Veen, H., & Veltman, H. (2006). Vibrotactile and visual threat cueing with high G threat intercept in dynamic flight simulation. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1547-1551). Santa Monica, CA, USA.

Goff, C., Ball, A., Yapp, J., Pearce, S., & Ceawford, F. (2010). Developing a 3-D Landing Symbology Solution for Brownout. In *American Helicopter Society 66th Annual Forum*.

Greenwood, T., & Tsang, C. C. (2012). *Wickens' Multiple Resource Model*. Retrieved 2012-05-15, from https://wiki.ucl.ac.uk/display/UCLICACS/Multitasking

Haskell, I., & Wickens, C. (1993). Two- and Three-Dimensional Displays for Aviation: A Theoretical and Empirical Comparison. *The International Journal of Aviation Psychology, 3*, 87-109.

Hwang, S., & Rye, J.-h. (2010). The Haptic steering Wheel: Vibro-tactile based navigation for the driving environment. In *In proceedings of Pervasive Computing and Communications Workshops (PERCOM Workshops), 8th IEEE International Conference*. Mannheim, Germany.

Krausman, A. S., & White, T. L. (2006). *Tactile Displays and Detectibility of Vibrotactile Patterns as Combat Assault Maneuvers are Being Performed* (No. ARL-TR-3998). Aberdeen, MD: U.S: Army Research Laboratory.

Lif, P., Oskarsson, P.-A., Lindahl, B., Hedström, J., & Svensson, J. (2011). Multimodal threat cueing in simulated combat vehicle with tactile information switching between threat and waypoint indication. In M. J. Smith & G. Salvendy (Eds.), *Human Interface, Part I, Proceedings of the HCII 2011, Orlando* (pp. 454-461). Berlin Heidelberg: Springer-Verlag.

Martin-Emerson, R., & Wickens, D. D. (1997). Superimposition, Symbology, Visual Attention, and the Head-Up Display. *Human Factors, 39*(4), 581-601.

Militello, L.G., & Hutton, R.J.B. (1998). Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, 41, 1618-1641.

Oskarsson, P.-A., Eriksson, L., & Carlander, O. (2012). Enhanced perception and performance by multimodal threat cueing in simulated combat vehicle. *Human Factors, 54*, 122-127.

Oskarsson, P.-A., Eriksson, L., Lif, P., Lindahl, B., & Hedström, J. (2008). Multimodal Threat Cueing in Simulated Combat Vehicle. In *Prceedings of the Human Factors and Ergonomics Sisciet Annual Meeting*. Santa Monica, CA.

Paivio, A. (1986). *Mental representations: A dual coding approach* (Vol. 9). New York: Oxford University Press.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: beyond human-computer interaction*. New York: Wiley.

Prinzel III, L., Kramer, L., Arthur, J., & Bailey, R. (2006). Multi-Dimensionality of Synthetic Cockpit Displays: Prevention of Controlled-Flight-Into-Terrain. In *Proceedings of the Human Factors and ergonomics Society 50th Annual Meeting* (pp. 2056-2060). San Fransisco, CA, USA.

Raj, A. K., Suri, N., Braithwaite, M. G., & Rupert, A. H. (1998). *The tactile situation awareness system in rotary wing aircraft: Flight test results.* Paper presented at the Proceedings of the RTA/HFM Symposium on Current Aeromedical Issues in Rotary Wing Operations, Neuilly-sur-Seine, France.

Rasmussen, J., Pejtersen, A. & Goodstein, L. (1994). Cognitive Systems Engineering. New York: John Wiley & Sons, Inc.

Sheridan, T.B. (1992). Telerobotics, Automation, and Human Supervisory Control, MIT Press.

Szoboszlay, Z., Albery, W. B., T., T., & Neiswander, G. M. (2008). Brown-Out Symbology Simulation (BOSS) on the NASA Ames Vertical Motion Simulator. In *American Helicopter Society 64th Annual Forum*.

Van. Erp, J. (2007). *Tactile displays for navigation and orientation: perception and behaviour.*, (Doctoral thesis). Utrecht University, The Netherlands.

Van Erp, J., Duistermaat, M., Jansen, C., Groen, E., & Hoedemaeker, M. (2006). Tele-Presence: Bringing the Operator Back in the Loop. In *In Proceedings of Virtual Media for Military Applications, meeting NATO RTO-MP-HFM-136.* (pp. 9-1 - 9-18). Neuilly-sur-Seine, France.

Van Erp, J., & Van Veen, H. (2004). Vibrotactile in-vehicle navigation system. *Transportation Research Part F, 7*, 247-256.

Van Erp, J., Veltman, J., van Veen, H., & Oving, A. (2002). Tactile Torso Display as Countermeasure to Reduce Night Vision Goggles Induced Drift. *Paper presented at the RTO HFM Symposium on "Spatial Disorientation in Military Vehicles: Causes, Consequences and Cures".* La Coruna, Spain, RTO-MP-086.

Veltman, J.E. & Jansen, C. (2004). The adaptive operator. In D.A. Vincenzi, M. Mouloua, & P.A. Hancock (Eds.), Human Performance, Situation Awareness and Automation Technology (HPSAA II). Daytona Beach, Vol II, pp. 7-11.

Wickens, C. (2008). Multiple Resources and Mental Workload. *Human Factors, 50*(3), 449-455.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomic science, 3*(2), 159-177.

Wickens, C., & Carswell, C. (1995). The Proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors, 37*(473-494).

Wickens, C., & Hollands, J. (2000). *Enginerring Psychology and Human Performance* New York: Ptentice-Hall Inc.

Woods, D. D. (1984). Visual Momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine. Studies, 21, 229-244*.

Woods, D. D. (1991). The Cognitive Engineering of Problem Representations. In G. R. S. Weir & J. L. Alty (Eds.), *Human-Computer Interaction and Complex Systems* (pp. 169-188). London Academic Press.

Woods, D. D. (1995). Towards a theoretical base for representation design in the computer medium: ecological perception and aiding human cognition. In J. Flach, P. Hancock, K. Caird & K. Vicente (Eds.), *An ecological approach to human machine systems I: A global perspective* (pp. 157-188). New Jersey, USA: Erlbaum: Hillsdale.