Applying Insights on Categorisation, Communication, and Dynamic Decision-Making: A Case Study of a 'Simple' Maritime Military Decision

Charlotte E. R. Edmunds[1] ⓘ, Adam J. L. Harris[2], and Magda Osman[3]

Abstract

A complete understanding of decision-making in military domains requires gathering insights from several fields of study. To make the task tractable, here we consider a specific example of short-term tactical decisions under uncertainty made by the military at sea. Through this lens, we sketch out relevant literature from three psychological tasks each underpinned by decision-making processes: categorisation, communication and choice. From the literature, we note two general cognitive tendencies that emerge across all three stages: the effect of cognitive load and individual differences. Drawing on these tendencies, we recommend strategies, tools and future research that could improve performance in military domains – but, by extension, would also generalise to other high-stakes contexts. In so doing, we show the extent to which domain general properties of high order cognition are sufficient in explaining behaviours in domain specific contexts.

To comprehensively understand the psychological mechanisms underpinning our behaviour in high stakes, applied, contexts (e.g. medical, legal, military or financial decision-making), we must integrate insights from many disparate fields. In this review, we demonstrate this by exploring psychological fields (e.g. judgement and decision-making, category learning, communication) and factors (e.g. cognitive load, individual differences) relevant to understanding the process of decision-making (i.e. choosing between different options) in military contexts. We are particularly interested in military tasks because the nature of the domain means relying on multiple actors, at different levels of a hierarchy, where effective coordination of information across actors is critical, because the stakes are high (Goodwin et al., 2018). The general approach we take is to draw domain general insights from psychology to explore how they apply to a specific domain, and then generalise back to consider the extent to which aspects of cognitive processes are found generally in high-stake, dynamic, real-world contexts,

The review is organised in the following way. We begin by outlining a particular scenario in maritime military decision-making. We structure our review by dividing the scenario into three distinct tasks and considering the relevant psychological literature of each in turn. We end by identifying and discussing the key psychological factors that are important to

consider for all three stages and discussing potential areas where improvements are possible.

The Scenario

So, what do we mean by 'maritime military decision-making'? By definition, maritime decision-making encompasses all the decisions that might be made at sea, from whether to enter potentially hostile waters, to whether to stock up the ship's freezer at the next port. In this review, we focus on one specific, narrowly defined, scenario – from which we subsequently draw broad parallels to other areas of high stakes decision-making. Specifically, we consider the key tactical decision of whether, where and how a ship engages a nearby

[1]Department of Experimental Psychology, University of Oxford, Oxford, UK
[2]Department of Experimental Psychology, University College London, London, UK
[3]Centre for Science and Policy, University of Cambridge, Cambridge, MA, USA

Corresponding Author:
C. E. R. Edmunds, Department of Experimental Psychology, University of Oxford, Anna Watts Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK.
Email: ceredmunds@gmail.com

enemy vessel (Holmes, 2001). For a simplified visualisation of the scenario, see Figure 1.

A widely accepted prerequisite for making tactical decisions is Maritime Domain Awareness (MDA); an understanding of how the ship relates to its surroundings (Carvalho et al., 2011; Hammond, 2006). One aspect of this requires commanders (or other decision makers) to have awareness of the physical locations of surrounding craft as well as other relevant information such as their type (e.g. whether it's an aircraft, submarine or ship) and especially their status (i.e. whether they are friendly or hostile). Thus, the first task in our scenario is to determine the status of surrounding entities – ships, aeroplanes and submarines. Entity status is typically estimated by integrating position data with information from the cooperative Automatic Identification System (AIS; Hammond, 2006). This is a signal emitted by all vessels over a certain tonnage and includes information such as the vessel's identity number, navigation status (at anchor, underway using engines etc.), speed, direction, and details about course (destination, estimated arrival time etc.).

On a Royal Navy vessel, junior operators situated in the operations room use available positional and AIS data to partition the surrounding entities into four categories: 'unknown', 'neutral', 'entity of concern' or 'entity of allies' ('NATO Joint Military Symbology', 2017). Unknown is the default category. New vessels that enter the monitored area are automatically assigned Unknown, until an operator decides otherwise. Entities might also be assigned to Unknown if the information available does not exceed the threshold for one of the other categories. Neutral craft are those which will not cause harm but are also not able to help if needed (e.g. a passenger airline). Entities of concern are those likely to be Hostile and, finally, entities of allies are those which are Friendly and could possibly provide or require assistance.

Once an entity's status has been assigned by operators in the operation room, this information is passed up the chain of command, firstly to the Principal Warfare Officer who integrates the information from multiple operators and then passes up the chain of command to those making tactical decisions. Tactical decisions involve evaluating which of a set of possible actions is the optimal one to take, given information about classifications and positions of surrounding entities (Cummings et al., 2010; Kobus et al., 2001; Matthews et al., 2009). Options are evaluated based on the overarching goals of the mission as well as the proximity and inferred intentions of the surrounding entities and other considerations such as costs and available equipment. Further, these short- and long-term goals may change over time, as does the surrounding situation. Thus, the military planner must evaluate and prioritise between conflicting goals, whilst maintaining an up-to-date awareness and understanding of the situation.

The 'simple' scenario which is the focus of the present review can thus be decomposed into three stages: classification of entities' statuses, communicating that information up the hierarchy and finally, choosing an appropriate action. Each of these stages has parallels in areas of psychological research. In the following, we examine each of these tasks, and their corresponding literatures, in turn.
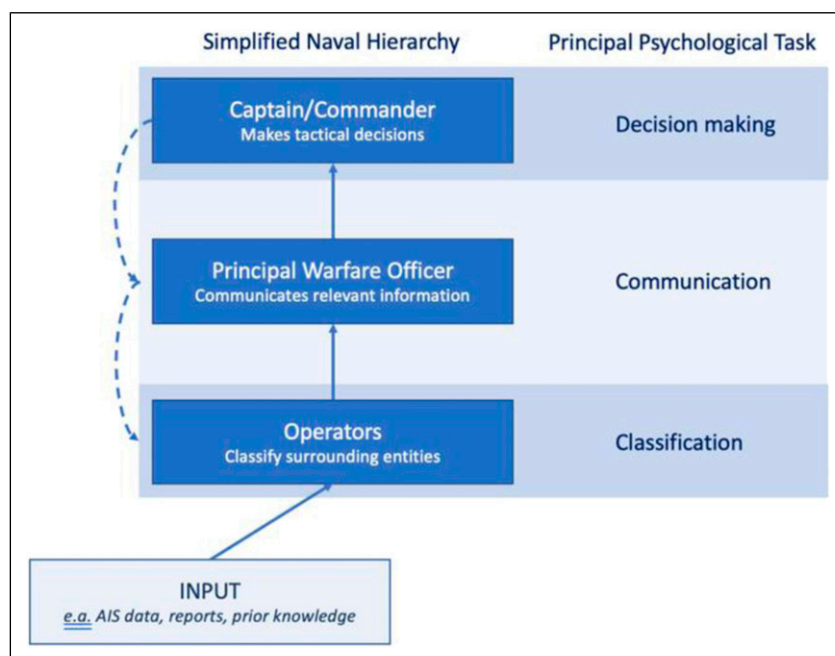

Figure 1. Simplified Naval structure shown alongside the principal psychological process shown at that point. Solid arrows represent the main direction of information flow as considered in the current article. Inclusion of the dashed arrows acknowledges that, in reality, there will be bidirectional communications.

Classification of Entities

The classification of entities, as described in the previous section, closely mirrors work from the categorisation (or 'category learning') literature. In a typical experiment in this field, naïve participants sort stimuli into categories; participants are shown a stimulus, asked to assign it to a category, then are given corrective feedback on every trial (Kurtz, 2015). Generally, categorisation studies aim to understand how participants learn, understand, and generalise the underlying category structure(s), the (usually) experimenter-defined mapping of stimuli to category labels. The category structure is usually chosen to explore the underlying psychological mechanisms of categorisation (Ashby & Valentin, 2018; Kurtz, 2015). Thus, these experiments examine how participants' performance is affected by varying the category structure factorially with other features of the task. For instance the experimenters may add a concurrent task or increase time pressure by only giving participants a limited time to respond. These manipulations affect the 'cognitive load' of the task – that is, how many cognitive resources are needed to complete the task effectively.

How are Classification Decisions Made?

In our maritime scenario, operators estimate entity status from diverse sources of information, principally AIS and position data. This data can be supplemented by information requested from others and by knowledge of the combat theatre. Reframing this as a categorisation problem, operators infer category membership by combining many stimulus dimensions. These stimulus dimensions could be continuous, such as bearing; ordinal, such as overall AIS transmission quality; or discrete, such as destination. This is not only difficult but also highly consequential as some of the objects of interest may be fast moving and pose a serious threat (Finger & Bisantz, 2002; Liebhaber & Feher, 2002; Riveiro et al., 2018). In the following, we will explore the possible mechanisms by which operators may make these categorisation judgements.

The dominant explanation, seen in many different theoretical accounts of categorisation (Pothos & Wills, 2011), is that operators compare information from *every* stimulus dimension to representations of *every* category (Nosofsky, 1986; Wills et al., 2020). The entity would then be assigned to the category to which it is most similar, where 'most similar' is typically determined geometrically, in terms of minimising the psychological distance between the stimuli and the representation of the category (e.g. Nosofsky, 1986).

To the unfamiliar, this account suggests that categorisation might be a long, tedious, and cognitively demanding process; not ideal during a high-stakes scenario when an enemy fighter jet may be flying towards you at speeds surpassing the speed of sound. However, seminal categorisation research suggests that increasing the cognitive load (by increasing time pressure or adding a second task) resulted in participants using more

attributes to categorise (e.g. Kemler Nelson, 1984; Smith & Kemler Nelson, 1984; Ward, 1983). From these results, researchers inferred that people represent stimuli holistically, as an undifferentiated 'mass'. Thus, comparing multiple attributes would be less cognitively demanding that differentiating a single attribute from that representation.

Another theoretical approach argues that whether people use holistic representations of entities depends on the underlying category structure. The COVIS (COmpetition between Verbal and Implicit Systems) model of category learning proposes a dual-system mechanism (Ashby et al., 1998; Edmunds & Wills, 2016). This theory argues that simple category structures are learnt using rules based on a subset of the available information. Structures that are difficult to verbalise are learnt through an implicit, overall similarity approach, where stimuli are represented holistically. Much evidence has been argued to support this perspective (for reviews see Ashby & Maddox, 2004, 2011; Ashby & Valentin, 2017; Smith & Church, 2018). Given that the operator's job of classifying an entity as hostile would be hard to verbalise, this approach supports the seminal research in recommending (somewhat paradoxically) that the way to optimise classification in a military setting would be to *increase* cognitive load.

In contrast, and of critical importance, Wills et al. (2015) showed two critical findings that conflict with earlier theoretical claims. Firstly, participants used a wider range of strategies than previously supposed (many of them resulting in reduced accuracy). Secondly, participants were actually more likely to use *fewer* dimensions with a *greater* cognitive load.

In a similar vein, there is no evidence that participants use different types of stimulus representation depending on the category structure (for partial reviews see Newell et al., 2011; Wills et al., 2019). Participants report using rule-based strategies in both simple and complex categorisation tasks (Edmunds et al., 2015, 2016, 2019, 2020). For instance they may categorise using a single attribute or perhaps generate more complex rules (such as conjunctions or other combinations of single dimension rules). Thus, recent evidence suggests that participants are most likely to use a subset of the available information using a rule-like strategy, even though this type of strategy often results in reduced accuracy.

In sum, the categorisation literature suggests that only when participants have enough time or resources do they use all the available dimensions to categorise. The rest of the time they use a 'satisficing' heuristic: they use less of the available information than would be optimum (Simon, 1947). This finding is mirrored in the defence literature. For instance Liebhaber et al. (2000) found that the number of potential cues available to inform the classification of an entity as a threat was independent of the actual threat rating given to an object. That is it appears that military planners classify entities based on a small subset of attributes rather than integrating all the information together. Further, the attributes that are used depend on the prior knowledge of threat profiles that

correspond to the available data the military planner has available to them. Altogether, this work suggests that minimising cognitive load may be critical in improving the performance of operators trying to estimate the status of surrounding entities.

On What Information are Classifications Made?

The evidence in the previous section suggests that, in both military and lab settings, it is unlikely that all the available information at any one time will be used to make a classification. Failing to use all the available information could result in serious errors being made in a defence context. For instance operators appear to be susceptible to focussing on information that confirms their initial hypotheses regarding the threat profile of particular entities (Matthews et al., 2009). Narrowing the focus of attention to data that is consistent with prior hypotheses comes at the expense of attending to evidence that contradicts and undermines the focal hypotheses that the planner(s) are working with. Thus, failing to incorporate all the available evidence could have significant negative consequences as potentially relevant information could be neglected.

However, it is also possible that satisficing in this context is an adaptive way of responding given the considerable time pressure. After all, sorting and identifying the most valuable information is key in military decision-making (Mishra et al., 2015). Therefore, if operators are responding using a subset of the available information, but that information is almost as accurate as using all the available information and is made faster, then this is more efficient overall. In the next section, we explore whether people are tactical in selecting which subset of attributes they use.

Do People Classify by Picking the Most Useful Dimensions, or Choosing at Random?

For both participants in psychology experiments and operators on Royal Navy warships, the best approach would be to focus on the information that is most predictive of the outcome(s). For simple categorisation tasks, there is plenty of evidence that participants can identify the most predictive attribute in a unidimensional categorisation task (e.g. Ashby & Valentin, 2017; Shepard et al., 1961; Nosofsky et al., 1994; Wills et al., 2015), and can do so quickly (e.g. Nosofsky et al., 1994; Shepard et al., 1961).

Yet, in the field, operators may have to classify stimuli that are far more complex, under considerable cognitive load. Laboratory studies suggest that, when there are many additional, non-predictive attributes (Vong et al., 2019), or when a cognitive load is added (Wills et al., 2015), participants are less likely to find the optimum simple rule and instead rely on less predictive attribute, resulting in poorer accuracy.

Two other features of our maritime military scenario are that the information that can be used to make classifications changes over time, and there is uncertainty in the data

received by the operator (Irandoust et al., 2010; McCloskey, 1996; Potter et al., 2012). Unfortunately, very few categorisation experiments have examined scenarios where the category rule is maintained (for instance Category A represents small stimuli), but the distribution of stimuli changes over time (for instance at the beginning of the experiment, Category A stimuli are about 1 cm, by the end they are around 0.1 cm). Rather, studies have investigated the category structure changing over time (for instance at the beginning of the experiment the participant needed to focus on size, later they needed to focus on orientation). Some have looked at analogical transfer where participants are first trained on one category structure and then in a second phase, they must transfer this knowledge to another set of stimuli, outside of the range of the training stimuli (Casale et al., 2012; Edmunds et al., 2020; Soto & Ashby, 2019; Zakrzewski et al., 2018). Helie et al. (2015) looked at categorisation performance after the category structure was changed (see also Cantwell et al., 2015). Navarro et al. (2013) examined whether people could learn a unidimensional 'height' category structure where the category boundary slowly changed over time. Generally, these studies show that people are poor at these tasks: participants are slow to adapt (Navarro et al., 2013), do not take advantage of previously acquired knowledge (Edmunds et al., 2020), or may even fail to notice the change (Cantwell et al., 2015; Helie et al., 2015).

Similarly, uncertainty is rarely explored in the categorisation literature, but the limited available evidence suggests that people find it difficult to incorporate it effectively. For instance people make slower and less accurate responses when unsure about category membership (Grinband et al., 2006). The predictiveness and uncertainty associated with stimulus features in category learning studies are also closely linked with reward. Typically, the studies follow the stimulus-response-feedback protocol (Kurtz, 2015) where reward is often contingent on the proportion of correct responses. Predictive dimensions are rewarded more often, whereas uncertain dimensions are likely to be rewarded less (or perhaps more erratically). This area is particularly relevant given the highly consequential nature of military decision-making.

Given the high stakes in maritime military contexts, the incentives that may influence classification are preventing or minimising significant losses. In the lab, the only way to approximate this ethically is to introduce high stakes by offering a greater reward for better responding. Fortunately for the maritime military context (and the ability to generalise results from the lab to this context) it appears that reward has a negligible effect of the quality of categorisation. Schlegelmilch and von Helversen (2020) examined categorisation of three-dimensional stimuli, where one stimulus in each category was rewarded more for a correct answer than the other stimuli. Generally, they found that adding a high reward reduced performance on the stimuli with lower rewards. However, they also found that the addition of a higher reward had no effect on participants' self-reported attention to

each stimulus dimension. This contrasts with other similar paradigms where the value of the outcomes modulate attention: predictors of high value rewards receive especially high levels of attention (Le Pelley et al., 2016). This evidence suggests that categorisation may only be slightly affected by reward magnitude. Thus, perhaps the extreme possible consequences of categorisation in our scenario may not affect performance. It remains an open question whether the pressure of anticipating a positive reward and the pressure of avoiding a negative consequence can be equated in a categorisation task.

Categorisation: Recommendations, Limitations and Future Work

Is the categorisation literature useful in providing insights to help field operators improve classification of entity status? Generally, the findings reviewed contribute some useful insights. The laboratory evidence suggests that it is unlikely that operators will use all available data in their categorisation decisions (e.g. Edmunds et al., 2020; Wills et al., 2015). Rather, they are likely to focus on a subset of the available information, probably that which they found worked on previous occasions (Liebhaber et al., 2000; Matthews et al., 2009). By being aware of these heuristics, we can design displays and formulate training to help operators make better classification decisions. For instance if we know which types of evidence operators are focussing on, we can predict the types of errors they will make. In other words, the types of entities that will be misclassified given the internal representation of the category structure they have constructed. Thus, we can go some way to alleviating errors that might arise by using suboptimum, but efficient, strategies. Also, the laboratory evidence from the work examining changes in category structure over time suggests that operators might fail to notice changes and thus, may be overconfident in their judgements. However, this is an extremely tentative conclusion that needs to be investigated in future research.

Transferring insights from the general categorisation literature to maritime operators, requires an assumption that those laboratory experiments reported in the literature generalise to the real-world maritime situation. There are, however, key differences between these situations. Generally, laboratory studies of categorisation are substantially more restricted because they must allow researchers to isolate experimental effects from participants' prior knowledge. In contrast, stimuli outside the lab are far more complex, with many more dimensions. Further, the dimensions in 'real-life' stimuli are likely to be correlated (for instance as size goes up, so does weight) whereas most experimenter-constructed categories deliberately avoid interactions between dimensions (Rehder & Murphy, 2003).

There are other differences between a typical category learning study (Kurtz, 2015) and 'real-world' classification.

Operators' estimation of a craft's status may be influenced by the status of other surrounding craft and other features of the environment that are subject to change. Additionally, the dimensions operators are using to classify entities may be unreliable and subject to changing levels of uncertainty. Finally, the data from the AIS signal must be carefully scrutinised for errors or logical inconsistencies and the reason for mismatches ascertained (Riveiro et al., 2018).

Therefore, conducting category learning studies with more complex stimuli and category structures is extremely important to check that the results generalise to military (and other high stakes) contexts. Moreover, devising tasks that incorporate dynamic uncertainty can provide useful insights into the ways in which people detect (or not) discrepancies in changing information, and the reliability of the classifications resulting from it.

Communication

In our scenario, once the incoming data has been classified, it must be communicated to those who need to use it to make tactical decisions. Maritime military decision-making is distributed across individuals and teams (Song & Kleinman, 1994; see Figure 1); information must be passed from operators near the bottom of the hierarchy to commanders near the top. An added difficulty is that this information may still be associated with some uncertainty. For instance operators may only be 80% sure that a craft is friendly or there may be some uncertainty surrounding the craft's position or direction. Failing to acknowledge and appropriately incorporate uncertainty into decision-making can result in failing to appropriately judge which options are least likely to result in negative outcomes (Kahneman, 2011; Pizer, 1999; Sunstein, 2002, 2003). This means that it is important for commanders to consider any uncertainty associated with the information they are using to make tactical decisions. In other words, successful maritime military decision-making ought to incorporate uncertainty, and clearly. The question is, how are uncertainties best communicated?

What are the Best Ways of Representing and Communicating Uncertainty?

The literature suggests three approaches to communicating uncertainty: verbal probability expressions (VPEs), numerical uncertainty statements, and visualisations. VPEs add a verbal qualifier to a factual statement to indicate uncertainty (Budescu & Wallsten, 1995). Numerical uncertainty statements go a step further and quantify the uncertainty (Joslyn et al., 2011). The simplest numerical statements simply add an explicit, numerical estimate of the likelihood of something, whereas more complex numerical statements might add range information around a numerical estimate. Finally, uncertainty visualisations represent uncertainty visually using symbols or

graphics rather than words or numbers (McDowell & Jacobs, 2017). Common visualisations of uncertainty include Euler diagrams, frequency grids, glyphs (discrete objects), hybrids of Euler circles and glyphs, and tree diagrams (see Figure 2 for examples).

These approaches all have their strengths and weaknesses. For instance VPEs are simple and flexible but also prone to misinterpretation (e.g. Budescu & Wallsten, 1995; Dhami & Mandel, 2020; Theil, 2002). Adding numbers tends to increase precision in understanding but perhaps relies on the communicators being proficient in mathematical reasoning (Dieckmann et al., 2012). Visualisations can help and are often spontaneously generated by both civilians and military personnel (King, 2006; B. Tversky et al., 2011; Zahner & Corter, 2010), but the precise format matters. Some visualisations have been found to aid reasoning, whilst others do not (Zahner & Corter, 2010). The general principle that appears to underlie these disparate results is that optimal decision-making is supported by formats that make uncertainty information easiest to use. As 'easiest' is highly subjective, this means the optimum communication format is considerably impacted by the attributes of the context, task, and communicators. The focus of the remainder of this section is on visualisation of uncertainty, which is often the common mode of representation of uncertainty that appears in military contexts (Chung & Wark, 2016).

Effectiveness of Visually Communicating Uncertainty

The task in the military scenario we have described relies on operators communicating information to commanders, for the latter to remain apprised of the spatial-temporal context on which tactical decisions are based (Carvalho et al., 2011; Hammond, 2006; John et al., 2000). Practically, commanders need to quickly understand what an entity is (status), where it is (position), where it is going (direction) as well as any uncertainty associated with these dimensions. The research below indicates that the uncertainty communicated by operators needs to be represented in a manner congruent to the dimension it is associated with.

Firstly, consider the uncertainty associated with entity status. This is an inherent property of the entity, and thus, the literature suggests that its uncertainty is best represented intrinsically, where features of the extant display are manipulated, such as changing the colour or shape (Kinkeldey et al., 2014, 2017). In contrast, extrinsic representations of uncertainty add new items to a data display. Finger and Bisantz (2002) found that participants were more efficient when only using icons that were degraded to a greater or lesser degree (representing uncertainty) rather than using them alongside numerical probability estimates, an additional extrinsic element. Kolbeinsson et al. (2015) and Bisantz et al. (2005) also found superior performance with representations that used intrinsic representations of uncertainty for status information.

Secondly, consider the assessment of the entity's position. Andre and Cutler (1998) showed that spatial uncertainty was communicated most reliably when it was represented spatially (i.e. matching the task at hand). In one of their tasks, participants navigated a spacecraft to a goal, whilst avoiding a meteor. The position of the meteor was noisy, and the authors explored representing this uncertainty in various formats. They found a spatial region of uncertainty around the target (an additional ring) was preferable to other representations (colour or text).

Thirdly, for the assessment of the entity's trajectory, again visualisations that matched the task enhanced communication of the uncertainty associated with the judgement. Andre and Cutler (1998) asked participants to control a gun turret and shoot down hostile, whilst avoiding friendly, entities. Here, uncertainty was associated with the direction of the craft. Participants who saw visual representations of this uncertainty were not as negatively affected when the uncertainty level was high compared to a version where the uncertainty was displayed numerically. Along the same lines, John et al. (2000) found that directional information was superior to text information when trying to predict where a unit would be in the future.

Thus, this literature suggests that optimal uncertainty communication in this context might be achieved by combining these representations of position, direction, and status. Indeed, this is frequently done in military contexts (see Figure 3 'NATO Joint Military Symbology', 2017).

Cognitive Abilities and Prior Experience. Visualisation formats vary in the amount they challenge cognitive abilities
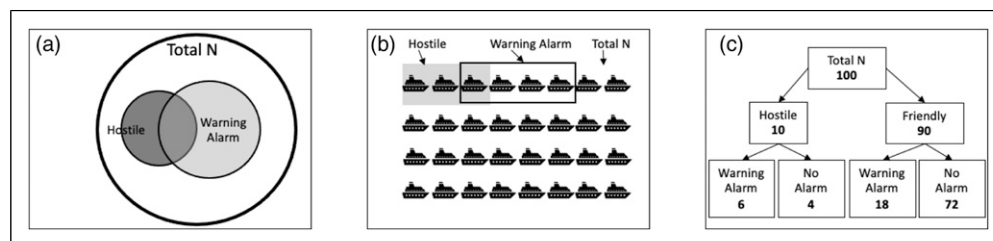

Figure 2. Examples of uncertainty communications showing the probability of potentially overlapping states in a military context. Here, the probability that a vessel is hostile (as opposed to Friendly), and the presence/absence of an automatic (but not perfect) warning alarm shown as (a) Euler diagram, (b) frequency grid and (c) tree diagram.
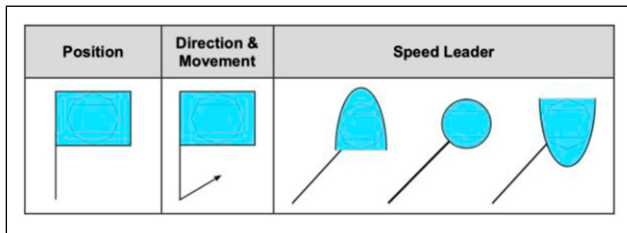
Figure 3. Example of a symbol that incorporates position, speed and direction in a single icon. Taken from 'NATO Joint Military Symbology' (2017). The direction of the arrow represents direction, and the length represents speed.

(Anderson et al., 2011). Bisantz et al. (2011) used a missile defence game where participants had to select and destroy missiles heading towards a city whilst not harming birds or planes. They found that adding numerical uncertainty information to the visualisations did not improve performance. Finger and Bisantz (2002) found similar results in a task where participants had to classify moving entities as either friendly or hostile (see also Bisantz et al., 2005). Others have found that including numbers can reduce the speed (Andre & Cutler, 1998; Finger & Bisantz, 2002; Hope & Hunter, 2007) and accuracy (Andre & Cutler, 1998) of decision making with uncertainty visualisations (for a review see Kinkeldey et al., 2017).

Some of the practical benefits of using visualisations can be directly tied to the fact they reduce cognitive load. Hegarty and Steinhoff (1997) found that participants with lower visual working memory capacity, compared to those with a higher working memory capacity, were helped more by diagrams when solving a mechanical reasoning problem. This suggests that the diagrams might perform as external or distributed memory. Additionally, visualisations, compared to verbal, auditory communications, allow for asynchronous communication. As mentioned above, maritime military decision makers are often not in the same place as those gathering and classifying the information needed for a decision. Thus, to obtain the required information, decision makers monitor both an auditory communication channel and visual displays. Performing a visual task can result in inattentional deafness, where people fail to hear auditory stimuli because of high visual perceptual load (Molloy et al., 2015). This suggests that serious problems may arise by decision makers missing key pieces of information from one or other of these sources of information. However, visualisations are easier to understand and change more slowly. Thus, decision makers can go back and double-check information they might have missed.

Thus, this literature suggests that by trying to include all the available uncertainty information simultaneously, we might increase the cognitive difficulty of the task beyond the point where communicating uncertainty is useful. However, people may be able to compensate for this limitation by using the representation as an external memory store. Indeed, this might explain why improvements associated with communication format are moderated by prior experience.

Kirschenbaum et al. (2014) looked at a submarine task where participants were asked to manoeuvre a submarine to torpedo another craft. Again, they found that participants performed better when positional uncertainty was represented spatially by a ring, rather than numerically in an associated table. However, in this experiment the benefit was limited to novices. This contrasts with previous evidence that found that experienced naval commanders also performed better when visual uncertainty information was included (Kirschenbaum & Arruda, 1994). One explanation for this apparent discrepancy is that the failure of spatial visualisations to improve performance for experts in Kirschenbaum et al. (2014) may be due a ceiling effect: the experts found the task easy, even with sub-optimum display formats.

Interpretation Errors and Biases. Although the evidence consistently suggests visualisations as the *medium* for communication in our scenario, the specific *format* of those visualisations still needs to be carefully considered. This is because in the literature there is no overall superior pictorial representation of uncertainties that has marked facilitative effects on decision-making performance (Binder et al., 2015; Böcherer-Linder & Eichler, 2017; Micallef et al., 2012; Spiegelhalter et al., 2011). For instance the use of icons (or glyphs) has been shown to be generally effective in some studies (Zikmund-Fisher et al., 2014) but not others (Sirota et al., 2014). There is also not clear evidence that Euler diagrams are reliably effective in generating superior decision making (Micallef et al., 2012; Sirota et al., 2015; Sloman et al., 2003). Similarly, the use of roulette wheels in some studies only show marginal improvements (Brase, 2014; Starns et al., 2019), whereas others show significant superior performance as compared to tree-diagrams and text-only representations (Yamagishi, 2003). Others again have shown that tree diagrams are extremely effective as compared to other visual aids and text-only information (Binder et al., 2015; Friederichs et al., 2014). This demonstrates the necessity to consider the requirements of the specific individual context when designing an appropriate visualisation.

The existence of cognitive biases is an important issue to consider when selecting a visualisation format. Here we consider three biases. *Deterministic construal error* is where someone misinterprets an uncertainty communication as a deterministic one (for alternatives see Gigerenzer et al., 2005; S. Joslyn & Savelli, 2010; Juanchich & Sirota, 2016; Morss et al., 2008, 2010; A. H. Murphy et al., 1980; Sink, 1995). The closely related *containment error* is where people perceive uncertain, fuzzy boundaries as representing fixed, deterministic points (Brown, 2004; Lundström et al., 2007; MacEachren et al., 2005; Ruginski et al., 2016). Finally, *anchoring* might guide the qualitative nature of such a construal (e.g. Broad et al., 2007; Oppenheimer et al., 2008; Tversky & Kahneman, 1973; Turner & Schley, 2016). Thus, the most effective visualisations will be those which minimises the likelihood of these errors being made.

Correcting biases. Identification of such biases can aid with the design of visualisations such that participants are guided towards the correct inferences and away from incorrect inferences (see also Card et al., 1999; Padilla, 2018). Some ways of doing this have begun to be explored. The basic approach is to use features of attentional processes to design visualisations for which questions about the underlying data can be answered more easily, more accurately or faster (Jänicke & Chen, 2010). In other words, designing the visual representation so that it matches the demands of the task (Vessey, 1991; 1994; 2006; Vessey & Galletta, 1991).

Human visual attention can be broadly defined as two processes: one bottom-up and stimulus-driven, the other top-down and goal-driven (Corbetta & Shulman, 2002; Petersen & Posner, 2012). *Bottom-up* attention is driven by external stimuli (Padilla, 2018) and is typically characterised as automatic, unconscious, and physiologically based (Connor et al., 2004). Visualisations can leverage bottom-up attention by manipulating the salience of visualisation components (Padilla, 2018). Visual salience is the degree to which an item stands out from those nearby (Jänicke & Chen, 2010). These items might 'pop-out' due to differences in colour, shape, orientation, size or movement (Fabrikant et al., 2010; Haroz & Whitney, 2012).

However, bottom-up attention is often modulated by top-down processes. *Top-down* attention involves deliberately searching for features in relation to a goal and is thought to be slow, conscious and effortful (Connor et al., 2004). Haider and Frensch (1999) found that task-redundant information tends to be ignored at the perceptual rather than the conceptual level. Visualisations that do not comply with existing schemas are generally less effective (Padilla, 2018), whereas those that respect commonly inferred meanings are more effective (Norman, 1988; B. Tversky et al., 2011). For example if designing a coloured 'danger' scale in the UK, it would be misleading to provide a key whereby green represented a dangerous location and red the safest location, since this is contrary to people's common expectations (Wogalter et al., 2002). However, the opposite would be true in China where red is more likely to have positive connotations (e.g. He, 2009; Yu, 2014).

The interaction of top-down and bottom-up processes can also improve how people interpret and use uncertainty information. Bisantz et al. (2009) tested the degree to which people agreed (with no instruction) on the mapping of different levels of saturation, brightness and transparency to uncertainty. They found good agreement between individuals that greater brightness, saturation, and opaqueness (i.e. more 'intense' colouring) was typically associated with more certain information. However, this result was found to be driven predominantly by the degree of contrast between the target object and the background, and to be qualified by the requirements of the task. Objects with greater contrast against the background were perceived as more relevant for the task at hand. That is if the task was to identify uncertain data points, greater contrast (and therefore visual saliency) was seen as representing greater uncertainty. This highlights the importance of considering the key requirements of the decision-maker (Doyle et al., 2019; Griethe, 2006; Hegarty et al., 2016; Kinkeldey et al., 2014; 2017; Loucks, 2003; Pang et al., 1997) who will be using the visualisation, such that the most task-relevant information can be highlighted within an uncertainty visualisation.

Communication: Recommendations, Limitations and Future Work

Is the communication literature useful in providing insights to help recommend how to represent uncertainty to military personnel? There are important insights from the vast literature on communication of uncertainty that can be applied to the military domain. The literature on effective communication suggests that the optimum format is one that best matches the information to be communicated. Thus, visualisations are likely to be superior for communicating the status, position and direction of surrounding entities to decision makers. In this maritime military context, visualisations are less likely to be misunderstood and additionally allow for asynchronous communication, thereby reducing the cognitive load associated with understanding incoming information.

Do laboratory experiments examining communication of uncertainty generalise to military contexts? The literature allows us to infer some general principles to be considered when designing displays for use in military contexts. However, the literature also suggests that it is very hard to predict exactly which factors are most important given a particular scenario, with a variety of personnel with different backgrounds. So, selecting the precise format these visualisations should take requires some finesse and further experimental study. Interpretation of visualisations can still be subject to cognitive biases, whether general pitfalls or those due to communicators' unique abilities or background. One approach to minimising misunderstanding is to take advantage of attentional processes. Visualisation designs need to make sure that both top-down and bottom-up attentional processes support optimal performance in tasks. This means that the information that is conceptually most important to the task should be the most salient (Vessey, 1991; 1994; 2006; Vessey & Galletta, 1991). This principle has shown to be effective in work by Hegarty et al. (2010). When trying to predict weather, participants were better able to use the (task-relevant) pressure information on a map when this information was highlighted with colour, rather than task-irrelevant temperature information. When these attentional processes are aligned, people will not have to work to inhibit irrelevant information.

Making Decisions

So far in our scenario, the operators have determined the statuses of surrounding entities and communicated this information, along with the associated uncertainty, to a key decision-maker (e.g. the Captain or their delegated Commander). In the final stage, the decision-maker must now decide what action to take (Szeligowski, 2018). They might decide to change course, engage an enemy craft, communicate with other vessels or do nothing (Lipshitz & Strauss, 1997). In the following, we review the literature relevant to making these types of dynamic decisions.

Dynamic Decision-Making

This scenario has much in common with tasks in the so-called dynamic decision-making literature (for reviews see Holt & Osman, 2017; Osman, 2010, 2011). Studies in this area are concerned with investigating how decision makers cope with an environment where events and their probabilities change over time. Dynamic decision-making is also iterative. It requires sequential decision-making, where each decision taken generates an outcome that requires a further decision to be made (making the decisions interdependent over time). The outcomes of the decisions change over time, their effects are cumulative and any changes experienced can result directly from the decisions made as well as independently of them (i.e. endogenous properties of the system; Brehmer, 1992; Dörner & Schaub, 1994; Holt & Osman, 2017; Osman, 2010). Thus, whether or not the changes experienced are stable or unstable (Osman, 2011), whether the outcomes of decisions are experienced in real time or only periodically (frequent vs. intermittent outcome feedback; Osman et al., 2017), whether or not the goals of the decision process are highly specific or general (Osman, 2008a), and whether or not the outcomes of the decisions taken also included additional information (augment positive vs. negative feedback, financial costs and benefits; Osman, 2012b); all of these features have implications for the effectiveness of the decisions made.

Here, we will discuss the literature about three key factors that affect decision-making in dynamic contexts. Firstly, how the goals of the task influences decision-making. Secondly, we consider the uncertainty inherent in the task. Finally, we discuss the role of feedback in dynamic tasks.

Goals. Research, both in dynamic and military decision-making, has shown that goals can have a substantial impact on the way people make decisions. Generally, studies have shown that in highly uncertain dynamic contexts, specifying a precise, static goal, such as controlling an outcome to precise criteria (for instance aiming for a chemical level of 5) leads to good decision-making performance in achieving that goal. However, when the decision-maker is required to adapt their decision-making to new goals in the same decision-making context, their performance suffers. In contrast, when the goals are broader to start with, then the decision-maker is better able to sample relevant information from the decision context to learn plans of actions to a variety of goals they might face in the future (Burns & Vollmeyer, 2002; Locke & Latham, 2006; Osman, 2008a; 2008b; 2008c; 2012a; Vollmeyer et al., 1996). Moreover, when the costs (e.g. financial penalties) in failing to reach the goal are high, this can lead to highly erroneous decisions over time, regardless of whether the dynamic changes to the outcome are stable or unstable. The reason for this is that the decision-maker is overly concerned with minimising the negative consequences which leads to sub-optimal sampling of the information and poorer decision-making over time (Kerstholt, 1996; Kerstholt & Raaijmakers, 1997).

These findings are echoed in the military literature. Indeed, goals can substantially influence situational awareness (van Westrenen & Praetorius, 2014). The military distinguishes between strategic, tactical and control decisions, each of which focus on different goals. Strategic decisions concern selecting the means to achieve a goal (e.g. ordering tugs), tactical decisions aim to deploy the means to achieve affordances (e.g. positioning the tugs or overtaking), and control decisions concern selecting the means to achieve a desired state (e.g. realising a speed and direction). Thus, the information a decision-maker needs to be aware of critically depends on the type of decision they are making.

However, unlike existing psychological research, military scenarios often have many competing goals at different levels. For instance typically the commander has been briefed on the overarching goals of the mission as well as more specific goals that their vessel is responsible for. In addition, they are responsible for the smooth running of the ship as well as other considerations such as limiting running costs. Thus, the psychological literature still has some ways to go before being able to completely understand how competing goals with varying stakes attached may impact decision-making in such a complex environment.

Uncertainty. Making good decisions in our dynamic decision-making scenario also requires dealing appropriately with uncertainty. Riveiro et al. (2014) examined dynamic decision-making in an air defence task. In this task, participants had to protect a radar station, whilst at the same time monitoring a map of the region to detect possible aerial threats. Practically, the decision makers had to identify and prioritise targets of interest to be communicated to a higher authority to determine the appropriate countermeasures. They found that uncertainty information helped participants make a final judgement more swiftly, although it had no effect on decision accuracy. In other words, in this task uncertainty appeared to make participants more confident their decisions.

However, in other military tasks, uncertainty appears to interact in unpredictable ways with expertise. Kobus et al. (2001) examined experts and non-experts in developing a battle plan in a dynamic tactical scenario. They found that the level of uncertainty interacted with experience in predicting

time to gain situational awareness and time to execute their plan. They found that for both high and low uncertainty conditions, experts took more time to gain situational awareness than non-experts. However, in the high uncertainty condition, experts were much faster to then execute their plan, although the level of experience made no difference when uncertainty was minimal. John et al. (2000) also found that increasing uncertainty influenced tactical decisions. However, in contrast to Kobus et al. (2001), they only found this for less experienced Marines in Combat Operation Centres: the *less* experienced officers tended to adopt a 'wait and see' approach showing greater levels of caution.

Perhaps part of the reason for these disparate results may come from how uncertainty is communicated to (and thus understood by) participants. These studies examine uncertainty using different tasks and thus, present the information in differing ways. This is an issue as many studies have found that the more salient the uncertainties around data are, the lower the confidence in decisions, and the greater the likelihood of making more conservative decisions (Finger & Bisantz, 2002; Riveiro et al., 2014; Svenson et al., 2010). Therefore, the differences between these studies could be driven by an interaction between the salience of information and expertise. After all, experts are obviously more experienced with the task. However, they may not have experience with how exactly the abstract research task is presented to them.

Dynamic tasks can also change their contingencies over time and there is preliminary evidence that participants are sensitive to this. When the fluctuations in outcome are moderately stable, decision makers tend to make small conservative changes to the system (Osman & Speekenbrink, 2011). In contrast, when experiencing large noisy fluctuations to the outcomes over time, decision makers often made multiple and dramatic changes simultaneously. It is important to note that, in these experiments, decision makers learn to make more effective decisions over time through extensive repeated exposure to the task environment.

In sum, the effect of uncertainty on performance varies with the attributes of the participants and the method by which uncertainty is communicated. These factors need to be explored in much greater depth before we can determine how best to facilitate optimum performance.

Feedback. Optimum decision-making requires decision makers to understand which options lead to which outcomes with what probabilities and select the option that draws them closest to their goal. Receiving feedback is an important part of learning these relationships. In several studies the findings show that outcome feedback (i.e. simply finding out the actual effects on the outcome experienced because of a decision taken) leads to better performance than the presence of either positive or negative feedback (i.e. value judgments attached to the outcomes; Osman, 2012a; Osman et al., 2017). The presence of positive or negative feedback has similar effects to decision makers being aware of the high stakes attached to their decisions, namely that they are overly focused on achieving positive feedback or reducing negative feedback than focussing on what information is relevant in achieving a desirable outcome over time. This aligns with a substantial body of literature looking at the role of feedback in many other decision-making contexts (for a review, see Kluger & DeNisi, 1996). This work suggests that unless the decision-making task is simple, additional feedback other than outcome feedback can, at best, add little additional benefits to decision-making performance, and at worst, impair decision-making performance.

The probabilistic nature of dynamic decision-making tasks adds an additional hurdle to interpreting feedback. Feedback on a single trial is not particularly diagnostic due to the large role chance may play on achieving an outcome. Rather, what is most important is the success of decisions *over time*. One way of adapting feedback to prevent overweighting of the outcomes on single trials is to give feedback intermittently. That is, providing feedback (as well as reward information, financial benefits, financial costs; or social rewards and costs) periodically at set intervals, rather than every time a decision is made. Unexpectedly, however, direct comparisons of the effect of intermittent versus consistent feedback show that overall dynamic decision-making performance suffered under intermittent outcome feedback (Osman et al., 2017). This is consistent with other work that shows intermittent feedback does not perform well, even with non-probabilistic outcomes (Le Pelley et al., 2019; Smith et al., 2014).

Decision-Making: Recommendations, Limitations and Future Work

Are laboratory studies on dynamic decision-making useful in providing insights to help recommend what strategies to adopt when facing military contexts? Work examining decision-making in dynamic contexts reveals that, whilst people can learn to manage uncertainty and to determine plans of action that require regular interaction with an environment that is constantly changing, there are factors that contribute to less effective decision-making. Directing decision makers to focus on achieving highly specific outcomes, especially in a decision-context that they are unfamiliar with, can overly constrain what they learn, and lead to less adaptive decision-making in the long run. Feedback has a significant role to play in guiding the way in which decision makers act in dynamic environments, because the environments often present a high level of uncertainty. Therefore, signals about the efficacy of decisions taken do provide useful guidance as to how to proceed, but can also overly constrain the focus of attention onto specific, narrow, features of the decision problem of the decision problem at the expense of other useful information.

Do laboratory experiments examining dynamic decision-making generalise to military contexts? Although the experimental research above suggests that participants can be

quite successful at dynamic decision-making tasks, there is one key (and likely obvious) difference between the experimental and military contexts that needs to be addressed. The dynamic decision-making scenarios are far simpler than the decision environments seen in military examples. For instance in Osman et al. (2017) participants manipulated the proportions of three inputs to control a single output. In contrast, military decision-making is only becoming more complex as the sophistication, complexity, volume and quality of information on which to base decisions increases (Mishra et al., 2015; Riveiro et al., 2014; Szeligowski, 2018; van Westrenen & Praetorius, 2014). This difference in complexity has implications for interpretation of the results. For example within a simplified laboratory setup it is reasonable simple to determine whether a particular decision was optimal or not. Outside the lab, it is much more difficult.

One approach to judging whether a decision was optimum or not is to look at the strategy that people use to come to their decision. Military decision makers are trained to weigh the pros and cons of each possible option and to select the alternative that leads to the best outcome (Kobus et al., 2001; Lipshitz & Strauss, 1997). However, the evidence suggests that often military decisions in the field are not optimal in this sense (Kobus et al., 2001; Riveiro et al., 2014). In a qualitative study exploring military decision-making, Kaempf et al. (1993) reported that 95% of operators use suboptimum strategies based on matching the situation to past experiences. However, it is difficult to tell whether this is an appropriate way of dealing with uncertainty. After all, decision makers are limited. They can be limited by gaps in knowledge, computation capacity (e.g. Frühling, 2014; van Westrenen & Praetorius, 2014; Yang et al., 2009), and motivational factors, such as time pressure (e.g. Frühling, 2014) and competing goals (Hammond, 2006; T. Murphy, 2010; van Westrenen & Praetorius, 2014). Further, given the dynamic nature of the environment, it may be the case that using a suboptimum strategy or heuristic that provides a quick answer may be better than carefully weighing the evidence, given that the scenario may change in that time. This is especially so as decisions in dynamic scenarios are interdependent: if the ship goes down because you failed to take prompt action, there will not be an opportunity to fix the mistake.

One possible way future research might attempt to judge the efficacy of decisions, and thereby improving training and understanding, is by using causal models. We know from the empirical literature is that causal representations are essential in developing plans of actions in learning, decision-making and reasoning (Bramley et al., 2017; Meder et al., 2014; Sloman & Lagnado, 2015). By mathematically modelling the causal relations of a problem, we can better determine the optimal solution. This may be especially important given that there are a host of biases and heuristics found even when studying decision making in static environment (Kahneman, 2011). In addition, fleshing out causal models has been shown to reduce bias (Krynski & Tenenbaum, 2007).

We could also use causal models as a decision-support tool. Much recent work in military decision-making talks about using decision-support tools. However, much of this work focuses on uncertainty visualisation, including uncertainty information so that it can be used in decision-making (e.g. Bisantz et al., 1999), rather than tools to help people make decisions. One approach to improving decision makers' understanding of the underlying causal structure of the system is to use visualisations to represent that system. Some researchers have begun to explore the effectiveness of diagrams that represent the available options and their associated probabilities in improving decision-making (e.g. Bae et al., 2019; Hänninen et al., 2014; Laskey et al., 2011; Pilato et al., 2012; Riveiro et al., 2014; Snidaro et al., 2015; Svenson et al., 2010; Zhang et al., 2008). Generally, these representations of the available options help decision makers improve their decision-making. However, like the visualisations mentioned above, they can still be subject to misinterpretations. Without training, the decision-maker is liable to impose their own understanding of the decision-problem and re-interpret the diagram to support their own preconceptions.

General Discussion

In this review, we aimed to fully explore the psychological processes involved in a complex decision-making task. Specifically, we focussed on a simplified military scenario: how information is gathered and used in short-term tactical decisions (Figure 1). To make the task tractable, we firstly divided this overarching scenario into three parts (classification, communication and choice) and then examined in detail the relevant literature at each stage. Like many other decision-making tasks, the complexity of the psychological processes involved in the scenario is demonstrated by the wide range of literature reviewed, from basic perceptual processes to the high-level influences of goal format. However, two core psychological factors repeatedly emerged as key factors in performance: cognitive load and individual differences.

Throughout the three stages, we consistently found that task performance was related to participants' ability to deal effectively with cognitive load. For instance in the communication literature the most successful uncertainty communications are those that minimise the amount of cognitive effort required for understanding (e.g. Andre & Cutler, 1998; Finger & Bisantz, 2002; Kirschenbaum et al., 2014). Participants also tend to ignore uncertainty information. Deterministic construal errors (Gigerenzer et al., 2005), containment (MacEachren et al., 2005) and anchoring (Broad et al., 2007) are all examples of biases where participants mistakenly treat uncertain information as certain and thereby reduce the amount of information they must consider. Further, adding more information does not necessarily improve understanding or performance (e.g. Andre & Cutler, 1998; Bisantz et al., 2011; Finger & Bisantz, 2002; Hope & Hunter, 2007; S. L. Joslyn & Grounds, 2015).

In categorisation and decision-making tasks, the role of cognitive load emerges as a tendency to satisfice: to use less of the available information than would be optimum (Simon, 1947). For instance in categorisation the evidence suggests participants are most likely to use simple rules based on a subset of the available information (e.g. Edmunds et al., 2015, 2018, 2019; Rehder & Hoffman, 2005a; Wills et al., 2015, 2020). Similarly, in dynamic decision-making participants will systematically vary one variable at a time to work out the underlying causal structure of a dynamic task (e.g. Osman & Speekenbrink, 2011; Osman et al., 2017). Therefore, at all levels of the scenario, people are likely to try and minimise the amount of information they use at any one time, at least when learning to figure out what is relevant in the task at hand.

There is nothing wrong with this approach. After all, expertise is often associated with subjectively evaluating the cognitive load of a task as less. However, by considering this tendency carefully we can leverage it to improve performance in these tasks in military contexts: both by using them to promote desirable behaviour and devising strategies to attenuate undesirable behaviours. Some ways of doing this have begun to be explored. For instance in the communication literature some have leveraged features of attentional processes to design visualisations for which questions about the underlying data can be answered more easily, more accurately or faster (Jänicke & Chen, 2010). In other words, designing the visual representation so that it matches the demands of the task (Vessey, 1991; 1994; 2006; Vessey & Galletta, 1991). This suggests that similar approaches may be fruitful in categorisation and decision-making experiments. For instance one could include explicit instructions to attend to a specific, randomly selected attribute every so often to avoid participants focussing on a small subset of the available information.

These results also suggest that cognitive load is likely to be a key factor to consider in other complex decision-making tasks. The literature we reviewed here suggests that one approach to improving performance is by leveraging any small opportunity to reduce load. Such an approach will be beneficial in any high stakes, high load decision making context. For instance the effectiveness of the compounding influence of numerous small changes (e.g. changing visual displays and alert sounds) to reduce load has been demonstrated within medical decision making (Phansalkar et al., 2010). Further, the literature reviewed here suggests that it is often difficult to know a priori which method of communication is most intuitive given a task. Thus, running (adequately powered, Bartlett et al., Under review) user studies that compare communication formats in scenarios as representative as possible is likely to improve final outcomes. For instance military exercises could be used as an opportunity to test different ways of representing information. There is a similar move in medical fields: virtual reality training is becoming more common and more realistic (Ruthenbeck & Reynolds, 2015).

Another tendency that has emerged from the literature is that task performance can vary greatly due to differences between individuals. Although overall there is a tendency to try and minimise the amount of information needed to complete a task, people vary in the information they select to use (e.g. Edmunds et al., 2015; Haider & Frensch, 1996). Similarly, the success of particular uncertainty communication formats often depends on attributes of the recipient such as their background (e.g. Brun & Teigen, 1988; Doupnik & Richter, 2003; Harris et al., 2013), expertise (e.g. Kirschenbaum & Arruda, 1994; Kirschenbaum et al., 2014; Willems et al., 2020), expectations (e.g. Norman, 1988; Padilla, 2018; B. Tversky et al., 2011; Wogalter et al., 2002), cognitive skills (e.g. Hegarty & Steinhoff, 1997). Thus, in some way, successful maritime military decision-making requires adapting to the individual differences of the personnel involved.

So, how do we overcome the difficulties raised by individual differences? Perhaps the most obvious answer is through training. Training is a core part of the military organisation and for the most part these training procedures are extremely effective. However, when it comes to decision-making there still appears to be a discrepancy between the strategies that military personnel are trained to use (Kobus et al., 2001; Lipshitz & Strauss, 1997) and those they actually use (Kaempf et al., 1993). Personnel are trained in how to determine the optimum option (Kobus et al., 2001; Lipshitz & Strauss, 1997). By weighing the pros and cons, personnel are likely to gain a deeper understanding of the situation and thereby, improve their decision-making (Osman & Speekenbrink, 2012). However, this strategy is not consistent with how people make decisions: personnel will very rarely have the time or resources to consider every possible action. Rather, the evidence suggests that they focus on key information and the similarity to past events and choose the best seeming option.

To some extent, the mismatch between the optimal strategy given in training and the ad-hoc strategies generated in the moment may be alleviated by many rounds of training operations. By repeating the same thing over and over, personnel can now use their imperfect strategy (matching to past experience), which will hopefully also match the optimum solution. However, this approach does not take into consideration that, due to their differences, people may interpret these training opportunities differently and thus, inadvertently learn different representations of the task. This suggests that perhaps more individualised training routines may help in making sure that military personnel are all on the same page. Indeed, work on developing more individualised training has begun in the decision-making literature. For instance Parpart et al. (2015) developed active learning algorithms that could better determine participants' strategies whilst completing the task. Future work could take these algorithms and use them to guide participants to the optimum solution.

Again, this suggests that individual differences need to be carefully considered in other complex decision-making tasks. The literature suggests that assuming that everyone uses the same approach (even broadly) in psychological tasks is something that needs to be demonstrated using appropriate evidence.

Other Factors

This review focused on the flow of information from the bottom of the hierarchy to the top. It implied that the end of the scenario occurs when the commander decides what to do. However, military hierarchies are not one-way systems; information also travels from the top to the bottom, usually via orders (Szeligowski, 2018). Thus, decision-making is not the end of the scenario, rather command is. In other words, a commander's success depends not only on making the optimum choice but also on implementing it effectively. This highlights a final key factor that influences behaviour in this scenario which we have avoided discussing in the current work: interpersonal dynamics. However, the scope of this literature is wide and would take too much space to explore here, although this is something critically important for future work. After all, making a good tactical decision but being unable to implement it because your crew has mutinied would in the end be equivalent to an incredibly poor decision.

Conclusion

To conclude, maritime military decision-making is an incredibly complex task. However, by drawing together diverse research from three different fields (categorisation, communicating uncertainty and dynamic decision making) we have gained a much greater understanding of the general tendencies that affect performance at each stage. Throughout the three tasks, the evidence shows that people generally try to minimise cognitive load by simplifying the complex environment around them, but that the way they do this varies between individuals. By identifying these two features of complex decision making, we can, in future, exploit them to improve performance in maritime military decision-making as well as other high-stakes scenarios.

Declaration of Conflicting Interests

Funding

ORCID iD

Charlotte E. R. Edmunds ⬤ https://orcid.org/0000-0002-0524-8756

References

Anderson, E. W., Potter, K. C., Matzen, L. E., Shepherd, J. F., Preston, G. A., & Silva, C. T. (2011). A user study of visualization effectiveness using EEG and cognitive load. In H. Hauser, H. Pfister, & J. J. van Wijk (Eds.), *Eurographics/IEEE symposium on visualization 2011 (eurovis 2011)* (*30*(3) pp. 791–800). https://doi.org/10.1111/j.1467-8659.2011.01928.x.

Andre, A. D., & Cutler, H. A. (1998). Displaying uncertainty in advanced navigation systems. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, *42*(1), 31–35. https://doi.org/10.1177/154193129804200108.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442–481. https://doi.org/10.1037/0033-295x.105.3.442.

Ashby, F. G., & Maddox, W. T. (2004). Human category learning. *Annual Review of Psychology*, *56*(1), 149–178. https://doi.org/10.1146/annurev.psych.56.091103.070217.

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*(1), 147–161. https://doi.org/10.1111/j.1749-6632.2010.05874.x.

Ashby, F. G., & Valentin, V. V. (2017). Multiple systems of perceptual category learning: Theory and cognitive tests. In H. Cohen, & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 157–188). Elsevier. https://doi.org/10.1016/b978-0-08-101107-2.00007-5.

Ashby, F. G., & Valentin, V. V. (2018). The categorization experiment: experimental design and data analysis. In J. T. Wixted (Ed.), *Stevens' Handbook of experimental psychology and cognitive neuroscience* (pp. 1–41). Wiley. https://doi.org/10.1002/9781119170174.epcn508.

Bae, J., Falkman, G., Helldin, T., & Riveiro, M. (2019). Visual data analysis. In A. Said, & V. Torra (Eds.), *Data science in practice. Studies in big data* (pp. 133–155). Springer. https://doi.org/10.1007/978-3-319-97556-6_8.

Bartlett, M. B., Edmunds, C. E. R., Belpaeme, T., & Thill, S. (Under review). *Have I got the power? A comparison of past and present reporting practice for replication in human-robot interaction*. Transactions on Human-Robot Interaction.

Binder, K., Krauss, S., & Bruckmaier, G. (2015). Effects of visualizing statistical information-an empirical study on tree diagrams and 2×2 tables. *Frontiers in Psychology*, *6*(3), 1–9. https://doi.org/10.3389/fpsyg.2015.01186.

Bisantz, A. M., Cao, D., Jenkins, M., Pennathur, P. R., Farry, M., Roth, E., Potter, S. S., & Pfautz, J. (2011). Comparing uncertainty visualizations for a dynamic decision-making task. *Journal of Cognitive Engineering and Decision Making*, *5*(3), 277–293. https://doi.org/10.1177/1555343411415793.

Bisantz, A. M., Finger, R., Seong, Y., & Llinas, J. (1999). Human performance and data fusion based decision aids. Proceedings

of the 2nd international conference on information fusion, Sunnyvale, CA

Bisantz, A. M., Marsiglio, S. S., & Munch, J. (2005). Displaying uncertainty: Investigating the effects of display format and specificity. *Human Factors*, 47(4), 777–796. https://doi.org/10.1518/001872005775570916.

Bisantz, A. M., Stone, R. T., Pfautz, J., Fouse, A., Farry, M., Roth, E., Nagy, A. L., & Thomas, G. (2009). Visual representations of meta-information. *Journal of Cognitive Engineering and Decision Making*, 3(1), 67–91. https://doi.org/10.1518/155534309x433726.

Böcherer-Linder, K., & Eichler, A. (2017). The impact of visualizing nested sets. An empirical study on tree diagrams and unit squares. *Frontiers in Psychology*, 7, 26. https://doi.org/10.3389/fpsyg.2016.02026.

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing Neurath's ship: Approximate algorithms for online causal learning. *Psychological Review*, 124(3), 301–338. https://doi.org/10.1037/rev0000061.

Brase, G. L. (2014). The power of representation and interpretation: Doubling statistical reasoning performance with icons and frequentist interpretations of ambiguous numbers. *Journal of Cognitive Psychology*, 26(1), 81–97. https://doi.org/10.1080/20445911.2013.861840.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81(3), 211–241. https://doi.org/10.1016/0001-6918(92)90019-a.

Broad, K., Leiserowitz, A., Weinkle, J., & Steketee, M. (2007). Misinterpretations of the "cone of uncertainty" in Florida during the 2004 hurricane season. *Bulletin of the American Meteorological Society*, 88(5), 651–668. https://doi.org/10.1175/bams-88-5-651.

Brown, R. (2004). Animated visual vibrations as an uncertainty visualisation technique. Proceedings International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia, Singapore, 15–18 June, 2004, pp. 84–89. https://doi.org/10.1145/988834.988849.

Brun, W., & Teigen, K. H. (1988). Verbal probabilities: Ambiguous, context-dependent, or both? *Organizational Behaviour and Human Decision Processes*, 41(3), 390–404. https://doi.org/10.1016/0749-5978(88)90036-2.

Budescu, D. V., & Wallsten, T. S. (1995). Processing linguistic probabilities: General principles and empirical evidence. *Psychology of Learning and Motivation*, 32(2), 275–318. https://doi.org/10.1016/s0079-7421(08)60313-8.

Burns, B. D., & Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *Quarterly Journal of Experimental Psychology*, 55(1), 241-261. https://doi.org/10.1080/02724980143000262.

Cantwell, G., Crossley, M. J., & Ashby, F. G. (2015). Multiple stages of learning in perceptual categorization: Evidence and neurocomputational theory. *Psychological Bulletin & Review*, 22(6), 1598–1613. https://doi.org/10.3758/s13423-015-0827-2.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (Eds.). (1999). *Readings in information visualization: Using vision to think.*

Morgan Kaufmann Publishers Inc. https://doi.org/10.5555/300679.

Carvalho, R. N., Haberlin, R., Costa, P. C. G., Laskey, K. B., & Chang, K., (2011). Modeling a probabilistic ontology for maritime domain awareness. In Proceedings of the 14th international conference on information fusion (p. 1-8).

Casale, M. B., Roeder, J. L., & Ashby, F. G. (2012). Analogical transfer in perceptual categorization. *Memory & Cognition*, 40(3), 434–449. https://doi.org/10.3758/s13421-011-0154-4.

Chung, J., & Wark, S. (2016). *Visualising uncertainty for decision support*. Department of Science and Technology. https://www.dst.defence.gov.au/sites/default/files/publications/documents/DST-Group-TR-3325.pdf.

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology*, 14(19), R850–R852. https://doi.org/10.1016/j.cub.2004.09.041.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201–215. https://doi.org/10.1038/nrn755.

Cummings, M. L., Buchin, M., Carrigan, G., & Donmez, B. (2010). Supporting intelligent and trustworthy maritime path planning decisions. *International Journal of Human Computer Studies*, 68(10), 616–626. https://doi.org/10.1016/j.ijhcs.2010.05.002.

Dhami, M. K., & Mandel, D. R. (2020). Words or numbers? Communicating probability in intelligence analysis. *American Psychologist*, 76(3), 549–560. https://doi.org/10.31234/osf.io/zkvcf.

Dieckmann, N. F., Peters, E., Gregory, R., & Tusler, M. (2012). Making sense of uncertainty: Advantages and disadvantages of providing an evaluative structure. *Risk Analysis*, 15(7), 717–735. https://doi.org/10.1080/13669877.2012.666760.

Dörner, D., & Schaub, H. (1994). Errors in planning and decision-making and the nature of human information processing. *Applied Psychology*, 43(4), 433–453. https://doi.org/10.1111/j.1464-0597.1994.tb00839.x.

Doupnik, T. S., & Richter, M. (2003). Interpretation of uncertainty expressions. *Accounting, Organizations and Society*, 28(1), 15–35. https://doi.org/10.1016/s0361-3682(02)00010-7.

Doyle, E. E. H., Johnston, D. M., Smith, R., & Paton, D. (2019). Communicating model uncertainty for natural hazards: A qualitative systematic thematic review. *International Journal of Disaster Risk Reduction*, 33(5), 449–476. https://doi.org/10.1016/j.ijdrr.2018.10.023.

Edmunds, C. E. R., Inkster, A. B., Jones, P. M., Milton, F., & Wills, A. J. (2020). Absence of cross-modality analogical transfer in perceptual categorization. *Open Journal of Experimental Psychology and Neuroscience*, 1(2), 3–13. https://doi.org/10.46221/ojepn.2020.8639.

Edmunds, C. E. R., Milton, F., & Wills, A. J. (2015). Feedback can be superior to observational training for both rule-based and information-integration category structures. *Quarterly Journal of Experimental Psychology*, 68(6), 1203–1222. https://doi.org/10.1080/17470218.2014.978875.

Edmunds, C. E. R., Milton, F., & Wills, A. J. (2018). Due process in dual process: Model-recovery simulations of decision-bound

strategy analysis in category learning. *Cognitive Science*, *42*(Suppl 3), 833–860. https://doi.org/10.1111/cogs.12607.

Edmunds, C. E. R., & Wills, A. J. (2016). Modeling category learning using a dual system approach: A simulation of Shepard, Hovland and Jenkins (1961) by COVIS. Proceedings of 38th Annual Conference of the Cognitive Science Society, Philadelphia, PA, 10–13 August, 2016

Edmunds, C. E. R., Wills, A. J., & Milton, F. (2016). Memory for exemplars in category learning. Proceedings of the 38th Annual Conference of the Cognitive Science Society, Philadelphia, PA, 10–13 August, 2016, pp. 2243–2248.

Edmunds, C. E. R., Wills, A. J., & Milton, F. (2019). Initial training with difficult items does not facilitate category learning. *Quarterly Journal of Experimental Psychology*, *72*(2), 151–167. https://doi.org/10.1080/17470218.2017.1370477.

Fabrikant, S. I., Hespanha, S. R., & Hegarty, M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making. *Annals of the Association of American Geographers*, *100*(1), 13–29. https://doi.org/10.1080/00045600903362378.

Finger, R., & Bisantz, A. M. (2002). Utilizing graphical formats to convey uncertainty in a decision-making task. *Theoretical Issues in Ergonomics Science*, *3*(1), 1–25. https://doi.org/10.1080/14639220110110324.

Friederichs, H., Ligges, S., & Weissenstein, A. (2014). Using tree diagrams without numerical values in addition to relative numbers improves students' numeracy skills. *Medical Decision Making*, *34*(2), 253–257. https://doi.org/10.1177/0272989x13504499.

Frühling, S. (2014). *Defence planning and uncertainty*. Routledge. https://doi.org/10.4324/9781315818764.

Gigerenzer, G., Hertwig, R., van den Broek, E., Fasolo, B., & Katsikopoulos, K. V. (2005). "A 30tomorrow": How does the public understand probabilistic weather forecasts? *Risk Analysis*, *25*(3), 623–629. https://doi.org/10.1111/j.1539-6924.2005.00608.x.

Goodwin, G. F., Blacksmith, N., & Coats, M. R. (2018). The science of teams in the military: Contributions from over 60 years of research. *American Psychologist*, *73*(4), 322–333. https://doi.org/10.1037/amp0000259.

Griethe, S. H. H. (2006). *The visualization and of uncertain and data: Methods and problems*. Simvis.

Grinband, J., Hirsch, J., & Ferrera, V. P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron*, *49*(5), 757–763. https://doi.org/10.1016/j.neuron.2006.01.032.

Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition. *Cognitive Psychology*, *30*(3), 304–337. https://doi.org/10.1006/cogp.1996.0009.

Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 172–190. https://doi.org/10.1037/0278-7393.25.1.172.

Hammond, M. (2006). *The implications of self-reporting systems for maritime domain awareness*. Defence Research and Development.

Hänninen, M., Banda, O. A. V., & Kujala, P. (2014). Bayesian network model of maritime safety management. *Expert Systems with Applications*, *41*(17), 7837–7846. https://doi.org/10.1016/j.eswa.2014.06.029.

Haroz, S., & Whitney, D. (2012). How capacity limits of attention influence information visualization effectiveness. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2402–2410. https://doi.org/10.1109/tvcg.2012.233.

Harris, A. J. L., Corner, A., Xu, J., & Du, X. (2013). Lost in translation? Interpretations of the probability phrases used by the intergovernmental panel on climate change in china and the UK. *Climatic Change*, *121*(2), 415–425. https://doi.org/10.1007/s10584-013-0975-1.

He, G. (2009). English and Chinese cultural connotation of color words in comparison. *Asian Social Science*, *5*(7), 160–163. https://doi.org/10.5539/ass.v5n7p160.

Hegarty, M., Canham, M. S., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(7), 37–53. https://doi.org/10.1037/a0017683.

Hegarty, M., Friedman, A., Boone, A. P., & Barrett, T. J. (2016). Where are you? The effect of uncertainty and its visual representation on location judgments in GPS-like displays. *Journal of Experimental Psychology: Applied*, *22*(4), 381–392. https://doi.org/10.1037/xap0000103.

Hegarty, M., & Steinhoff, K. (1997). Individual differences in use of diagrams as external memory in mechanical reasoning. *Learning and Individual Differences*, *9*(1), 19–42. https://doi.org/10.1016/s1041-6080(97)90018-2.

Helie, S., Ell, S. W., Filoteo, J. V., & Maddox, W. T. (2015). Criterion learning in rule-based categorization: Simulation of neural mechanism and new data. *Brain and Cognition*, *95*, 19–34. https://doi.org/10.1016/j.bandc.2015.01.009.

Holmes, R. (Ed.). (2001). *The Oxford companion to military history*. Oxford University Press. https://doi.org/10.1093/acref/9780198606963.001.0001.

Holt, D. V., & Osman, M. (2017). Approaches to cognitive modeling in dynamic systems control. *Frontiers in Psychology*, *8*, 2032. https://doi.org/10.3389/fpsyg.2017.02032.

Hope, S., & Hunter, G. J. (2007). Testing the effects of positional uncertainty on spatial decision-making. *International Journal of Geographical Information Science*, *21*(6), 645–665. https://doi.org/10.1080/13658810601073273.

Irandoust, H., Benaskeur, A., Kabanza, F., & Bellefeuille, P. (2010). A mixed-initiative advisory system for threat evaluation. Proceedings of the 15th international command and control research and technology symposium: The evolution of c2, Santa Monica, CA, 22–24 June, 2010. (Vol. *2*, p. 2-20).

Jänicke, H., & Chen, M. (2010). A salience-based quality metric for visualization. *Eurographics/IEEE-VGTC Symposium on Visualization 2010*, *29*(3), 1183–1192. https://doi.org/10.1111/j.1467-8659.2009.01667.x.

John, M. S., Callan, J., Proctor, S., & Holste, S. T. (2000). *Tactical decision-making under uncertainty: Experiments i and ii.*

(Tech. Rep.). SPAWAR Space and Naval Warfare Systems Center. https://doi.org/10.21236/ada378170.

Joslyn, S. L., & Grounds, M. A. (2015). The use of uncertainty forecasts in complex decision tasks and various weather conditions. *Journal of Experimental Psychology: Applied*, *21*(4), 407–417. https://doi.org/10.1037/xap0000064.

Joslyn, S., & Savelli, S. (2010). Communicating forecast uncertainty: Public perception of weather forecast uncertainty. *Meteorological Applications*, *17*(2), 180–195. https://doi.org/10.1002/met.190.

Joslyn, S., Savelli, S., & Nadav-Greenberg, L. (2011). Reducing probabilistic weather forecasts to the worst-case scenario: Anchoring effects. *Journal of Experimental Psychology: Applied*, *17*(4), 342–353. https://doi.org/10.1037/a0025901.

Juanchich, M., & Sirota, M. (2016). How to improve people's interpretation of probabilities of precipitation. *Journal of Risk Research*, *19*(3), 388–404. https://doi.org/10.1080/13669877.2014.983945.

Kaempf, G. L., Wolf, S., & Miller, T. E. (1993). Decision making in the aegis combat information center In *Proceedings of the human factors and ergonomics society annual meeting* (*37*, pp. 1107–1111). SAGE Publications. https://doi.org/10.1177/154193129303701615.16

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior*, *23*(6), 734–759. https://doi.org/10.1016/s0022-5371(84)90442-0.

Kerstholt, J. (1996). The effect of information costs on strategy selection in dynamic tasks. *Acta Psychologica*, *94*(3), 273–290. https://doi.org/10.1016/s0001-6918(96)00011-x.

Kerstholt, J. H., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In R. Ranyard, W. Crozier, & O. Svenson (Eds.), *Decision making: Cognitive models and explanations*. Routledge.

King, A. (2006). The word of command. *Armed Forces & Society*, *32*(4), 493–512. https://doi.org/10.1177/0095327x05283041.

Kinkeldey, C., MacEachren, A. M., Riveiro, M., & Schiewe, J. (2017). Evaluating the effect of visually represented geodata uncertainty on decision-making: Systematic review, lessons learned, and recommendations. *Cartography and Geographic Information Science*, *44*(1), 1–21. https://doi.org/10.1080/15230406.2015.1089792.

Kinkeldey, C., MacEachren, A. M., & Schiewe, J. (2014). How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies. *The Cartographic Journal*, *51*(4), 372–386. https://doi.org/10.1179/1743277414y.0000000099.

Kirschenbaum, S. S., & Arruda, J. E. (1994). Effects of graphic and verbal probability information on command decision making. *Human Factors*, *36*(3), 406–418. https://doi.org/10.1177/001872089403600302.

Kirschenbaum, S. S., Trafton, J. G., Schunn, C. D., & Trickett, S. B. (2014). Visualizing uncertainty: The impact on performance.

Human Factors, *56*(3), 509–520. https://doi.org/10.1177/0018720813498093.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Review*, *119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254.

Kobus, D. A., Proctor, S., & Holste, S. (2001). Effects of experience and uncertainty during dynamic decision making. *Proceedings of the IEA 2000/hfes 2000 Congress*, *28*(5), 275–290. https://doi.org/10.1016/s0169-8141(01)00022-1.

Kolbeinsson, A., Falkman, G., & Lindblom, J. (2015). Showing uncertainty in aircraft cockpits using icons. *Procedia Manufacturing*, *3*(5), 2905–2912. https://doi.org/10.1016/j.promfg.2015.07.805.

Kurtz, K. J. (2015). Human category learning: Toward a broader explanatory account. In B. H. Ross (Ed.), *Psychology of learning and motivation* (*63*, pp. 77–114). Elsevier. https://doi.org/10.1016/bs.plm.2015.03.001.

Laskey, K. B., Haberlin, R., Carvalho, R. N., & da Costa, P. C. G. (2011). PR-OWL 2 case study: A maritime domain probabilistic ontology. Proceedings of the Sixth International Conference on Semantic Technologies for Intelligence, Defense, and Security, Fairfax, VA, 16–17 November, 2011. (pp. 76–83).

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin*, *142*(10), 1111–1140. https://doi.org/10.1037/bul0000064.

Le Pelley, M. E. L., Newell, B. R., & Nosofsky, R. M. (2019). Deferred feedback does not dissociate implicit and explicit category-learning systems: Commentary on smith et al. (2014). *Psychological Science*, *30*(9), 1403–1409. https://doi.org/10.1177/0956797619841264.

Liebhaber, M. J., & Feher, B. A. (2002). *Surface warfare threat assessment: Requirements definition*. (Tech. Rep.). SPAWAR Systems Center San Diego. https://doi.org/10.21236/ada406337.

Liebhaber, M. J., Kobus, D. A., & Smith, C. A. P. (2000). *Empirical investigation of air threat assessment*. (Tech. Rep.). Pacific Science and Engineering Group.

Lipshitz, R., & Strauss, O. (1997). Coping with uncertainty: A naturalistic decision-making analysis. *Organizational Behavior and Human Decision Processes*, *69*(2), 149–163. https://doi.org/10.1006/obhd.1997.2679.

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*, *15*(5), 265–268. https://doi.org/10.1111/j.1467-8721.2006.00449.x.

Loucks, D. P. (2003). Quantifying and communicating model uncertainty for decision-making in the everglades. In Y. Y. Haimes, D. A. Moser, & E. Z. Stackhiv (Eds.), *Risk-based decision making in water resources x* (pp. 40–58). American Society of Civil Engineers. https://doi.org/10.1061/40694(2003)4.

Lundström, C., Ljung, P., Persson, A., & Ynnerman, A. (2007). Uncertainty visualization in medical volume rendering using probabilistic animation. *IEEE Transactions on Visualization*

and Computer Graphics, *13*(6), 1648–1655. https://doi.org/10.1109/TVCG.2007.70518.

MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., & Hetzler, E. (2005). Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science*, *32*(3), 139–160. https://doi.org/10.1559/1523040054738936.

Matthews, M. L., Rehak, L., Lapinski, A.-L. S., & McFadden, S. (2009). Improving the maritime surface picture with a visualization aid to provide rapid situation awareness of information uncertainty. 2009 IEEE Toronto International Conference Science and Technology for Humanity, Toronto, Canada, 26–27 September, 2009, pp. 533–538. https://doi.org/10.1109/tic-sth.2009.5444441.

McCloskey, M. J. (1996). An analysis of uncertainty in the marine corps. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *40*(4), 194–198. https://doi.org/10.1177/154193129604000411.

McDowell, M., & Jacobs, P. (2017). Meta-analysis of the effect of natural frequencies on Bayesian reasoning. *Psychological Bulletin*, *143*(12), 1273–1312. https://doi.org/10.1037/bul0000126.

Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological Review*, *121*(3), 277–301. https://doi.org/10.1037/a0035944.

Micallef, L., Dragicevic, P., & Fekete, J.-D. (2012). Assessing the effect of visualizations on Bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Scientific Visualization/Information Visualization 2012)*, *18*(12), 2536–2545. https://doi.org/10.1109/tvcg.2012.199.

Mishra, M., Sidoti, D., Ayala, D. F. M., Han, X., Avvari, G. V., Zhang, L., & Kleinman, D. L. (2015). Dynamic resource management and information integration for proactive decision support and planning. 2015 18th international conference on information fusion (fusion), Online, 07 June, 2015–07 September, 2015

Molloy, K., Griffiths, T. D., Chait, M., & Lavie, N. (2015). Inattentional deafness: Visual load leads to time-specific suppression of auditory evoked responses. *The Journal of Neuroscience*, *35*(49), 16046–16054. https://doi.org/10.1523/jneurosci.2931-15.2015.

Morss, R. E., Demuth, J. L., & Lazo, J. K. (2008). Communicating uncertainty in weather forecasts: A survey of the U.S. public. *Weather and Forecasting*, *23*(5), 974–991. https://doi.org/10.1175/2008waf2007088.1.

Morss, R. E., Lazo, J. K., & Demuth, J. L. (2010). Examining the use of weather forecasts in decision scenarios: Results from a us survey with implications for uncertainty communication. *Meteorological Applications*, *17*(2), 149–162. https://doi.org/10.1002/met.196.

Murphy, T. (2010). Security challenges in the 21st century global commons. *Yale Journal of International Affairs*.

Murphy, A. H., Lichtenstein, S., Fischhoff, B., & Winkler, R. L. (1980). Misinterpretations of precipitation probability forecasts. *Bulletin American Meteorological Society*, *61*(7), 695–701. https://doi.org/10.1175/1520-0477(1980)061<0695:moppf>2.0.co;2.

NATO Joint Military Symbology. (2017). [Computer software manual].

Navarro, D. J., Perfors, A., & Vong, W. K. (2013). Learning time-varying categories. *Memory & Cognition*, *41*(6), 917–927. https://doi.org/10.3758/s13421-013-0309-6.

Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of category learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 167–215). Academic Press. https://doi.org/10.1016/b978-0-12-385527-5.00006-1.

Norman, D. A. (1988). *The psychology of everyday things*. Basic Books.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–61. https://doi.org/10.1037//0096-3445.115.1.39.

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., Mckinley, S. C., & Glauthier, P. (1994). Comparing modes of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, *22*(3), 352–369. https://doi.org/10.3758/bf03200862.

Oppenheimer, D. M., LeBoeuf, R. A., & Brewer, N. T. (2008). Anchors aweigh: A demonstration of cross-modality anchoring and magnitude priming. *Cognition*, *106*(1), 13–26. https://doi.org/10.1016/j.cognition.2006.12.008.

Osman, M. (2008a). Observation can be as effective as action in problem solving. *Cognitive Science*, *32*(1), 162–183. https://doi.org/10.1080/03640210701703683.

Osman, M. (2008b). Positive transfer and negative transfer/antilearning of problem-solving skills. *Journal of Experimental Psychology: General*, *137*(1), 97–115. https://doi.org/10.1037/0096-3445.137.1.97.

Osman, M. (2008c). Seeing is as good as doing. *The Journal of Problem Solving*, *2*(1), 29–40. https://doi.org/10.7771/1932-6246.1029.

Osman, M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin*, *136*(1), 65–86. https://doi.org/10.1037/a0017815.

Osman, M. (2011). *Controlling uncertainty: Decision making and learning in complex worlds*. John Wiley & Sons. https://doi.org/10.1002/9781444328226.

Osman, M. (2012a). The effects of self set or externally set goals on learning in an uncertain environment. *Learning and Individual Differences*, *22*(5), 575–584. https://doi.org/10.1016/j.lindif.2011.09.012.

Osman, M. (2012b). The role of reward in dynamic decision making. *Frontiers in Neuroscience*, *6*(1), 35. https://doi.org/10.3389/fnins.2012.00035.

Osman, M., Glass, B. D., Hola, Z., & Stollewerk, S. (2017). Reward and feedback in the control over dynamic events. *Psychology*, *08*(07), 1063–1089. https://doi.org/10.4236/psych.2017.87070.

Osman, M., & Speekenbrink, M. (2011). Cue utilization and strategy application in stable and unstable dynamic environments.

Cognitive Systems Research, *12*(3–4), 355–364. https://doi.org/10.1016/j.cogsys.2010.12.004.

Osman, M., & Speekenbrink, M (2012). Prediction and control in a dynamic environment. *Frontiers in Psychology*, *3*(2), 1–12. https://doi.org/10.3389/fpsyg.2012.00068.

Padilla, L. M. (2018). A case for cognitive models in visualization research: Position paper. 2018 IEEE Evaluation and Beyond - Methodological Approaches for Visualization, Berlin, Germany, 21 October, 2018. https://doi.org/10.1109/beliv.2018.8634267.

Pang, A. T., Wittenbrink, C. M., & Lodha, S. K. (1997). Approaches to uncertainty visualization. *Visual Computer*, *13*(8), 370–390. https://doi.org/10.1007/s003710050111.

Parpart, P., Schulz, E., Speekenbrink, M., & Love, B. C. (2015). Active learning as a means to distinguish among prominent decision strategies. Proceedings of the Cognitive Science Society, Pasadena, CA, 22–25 July, 2015

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*(1), 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525.

Phansalkar, S., Edworthy, J., Hellier, E., Seger, D. L., Schedlbauer, A., Avery, A. J., & Bates, D. W. (2010). A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. *Journal of the American Medical Informatics Association*, *17*(5), 493–501. https://doi.org/10.1136/jamia.2010.005264.

Pilato, G., Augello, A., Missikoff, M., & Taglino, F. (2012). Integration of ontologies and Bayesian networks for maritime situation awareness. 2012 IEEE Sixth International Conference on Semantic Computing, Palermo, Italy, 19–21 September, 2012. https://doi.org/10.1109/icsc.2012.25.

Pizer, W. A. (1999). The optimal choice of climate change policy in the presence of uncertainty. *Resource and Energy Economics*, *21*(3–4), 255–287. https://doi.org/10.1016/s0928-7655(99)00005-6.

Pothos, E. M., & Wills, A. J. (Eds.). (2011). *Formal approaches in categorization*. Cambridge University Press. https://doi.org/10.1017/cbo9780511921322.

Potter, K., Rosen, P., & Johnson, C. R. (2012). *Expanding the frontiers of visual analytics and visualization* (pp. 226–249). Springer. https://doi.org/10.1007/978-3-642-32677-6_15.

Rehder, B., & Hoffman, A. B. (2005a). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*(1), 1–41. https://doi.org/10.1016/j.cogpsych.2004.11.001.

Rehder, B., & Murphy, G. L. (2003). A knowledge-resonance (KRES) model of category learning. *Psychonomic Bulletin & Review*, *10*(4), 759–784. https://doi.org/10.3758/bf03196543.

Riveiro, M., Helldin, T., Falkman, G., & Lebram, M. (2014). Effects of visualizing uncertainty on decision-making in a target identification scenario. *Computers & Graphics*, *41*(1), 84–98. https://doi.org/10.1016/j.cag.2014.02.006.

Riveiro, M., Pallotta, G., & Vespe, M. (2018). Maritime anomaly detection: A review. Wiley. *Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(5), e1266. https://doi.org/10.1002/widm.1266.

Ruginski, I. T., Boone, A. P., Padilla, L. M., Liu, L., Heydari, N., Kramer, H. S., Hegarty, M., Thompson, W. B., House, D. H., & Creem-Regehr, S. H. (2016). Non-expert interpretations of hurricane forecast uncertainty visualizations. *Spatial Cognition & Computation*, *16*(2), 154–172. https://doi.org/10.1080/13875868.2015.1137577.

Ruthenbeck, G. S., & Reynolds, K. J. (2015). Virtual reality for medical training: The state-of-the-art. *Journal of Simulation*, *9*(1), 16–26. https://doi.org/10.1057/jos.2014.14.

Schlegelmilch, R., & von Helversen, B. (2020). The influence of reward magnitude on stimulus memory and stimulus generalization in categorization decisions. *Journal of Experimental Psychology: General*, *149*(10), 1823–1854. https://doi.org/10.1037/xge0000747.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13), 1–42. https://doi.org/10.1037/h0093825.

Simon, H. A. (1947). *Administrative behavior* (*41*, p. 1215). Macmillan. https://doi.org/10.2307/1950596.6

Sink, S. A. (1995). *Determining the public's understanding of precipitation forecasts; Results of a survey*. National Weather Digest.

Sirota, M., Kostovičová, L., & Juanchich, M. (2014). The effect of iconicity of visual displays on statistical reasoning: evidence in favour of the null hypothesis. *Psychonomic Bulletin & Review*, *21*(4), 961–968. https://doi.org/10.3758/s13423-013-0555-4.

Sirota, M., Kostovičová, L., & Vallée-Tourangeau, F. (2015). Now you bayes, now you don't: Effects of set-problem and frequency-format mental representations on statistical reasoning. *Psychological Bulletin & Review*, *22*(5), 1465–1473. https://doi.org/10.3758/s13423-015-0810-y.

Sloman, SA, & Lagnado, D (2015). Causality in thought. *Annual Review of Psychology*, *66*(1), 223–247. https://doi.org/10.1146/annurev-psych-010814-015135.

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, *91*(2), 296–309. https://doi.org/10.1016/s0749-5978(03)00021-9.

Smith, J. D., Boomer, J., Zakrzewski, A. C., Roeder, J. L., Church, B. A., & Ashby, F. G. (2014). Deferred feedback sharply dissociates implicit and explicit category learning. *Psychological Science*, *25*(2), 447–457. https://doi.org/10.1177/0956797613509112.

Smith, J. D., & Church, B. A. (2018). Dissociable learning processes in comparative psychology. *Psychonomic Bulletin & Review*, *25*(5), 1565–1584. https://doi.org/10.3758/s13423-017-1353-1.

Smith, J. D., & Kemler, D. G. (1984). Overall similarity in adults' classification: The child in all of us. *Journal of Experimental Psychology: General*, *113*(1), 137–159. https://doi.org/10.1037/0096-3445.113.1.137.

Snidaro, L., Garcia, J., & Corchado, J. M. (2015). Context-based information fusion. *Information Fusion*, *21*(1), 82–84. https://doi.org/10.1016/j.inffus.2014.02.001.

Song, A. A., & Kleinman, D. L. (1994). A distributed simulation system for team decision-making. In *Fifth annual conference on AI, and planning in high autonomy systems* (pp. 129–135). IEEE. https://doi.org/10.1109/aihas.1994.390491.

Soto, F., & Ashby, F. G. (2019). Novel representations that support rule-based categorization are acquired on-the-fly during category learning. *Psychological Research*, *83*(3), 544–566. https://doi.org/10.1007/s00426-019-01157-7.

Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science*, *333*(6048), 1393–1400. https://doi.org/10.1126/science.1191181.

Starns, J. J., Cohen, A. L., Bosco, C., & Hirst, J. (2019). A visualization technique for Bayesian reasoning. *Applied Cognitive Psychology*, *33*(2), 234–251. https://doi.org/10.1002/acp.3470.

Sunstein, C. R. (2002). Probability neglect: Emotions, worst cases, and law. *The Yale Law Journal*. https://doi.org/10.2139/ssrn.292149.

Sunstein, C. R. (2003). Terrorism and probability neglect. *The Journal of Risk and Uncertainty*, *26*(2/3), 23–38. https://doi.org/10.1007/978-1-4757-6787-2_2.

Svenson, P., Forsgren, R., Kylesten, B., Berggren, P., Wong Rong Fah, W. R., Choo, M. S., & Hann, J. K. Y. (2010). Swedish-Singapore studies of Bayesian modelling techniques for tactical intelligence analysis. In *2010 13th international conference on information fusion*. IEEE. https://doi.org/10.1109/icif.2010.5711904.

Szeligowski, R. (2018). *Cognifying the OODA loop: Improved maritime decision making*. (Tech. Rep.). Gravely Naval Research Group, Naval War College Newport United States.

Theil, M. (2002). The role of translations of verbal into numerical probability expressions in risk management: A meta-analysis. *Journal of Risk Research*, *5*(2), 177–186. https://doi.org/10.1080/13669870110038179.

Turner, B. M., & Schley, D. R. (2016). The anchor integration model: A descriptive model of anchoring effects. *Cognitive Psychology*, *90*(2), 1–47. https://doi.org/10.1016/j.cogpsych.2016.07.003.

Tversky, B., Corter, J., Yu, L., Mason, D., & Nickerson, J. (2011). Visualizing thought: Mapping category and continuum. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *3*(33), 499–535. https://doi.org/10.1111/j.1756-8765.2010.01113.x.

Tversky, A., & Kahneman, D. (1973). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press. https://doi.org/10.21236/ad0767426.

van Westrenen, F., & Praetorius, G. (2014). Situation awareness and maritime traffic: Having awareness or being in control? *Theoretical Issues in Ergonomics Science*, *15*(2), 161–180. https://doi.org/10.1080/1463922x.2012.698661.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences*, *22*(2), 219–240. https://doi.org/10.1111/j.1540-5915.1991.tb00344.x.

Vessey, I. (1994). The effect of information presentation on decision making: A cost-benefit analysis. *Information Management*, *27*(2), 103–119. https://doi.org/10.1016/0378-7206(94)90010-8.

Vessey, I. (2006). The theory of cognitive fit: One aspect of a general theory of problem solving? In B. Shneiderman, P. Zhang, & D. Galletta (Eds.), *Human computer interaction and management information systems: Foundations (advances in management series)*. Routledge.

Vessey, I., & Galletta, D. (1991). Cognitive fit: An empirical study of information acquisition. *Information Systems Research*, *2*(1), 63–84. https://doi.org/10.1287/isre.2.1.63.

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science*, *20*(1), 75–100. https://doi.org/10.1207/s15516709cog2001_3.

Vong, W. K., Hendrickson, A. T., Navarro, D. J., & Perfors, A. (2019). Do additional features help or hurt category learning? The curse of dimensionality in human learners. *Cognitive Science*, *43*(3), Article e12724. https://doi.org/10.1111/cogs.12724.

Ward, T. (1983). Response temp and separable-integral responding: Evidence for an integral-to-separable processing sequencing in visual perception. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(1), 103–112. https://doi.org/10.1037/0096-1523.9.1.103.

Willems, S., Albers, C., & Smeets, I. (2020). Variability in the interpretation of probability phrases used in Dutch news articles — a risk for miscommunicati, *Journal of Science and Communication*, *19*(02), A03. https://doi.org/10.22323/2.19020203.

Wills, A. J., Edmunds, C. E. R., Le Pelley, M. E., Milton, F., Newell, B. R., Dwyer, D. M., & Shanks, D. R. (2019). Dissociable learning processes, associative theory, and testimonial reviews: A comment on Smith and Church (2018). *Psychonomic Bulletin & Review*, *26*, 1988–1993. https://doi.org/10.3758/s13423-019-01644-3.

Wills, A. J., Ellett, L., Milton, F., Croft, G., & Beesley, T. (2020). A dimensional summation account of polymorphous category learning. *Learning & Behavior*, *48*(1), 66–83. https://doi.org/10.3758/s13420-020-00409-6.

Wills, A. J., Inkster, A. B., & Milton, F. (2015). Combination or differentiation? Two theories of processing order in classification. *Cognitive Psychology*, *80*, 1–33. https://doi.org/10.1016/j.cogpsych.2015.04.002.

Wogalter, M. S., Conzola, V. C., & Smith-Jackson, T. L. (2002). Research-based guidelines for warning design and evaluation. *Applied Ergonomics*, *33*(3), 219–230. https://doi.org/10.1016/s0003-6870(02)00009-1.

Yamagishi, K. (2003). Facilitating normative judgments of conditional probability: Frequency or nested sets? *Experimental Psychology*, *50*(2), 97–106. https://doi.org/10.1026//1618-3169.50.2.97.

Yang, Z. L., Wang, J., Bonsall, S., & Fang, Q. G. (2009). Use of fuzzy evidential reasoning in maritime security assessment. *Risk Analysis: An International Journal*, *29*(1), 95–120. https://doi.org/10.1111/j.1539-6924.2008.01158.x.

Yu, H. C. (2014). A cross-cultural analysis of symbolic meanings of color. *Chang Gung Journal of Humanities and Social Sciences*, *7*(1), 49–74.

Zahner, D., & Corter, J. E. (2010). The process of probability problem solving: Use of external visual representations.

Mathematical Thinking and Learning, 12(2), 177–204. https://doi.org/10.1080/10986061003654240.

Zakrzewski, A. C., Church, B. A., & Smith, J. D. (2018). The transfer of category knowledge by macaques (Macaca mulatta) and humans (Homo sapiens). *Journal of Comparative Psychology*, 132(1), 58–74. https://doi.org/10.1037/com0000095.

Zhang, S., Buurman, J., & Babovic, V. (2008). Design of a maritime security system under uncertainty using an evolutionary real options approach. Proceedings of 12th annual real options conference, Rio de Janeiro, Brazil, 9–10 July, 2008

Zikmund-Fisher, B. J., Witteman, H. O., Dickson, M., Fuhrel-Forbis, A., Kahn, V. C., Exe, N. L., Valerio, M., Holtzman, L. G., Scherer, L. D., & Fagerlin, A. (2014). Blocks, ovals, or people? Icon type affects risk perceptions and recall of pictographs. *Medical Decision Making*, 34(4), 443–453. https://doi.org/10.1177/0272989x13511706.