TOO MUCH, TOO LITTLE OR JUST RIGHT: DESIGNING DATA FUSION FOR SITUATION AWARENESS

Geoffrey B. Duggan  Simon Banbury  Andrew Howes  John Patrick  Samuel M. Waldron

School of Psychology, Cardiff University, Cardiff, U.K.

Many operators within the battlefield find themselves unable to process all the data presented to them in the limited time available. Data fusion provides a means of reducing their workload, but can also reduce system transparency. Thus, either too much or too little fusion can lead to reduced operator situation awareness. A framework is proposed that incorporates a role for more sophisticated psychological theory when attempting to understand the consequences of data fusion technologies on SA. Four key questions are identified: How should the level of certainty in fused information be presented? How much does it cost the operator to “drill-down”? To what extent does data fusion inhibit representation change? Does data fusion ameliorate or exacerbate the consequences of interruption?

INTRODUCTION

The battlefield is typically a highly dynamic environment and the proliferation of information sources, both onboard (radar, infra-red etc.) and offboard (e.g. data-link), has meant that operators (e.g. pilots or ground commanders) often find themselves unable to process all the data presented to them in the limited time available. This phenomenon is often described as having poor “Situation Awareness” (SA), where SA refers to the complex set of information that must be maintained in real-time tasks. Reviewing the different perspectives upon SA is beyond the scope of this article (see Banbury & Tremblay, 2004, for a review), however, given the dynamic nature of many data fused environments and the psychological thrust of our arguments we find it helpful when discussing SA to consider not only the person’s knowledge of a situation but also the processes responsible for such knowledge. A number of technologies have been proposed to ameliorate the problems of data overload and the associated loss of SA. One option is the use of automated systems to handle more routine tasks. Another option is to use data fusion technology to try to summarise/collate/clean-up the streams of data from many disparate sources into one coherent picture that is presented to the operator.

The potential advantages of data fusion to facilitate SA are clear, however, many of the attendant issues and dangers are under researched and poorly understood. For example: How should data fused information be presented to the operator? Are there any unforeseen problems with operators using information that is fused? Given that the operator is considered to be a source of data, and that this information is propagated back into the “battlefield network”, what impact does data obtained or modified by one operator have on another operator trying to interpret these data?

One particular, perhaps paradoxical, danger of data fusion is that too much fusion may actually reduce SA because of the very fact that it reduces data. The extent to which data fusion will improve SA is contingent on the amount of “system transparency” achieved. (System transparency is an important concept within the automation literature, see e.g. Bainbridge, 1983.) Thus, although greater clarity can be achieved through filtering out unnecessary data, it is important that useful information is not also discarded along the way. Good SA is dependent on understanding the displayed information in context. Too much fusion can reduce system transparency and deprive the operator of this context. As shown in Figure 1 there is a “Goldilocks effect” of data fusion upon SA. Too little fusion leads to an excessively high workload for the operator with poor SA as a result. However, too much fusion will reduce system transparency also producing poor SA. Happily, just the right amount of data fusion should reduce workload, maintain system transparency and optimise...
SA. That is not to deny that the optimum level of data fusion is, however, dynamic as it will depend on the task, the environment and the operator. Thus, the design of any system must be flexible and able to accommodate this inevitable variability. (It should also be noted that the use of the term “poor SA” should not be understood to imply that SA is quantifiable upon one dimension. SA can be represented along qualitatively different types of knowledge which are difficult to summate across.)

Evaluation of data fused entities must, therefore, account for the impact upon operator SA, as it provides both an index of their effectiveness (concentration of relevant information) and a check upon the dangers associated with too much fusion (reduction of system transparency). Furthermore, given that data fusion exists to help an operator assimilate large amounts of incoming information, it is unsurprising that it is typically employed within highly dynamic environments where the content of SA is constantly changing too. Thus, we seek to investigate the impact of data fusion techniques upon the processes that constitute SA enabling a fuller appreciation of the overall effect of data fusion.

The standard approach to enhancing SA through data fusion is to focus upon technological improvements that reduce the volume of data presented to the operator. Moreover, human factors research (e.g. Endsley, 1995) has often concentrated upon developing effective tools to measure the impact of other manipulations upon SA. We propose an alternative framework that incorporates a role for more sophisticated psychological theory when attempting to enhance SA through data fusion. The loss of system transparency associated with too much data fusion can be addressed by a fuller understanding of how the operator completes the task and then designing the system in those terms. Rather than relying solely upon technological solutions to the difficulties presented by large amounts of data the aim is to also exploit the competencies of the operator.

This approach is consciously aligned with the current trend within cognitive science to emphasise the interaction between internal and external representations (see, e.g., Ballard, Hayhoe, Pook & Rao, 1997; Gray & Fu, 2004). The distinction between knowledge in-the-head and knowledge in-the-world captures many of the features of successful interface design (Payne, Howes & Reader, 2001). For example, representing information about both the range and location of enemy missile sites upon a map removes the need for a pilot or ground commander to remember or store this knowledge in-the-head. Moreover, due to this representation less cognitive effort is required to compute whether other units represented upon the same map are in danger of attack. Restated from this perspective the benefits of data fusion result from representing knowledge in-the-world in an appropriate fashion. However, knowledge in-the-world should not be presumed to be necessarily more

![Figure 1: The Relationship Between Data Fusion and SA](image-url)
efficacious than knowledge in-the-head. Across tasks such as list-learning (Slamecka & Graf, 1978), text comprehension (McNamara, Kintsch, Songer, & Kintsch, 1996) and learning to use a computer application (Palmiter, Elkerton, & Bagget, 1991) inference making or internal processing has been shown to improve subsequent task performance relative to a passive reliance upon equivalent information provided within the environment. McNamara et al. (1996) suggest that the process of inference making may lead to a more elaborated situation model. Or, to put it another way, the process of developing knowledge in-the-head can lead to better situation awareness than reliance upon knowledge in-the-world.

Ultimately, following Anderson (1990) and Payne et al. (2001), we would argue the relationship between internal and external representation can be more usefully understood as an adaptation by the user to the information structure of the environment. Anderson suggests that cognitive systems and behaviours can be understood as adaptations that seek to maximise the value of the expression PG-C, where P is the probability of achieving the current goal, G is the gain associated with achieving that goal and C is the cost of achieving the goal. People may offload cognitive effort to the environment but this decision will be determined by the relative costs and benefits of internal and external representations in each given situation.

Gray and Fu (2004) provide evidence for this position by demonstrating that even small scale alterations to an interface can affect the extent to which users externalise their representations while completing a task. The sensitivity of this adaptivity further highlights the need to ensure the level of data fusion is appropriate to the task, the environment and the operator. Data fusion should seek to maintain system transparency, moreover, where appropriate it should also exploit the flexibility and sophistication it is possible to attain through the creation of internal representations.

THE FRAMEWORK

We introduce four key questions where psychological theory can inform the assessment of a data fusion technology designed to improve SA.

1. How much does it cost the operator to “drill-down”?

Typically data fused information is more accessible than unfused information, however, there is potentially an associated loss of system transparency. One solution is to provide the operator with the capability to drill-down into any fused data, that is, the interface provides a means by which they can explore the original sources from which the fused entity was constructed. The act of drilling-down does, however, contain its own costs (e.g. interruption from primary task, time costs) that must be balanced against the higher level goals (e.g. mission effectiveness, survivability).

As noted above knowledge can be stored in-the-head as well as in-the-world which can have a positive or negative affect on task performance. Moreover, the work by Gray and Fu (2004) demonstrated that users are sensitive to very small differences within interfaces and as a consequence may adopt different microstrategies. These microstrategies can affect the extent to which users internalise information or leave it within the environment. Duggan and Payne (2001) demonstrated that participants internalised larger sets of information if the information display was more expensive to visit. Conversely, if information is readily available they found that people appeared to put less effort into maintaining an internal representation and that task performance was subsequently degraded. These findings indicate the potential negative consequences of information displays that make information too easy to acquire (i.e. by reducing the probability that an internal representation is maintained).

Two specific scenarios where difficulties associated with the cost of information access could arise are (1) If a piece of information is particularly difficult to acquire an operator may be unwilling to attend to other, more critical, tasks due to their determination to complete the process of “drilling-down”. (2) Having invested a lot of effort into acquiring a piece of information operators may attach disproportionate importance to it partly to ensure it does not have to be retrieved again.

2. How should the level of certainty in fused information be presented?

Where there is a lot of data to be combined or “fused”, there is usually a degree of ambiguity within the final picture or situation to be conveyed to the operator. No sensor produces entirely accurate information, and degree of accuracy varies with time. Atmospheric conditions, target characteristics, and intrinsic design limitations can all introduce a degree of error into displayed information. Thus, data received from a
number of different sources can be contradictory, or at the least approximate. How this ambiguity or uncertainty is represented to the operator and the amount of system transparency provided are pertinent issues within data fusion. Moreover, only a small body of research has explored the issue of displaying uncertainty in the context of aviation (e.g. Banbury, Selcon, Endsley, Gorton & Tatlock 1998).

By contrast, there is a wealth of laboratory studies examining the effects of uncertainty on decision-making. A wide range of laboratory investigations have demonstrated that humans are poor decision-makers under uncertainty (for a review see Kahneman, Slovic & Tversky, 1982). However, Gigerenzer and Goldstein (1996) have argued that many of these results are attributable to simple pragmatic heuristics that are effective outside of the laboratory and within naturalistic settings where time and processing resources are limited. Interfaces designed for uncertain, dynamic environments must exploit this adaptivity whilst compensating for the risks associated with this type of decision making.

3. To what extent does data fusion inhibit representation change?

Expert operators, by definition, must have some mental representation of the system they use. However, within a dynamic environment a crucial aspect of effective performance is the capability to adapt or change the current representation to accommodate new information. The Cali aircraft accident report (1995) outlines how after losing awareness of the current position the pilot ignored repeated warnings to gain altitude and crashed into the side of a mountain. Wason (1960) observed that when testing a hypothesis participants demonstrated a bias to seek evidence that supported rather than disconfirmed the hypothesis. In instances like Cali this “confirmation bias” can lead to the overlooking of information that is relevant but does not confirm previously held beliefs.

Constraints upon performance are built into any task and one potential method for producing representation change is the elimination of constraints (Richard, Poitrenaud & Tijus, 1993). This approach may have particular relevance for data fused displays where any loss of system transparency could hinder the operators’ awareness of constraints they are unwittingly operating within, thus discouraging potential representation change. Furthermore, the work of Richard et al. (1993) indicates the actual process of eliminating constraints is closely related to representation change. Thus, the extent to which a representation is externalised may also affect its resistance to change.

4. Does data fusion ameliorate or exacerbate the consequences of interruption?

The advanced technology used in the cockpit at present is able to carry out several tasks without constant human control. This increases the operator’s ability to multi-task. Although this is a large help to the operator it also increases the frequency with which the operator will be interrupted (McFarlane, 2002). Whenever an advanced system wishes to make contact with the human operator it must first interrupt him/her from their present activity. These interruptions can have a detrimental effect to the individual’s SA. Having been interrupted, the operator may “lose track” of the current situation, forget to resume the pre-interruption task, or be engrossed in the interruption and neglect another, perhaps more important, task. Although the traditional interruption literature is vast and is often applied to the study of human-computer interaction, few studies have systematically investigated interruptions in terms of SA.

Preliminary research has looked at how best to present an interruption to the operator. For example, McFarlane (2002) found that the method of coordination (negotiated, mediated, scheduled, or immediate) has a large effect on how the interruption is dealt with by the individual. Research needs to look at how data fusion techniques can inform the presentation of an interruption. That is, how much and in what fashion should information pertinent to the interruption be displayed. Using data fusion techniques it will be possible to design novel user-interfaces to support the operator in times of interruption allowing interruptions to be managed efficiently and safely. These display designs should enable the user to extract information about the interruption in an easier, more accessible fashion.

As well as using data fusion techniques as a means to present interruption relevant information in the least disruptive manner, it is also useful to examine the effect of data fused displays upon the user’s ability to manage interruptions. Distributing knowledge in-the-world within a fused display may support recovery from interruption as the user is required to memorise less information about the target task during the interruption. Alternatively, the heavier reliance upon knowledge in-the-head within unfused displays may produce a greater internal SA. This more elaborated internal representation may prove more robust during the interruption relative to less complete internal SA produced by a fused display.
Moreover, these effects may be mediated by the different methods of coordination highlighted by McFarlane (2002).

Thus, two pertinent research questions are: How can data fusion techniques be used to present interruption relevant information to the operator as efficiently and safely as possible without compromising an operator’s SA? Are the consequences of an interruption mediated by the degree of data fusion within the interrupted task?

**SUMMARY**

Data fusion enables the knowledge necessary to complete a task to be situated in-the-world thus reducing the operator’s workload. However, data fusion does not inevitably lead to an improvement in SA. There is an optimum level of fusion that varies according to the task, the operator and the environment. This optimum level should maintain system transparency and support the benefits concomitant with processing knowledge in-the-head. Four key questions are introduced where psychological theory can help identify the appropriate level of fusion and more generally inform the use of data fusion within system design.

**ACKNOWLEDGEMENTS**

This work was funded by the UK MOD’s DIF: DTC (Project 4.9).

**REFERENCES**


