Fire Safety Journal (in press)

Situation Awareness and Habitual or Resting Bias in High-Pressure Fire-Incident Training Command Decisions

Geoff Sallis[1,2], Di Catherwood[1], Graham K. Edgar[1], Steven Baker[1] & David Brookes[1]

[1]Centre for Research in Applied Cognition, Knowledge, Learning and Emotion (CRACKLE), Natural & Social Sciences, University of Gloucestershire

[2]Gloucestershire Fire and Rescue Service

Triservice Centre

Quedgeley

Gloucestershire

Corresponding author: Graham K. Edgar (gedgar@glos.ac.uk)

Highlights

- Despite expertise, FRS personnel may display either a *conservative* bias in decision-making (accepting limited information) or a *liberal* bias (accepting a broader span of information), the former linked to *miss* errors and the latter to *false alarms*.

- Characteristic personal bias patterns ("resting bias") may be likely to emerge under high pressure since stress promotes habitual responses.

- Characteristic or "resting" bias patterns may thus emerge under pressure and along with any disjunction between actual and perceived SA may explain why well- trained FRS professionals can make decision errors under demanding conditions.

Abstract

The investigation aimed to determine if Fire and Rescue Service (FRS) personnel display characteristic individual patterns of "bias" in decision-making during high-pressure simulated fire incidents. Research using the Quantitative Analysis of Situation Awareness (QASA) method revealed that despite expertise, FRS personnel display "bias" in how information is accepted for decision-making, showing either a *conservative* bias (accepting limited information) or a *liberal* bias (accepting a broader span of information), the former associated with *miss* errors and the latter with *false alarms*.

QASA measures of Actual and Perceived Situation Awareness and Bias were obtained for 19 operational FRS Incident Commanders during two peer-assessments (one year apart) requiring management of complex simulated fireground incidents. Poor peer ratings meant potential loss of salary and status, generating high pressure on the participants. There was a high level of Actual and Perceived SA, but no significant correlation between these for either exercise ($p > .05$). Individuals displayed either conservative or liberal bias, with bias tending to be consistent across the exercises: $r_\tau = 0.335$, $p = .046$.

The finding of characteristic or "resting" bias patterns under pressure and the disjunction between actual and perceived SA may help to explain why highly-trained FRS staff can make decision errors – although there are likely to be other factors in play also.

Keywords: Human error; Naturalistic decision making; Situation awareness; Signal detection theory

Abbreviations: Quantitative Analysis of Situation Awareness (QASA); Fire and Rescue Service (FRS)

Situation Awareness and Habitual or Resting Bias in High-Pressure Fire-Incident Training Command Decisions

1. Introduction

Effective and safe fireground operations do not rest solely on expertise or knowledge and even well-trained Fire and Rescue Service (FRS) professionals can make decision errors, possibly with fatal consequences [1]. The explanation may be that under pressured conditions, individuals can lose Situation Awareness (SA)[2, 3] or due to "information bias" may overlook key information or accept faulty or irrelevant information [4-14].

Information bias is the way that individuals accept the information available about a situation for decision-making. The concept of bias is formalised in "Signal Detection" models of SA such as QASA [8, 15, 16]. As represented in Figure 1, bias can shift towards a strict or conservative criterion (with risk of rejecting or overlooking true information and making *miss* errors) or towards a lax or liberal criterion (with risk of accepting more false or irrelevant information and making *false alarm* errors). Theoretically, bias is independent of SA and even with good SA, bias can affect decision-making by filtering the information potentially available from the situational environment or the individual's cognitive representation of the situation.

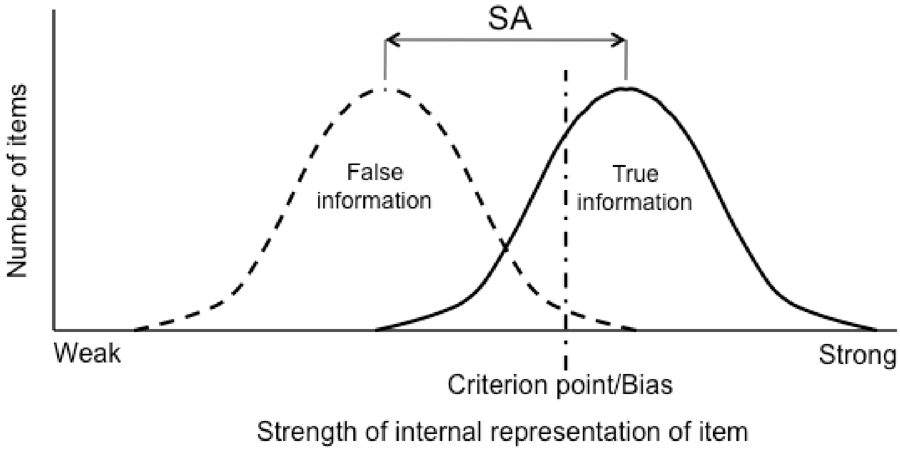

Figure 1. **A Model of the relationship between SA and Bias**: Bias can shift towards a lax or liberal criterion (accepting a broader range of information with the risk of accepting more false information and making false alarm errors) or towards a more conservative criterion (accepting a narrower band of information with a risk of rejecting true information and making miss errors). For a discussion of the distinction between criterion and bias, see Edgar *et al.*, (2018) [8].

The focus for perceptual, cognitive and attentional processing may necessarily vary between narrow and broad during an incident, but inappropriate bias regarding the *acceptance* of the information available from such processing for decision-making may precipitate errors. For these reasons, the concept of bias can explain why even highly trained fire personnel can still make errors of judgment [5, 13].

QASA measure bias by providing true and false statements (probes) about a situation and then calculating the rates of correct responses (*hits*: correctly identifying useful and true information and *correct rejections*: correctly discarding untrue or irrelevant information) and incorrect responses *(misses*: overlooking true and useful information and *false alarms*:

employing untrue or irrelevant information). From these data, individuals can be assigned a Bias score reflecting either conservative or liberal bias tendencies [few people show no bias: 5, 9, 14]. For example, in a study of 50 firefighters [5] only 6 out of 50 showed zero bias.

Although prior studies have established that bias is a key aspect in FRS decision-making exercises, it is not yet clear whether (a.) an individual's bias varies over incidents or (b.) reflects a habitual disposition that is constant across incidents or (c.) whether situational and individual factors interact, so that individual patterns invariably arise in certain situations but are less predictable in others. There are grounds for all three conclusions as follows.

In respect of possible *situational* effects on bias, degree of pressure or demand may be critical. Conditions that are complex, uncertain, time-pressured, risky, changeable with low control and high cost of errors may elicit stress, the psychological and physiological response to real or perceived threat, involving neurochemical reactions that can impair cognitive processing [17-21]. As demonstrated in demanding contexts such as air traffic control [22] and complex fireground incidents [23], high cognitive-perceptual workload can induce a narrow decision-making focus, with greater likelihood of oversight or *miss* errors [22, 24-27]. High-demand contexts may also create negative emotional arousal which can narrow attentional focus [28]. High-pressure circumstances may therefore produce cognitive processing associated with a narrow conservative bias and risk of *miss* errors. In contrast, more relaxed situations may promote positive arousal with less intensive filtering of information [29] and a more liberal bias with higher risk of *false alarms*.

Situational factors could therefore influence bias patterns, but bias could also be driven by factors internal to the *individual* that operate regardless of the situation. Personality traits have been linked to fireground command [30] and may affect bias. For example, affective state is linked to bias [26] and as noted, negative emotions such as anxiety can produce a narrow attentional focus [28], so an anxious individual could routinely display conservative bias in fireground decision-making. Other personality traits may also be associated with bias tendencies. For example, risk-taking or conversely risk-averse traits may respectively underlie relaxed or conservative bias tendencies [31]. Even unconscious memories of personal experiences ("somatic markers") [32] may influence bias. For example, a firefighter injured during a fire incident may experience either overt or unconscious emotional influences in a similar future situation [33], leading to a narrower bias towards aspects of the situation reminiscent of that which had previously caused harm, giving rise to the risk of missing other relevant features.

Other personal factors relating to past experience may also affect bias. Training and experience are of course essential to effective fireground performance, as evident in heuristic and "recognition-primed" decision making (RPD) [10, 11, 34-37]. Nevertheless, while experts may make more competent decisions [38], they may also have a narrower focus on selected features of a situation [13], possibly with a conservative bias and risk of *miss* errors. Years of firefighting experience are not related to bias tendencies or levels, with firefighters showing similar bias patterns to untrained individuals [5]. So training *per se* may not eradicate bias dispositions and may in fact be associated with particular bias tendencies that could bring errors.

Individual bias may therefore be a default pattern that operates regardless of context, with some individuals routinely exhibiting narrow conservative bias and others more liberal bias across all situations. There is however one additional possibility that *both* situational and individual characteristics *interact* to produce consistent personal bias patterns under some conditions but not others.

A "*pin-and-spring*" metaphor may be apt here. Individuals may have a "resting bias" that is a fixed disposition or point (the *pin)*, but the actual bias in any situation may be moderated by internal and external factors. Such moderating factors could act to move the bias away from the resting point (to either a more conservative or liberal bias), but there could always be a 'pull' back to that resting point (the *spring*) under certain circumstances.

One critical factor here may be the degree of pressure and associated stress in the situation. Stress can reduce the availability of high-order brain resources needed for evaluative and adaptive processing and instead promote reliance on more reflexive and habitual responses [39, 40]. If an individual loses understanding in a situation and experiences stress, there may be excessive demand on the prefrontal brain regions (e.g., orbito-frontal cortex) needed for the analytical and adaptive processing to regain understanding [41], while anxiety and stress hormones such as cortisol may impair such processing [20, 39, 40, 42]. Previous research has demonstrated that a narrow bias may be associated with a failure to regain ASA once lost [8]. The effect of stress may therefore be to overwhelm high-end brain resources. On the other hand, habitual modes of response, by definition, do not require extensive analytical processing by high-order centres, relying more on brain areas that support more automatic processing, such as the sub-cortical dorsolateral striatum [40, 43, 44].

High-pressure conditions may therefore predispose the individual towards default or habitual "resting bias" patterns that are less influenced by stress than more adaptive and analytic responses. Individuals may therefore show consistent habitual bias across similarly stressful contexts but under more relaxed conditions, the greater opportunity for higher-end processing may lead to more variable cognitive responses and less predictable bias patterns.

The consistency of individual bias under pressure has not been tested to date for FRS personnel, yet knowledge of such tendencies is of key importance to safe fireground decision-making. The only study to address bias consistency for FRS personnel so far used the QASA method in two tabletop exercises [14]. These showed both liberal and conservative bias patterns but there was no significant correlation between participants' QASA Bias scores across the exercises. Although such tabletop exercises are valuable in demonstrating that bias exists at all, habitual or resting bias tendencies may only emerge with higher pressure conditions.

Optimally this would of course involve assessment during actual FRS operations but as a first step in that direction, the current investigation assessed bias using the QASA tool with operational FRS Incident Commanders (Flexible Duty Managers, FDMs) in two simulated fireground incidents at different facilities approximately one year apart. Both were within state-of-the-art FRS training environments with a high degree of realism and capacity to develop the incident in a dynamic way. There was also a high degree of pressure for the following reasons, although we acknowledge that it is impossible to fully replicate the pressure of a real fireground.

The exercises were developed primarily to provide a professional assessment of performance needed to maintain FDM status and salary. The exercises were complex, dynamic, time-limited and with a high cost of failure. The QASA assessment was integrated into the exercises, but the performance primarily involved evaluation by FRS peers as to whether individuals were competent to maintain their current professional status or had to attend a training programme to improve specific aspects of performance or failed to meet minimum competency. The latter judgment carried serious consequences with loss of FDM status, with the requirement to re-qualify within three months or forfeit the FDM aspect of their salary and the vehicle provided for the role. For these reasons, the exercises were considered to be of substantially higher pressure than the tabletop studies in which bias consistency has been assessed to date.

The investigation employed the QASA method previously used in research across a range of decision-making contexts including FRS exercises [4, 5, 8, 9, 14, 15, 45, 46]. As noted, this approach is based on Signal Detection theory [16] and uses true-false decisions about statements (probes) containing information about a situation to provide an ASA (Actual Situation Awareness) or

Knowledge score (ability to discriminate true from false information) and a Bias score (the tendency to accept or reject the available information). Additionally, the participants rated their confidence in their responses, with this providing an estimate of Perceived SA (PSA). Any marked discrepancy between ASA and PSA is of concern as it could contribute to errors in decision-making (see Table 1, Appendix1 and Results for further details of QASA).

Alongside calculating ASA and PSA, the main aims here were also to determine (a.) whether the conservative or liberal Bias patterns displayed by operational FRS personnel in previous tabletop exercises were also in evidence in the more high-demand contexts here and (b.) whether there was any correlation in Bias scores across the two sessions for the participants.

2. Session 1

2.1. Method

2.1.1. Design.

As noted, this exercise involved a simulated fireground incident in a specialist training facility and was a mandatory professional requirement to which the QASA assessment was attached. Completion of the QASA assessment was however entirely voluntary and all participants were aware of this, and also that they could withdraw from this aspect of the assessment at any time and without giving any reason. This research complied with the British Psychological Society Code of Ethics and was approved by the Institutional Review Board at the University of Gloucestershire.

All participants individually completed the same exercise aimed at the role of operational FRS incident commanders in a realistic and developing simulated fireground incident. Each individual had to take over command from the first attendance commander and move towards a successful conclusion from an operational, environmental and social perspective (details below and in Appendix 2).

2.1.2. Sample.

The participants were 19 operational male UK FRS officers. There were 14 full-time and five retained (part-time) officers in the final sample although another three completed Session 1 but not Session 2. At time of testing in Session 1, the mean age was 49.1 years (SD 7.1) and mean years of experience 24.0 years (SD 6.3). All participants were Station Managers (SMs) working as Flexible Duty Managers (FDMs) in operational managerial roles for the FRS at incidents. The participants were tested individually in a session on one of seven days.

2.1.3. Materials and Procedure.

This first session was a dynamically unfolding Minerva exercise representing a school fire. The exercise was conducted in a FRS training environment using a fully staffed (containing all personnel that would be present in a 'real' incident) FRS "command vehicle" to present aspects and phases of the incident. Minerva exercises are highly realistic and immersive, resembling real operations (http://hydrafoundation.org/hydra-methodology/minerva). In this session, an operational command vehicle was used and all inputs to personnel within the command vehicle were designed to be as realistic as possible. The content of the exercise was designed by a FRS training team (including the lead investigator with FRS seniority) based on the job description / role map for a FDM.

The participant was "mobilised" within the facility by mobile phone/pager to a briefing room, shown a 20-minute video of the drive to the incident and given a brief via radio communications by the FRS control staff in the exercise. The scenario was then played out within a fixed time frame, with development and responses dependent on the IC's actions, within controlled parameters that allowed other role players (police, ambulance, media, parents, etc.) to introduce variables into the exercise, increasing the pressure on the IC and on the incident requirements.

For the QASA assessment, 24 probe statements regarding the scenario were presented, each requiring a True/False response and a confidence rating on a 4-point scale from *Guess* to *Certain*. There were 12 true and 12 false items in randomised order with respect to being true or false: e.g., *"The base pump overran the water supply T/F"* or *"There were six persons unaccounted for in the initial informative T/F"*.

Given that the probe statements define the situation of which awareness is being assessed, it is crucial that the T/F probe statements address relevant aspects of the situation. One or more of the investigators attended the exercises and confirmed there was good correspondence between the incident and the probes. That is, that the probe statements addressed relevant (to the role of the participant being assessed) aspects of the situation.

The QASA questionnaire was completed anonymously immediately after the exercise on a prepared answer sheet along with details of age and years of FRS experience. It is acknowledged that *post-hoc* completion of the questionnaire does introduce a memory component into the measurement, an issue that is of concern in the measurement of SA generally [47]. The use of T/F statements does, however, serve to reduce the memory load in the test. The participant is not required to remember the specific information presented in a probe, only whether they believe it to be true or false in the context of the scenario.

2.2. Results

The QASA scores. As noted, QASA scores of ASA, Bias and PSA were obtained and the raw scores (for Session 1 and Session 2) are presented in Appendix 3. The critical data are the proportion of correct responses (*hits* and *correct rejections*) and incorrect responses (*misses* and *false alarms*) to both the true (*signal*) and false (*noise*) probe statements. In practice, QASA uses just the *hits* and *false alarms* since the proportion of *correct rejections* or *misses* follow from these rates. The computation is represented by the formula in Appendix 1. [See 8, 15, 16 for the underlying Signal Detection Theory].

The ASA measure is based on a nominally non-parametric Signal Detection measure, A' (applicable even for unequal variances) with scores corrected for chance or guessing and the Bias measure is calculated using B''. These scores are re-scaled to respectively provide measures of ASA and Bias, each ranging from -100 to +100. The PSA is similarly scaled from -100 to +100. (See Table 1 for a summary of the measures.) Additionally, since the form of the underlying distribution for B" is logarithmic [48], a (natural) logarithmic transformation is applied to the Bias scores (only) to provide a linear relationship between Bias and the underlying distribution [see 8] and further analyses conducted on these transformed scores.

Table 1

Patterns of ASA, Bias and PSA using the QASA Measures, indicating the score implications.

SCORE	ACTUAL SITUATION AWARENESS (ASA)	BIAS	PERCEIVED SITUATION AWARENESS (PSA)
POSITIVE (MAX +100)	Good SA. Distinguishes true information from false: higher score is better.	Conservative/narrow/strict bias. Tends to reject information as false even if true: higher the score the greater this tendency	Strong confidence that SA is good: higher the score, the more confident
ZERO	No SA – possibly guessing	No bias towards accepting or rejecting information. A 'neutral' attitude.	Neither high nor low confidence
NEGATIVE (MAX -100)	Misguided. Judges false information as true and *vice versa.* More negative is worse.	Liberal/lax/broad bias. Tends to accept information as true even if false: the more negative the score the greater this tendency	Low confidence that SA is good; lower the score the less confident

The obtained ASA scores for both sessions and the PSA and Bias scores for Session 2 are normally distributed, but the PSA and Bias scores for Session 1 are not (details below). Hence both parametric and nonparametric analyses are used as appropriate. No Bonferroni corrections were applied in any analyses as this would be inappropriate for an exploratory study of this nature [49, 50].

Results for Session 1.

The general ASA level is high, consistent with the operational expertise of the participants: Mean = 69.42 (SD 10.31). The general PSA score is similarly high: Mean = 71.79 (SD 17.26). As noted, the ASA scores are normally distributed: $W = 0.948$, $p = 0.36$, but the PSA scores are not: $W = 0.893$, $p = .036$. It is notable that there is no significant correlation between ASA and PSA scores: $r_s = -0.042$, $p = 0.863$, so that confidence or perceived SA and actual SA may not be closely matched in all individuals.

Of most importance, all participants made errors and showed bias. There is a clearly bimodal (non-normal) distribution for these scores: $W = 0.858$, $p = 0.009$, with 12 individuals having positive Bias scores (Mean = 66.54, SD = 20.85) (and so reflecting a conservative bias and tending to make *miss* errors) and 7 showing negative scores (Mean = -67.20, SD = 25.75) (and so reflecting liberal bias and tending towards *false alarms*).

It is also important to note that there is no significant correlation between level of ASA and Bias: $r_s = 0.312$, $p = .0.194$, so that bias tendencies are not predictable from level of understanding or SA in this exercise.

2.3.Conclusions

This first session clearly establishes that the participants were skilled professionals who generally made competent judgments about the probe questions and had confidence in their abilities to do so, although this may have been less warranted in some cases. Nevertheless all participants made errors, tending to be either *misses* reflecting a conservative or narrow bias or *false alarms* reflecting a liberal or lax bias. The former individuals were at risk of overlooking key information, while the latter were at risk of using false or low-priority information. It is also important to note that Bias scores are not significantly correlated with ASA, so that bias is not simply a matter of poor understanding about the situation.

These results are consistent with prior research that confirms the operation of bias tendencies that may explain errors in the decision-making of skilled FRS professionals [4, 5, 7, 9, 14]. The current results also extend the previous evidence by indicating that such tendencies are apparent in exercises with a higher degree of realism and pressure than may have been the case in prior tabletop studies.

3. Session 2

This second session involved a repeat professional assessment with the same individuals approximately 12 months later, using a different scenario in a different FRS training environment. Using a different scenario, apart from controlling for practice effects, also allowed the assessment of common factors (such as tendencies to a certain type of bias) that were relatively independent of situation. This second session, while different to the first, also constituted another high-pressure context since there were the same negative professional consequences of poor perfomance as for Session 1. As well as the professional evaluation, the QASA method was again used to ascertain ASA, PSA and Bias.

Of key interest, this repeat session provided an opportunity to also determine whether there is consistency in Bias patterns across the two exercises after the 12 month interval. As noted, previous tabletop exercises did not reveal bias consistency [14], but the higher pressure here may provoke a greater reliance on habitual or resting bias, with more likelihood of correlation between Bias scores across the current exercises.

3.1. Method.

3.1.1. Sample.

The same 19 individuals from Session 1 participated for Session 2: mean age now being 50.1 years (SD 7.2) and mean years of experience now being 24.8 years (SD 6.1) . As before, the FRS evaluation was mandatory to maintain professional status, but the QASA assessment was optional, with all participants giving consent.

3.1.2. Materials and Procedure.

This session again involved a dynamically unfolding Minerva exercise conducted in another specialist UK FRS facility running. This facility provided a remote monitoring unit, allowing the researchers to review the probes and to build new ones in initial trial exercises. In this case the incident involved a fire at a care facility. Each participant was initially briefed about the exercise by the FRS facilitators and further detail again provided to participants following their mobilisation during a 20 minute drive-to video, with subsequent radio calls and messages (see Appendix 2 for details). As for Session 1, the scenario was then played out within a fixed time frame, with development and responses dependent on the IC's actions, with controlled parameters and other role players (police, ambulance, media, care centre manager, etc.) who provided additional information, queries, and so more pressure on the IC and the incident requirements.

The QASA assessment involved 28 probe statements judged as centrally relevant to the scenario and each required a True/False response and a confidence rating on a 4-point scale from *Guess* to *Certain*. There were 14 True and 14 False items in randomised order with respect to being true or false: e.g., *"Two Breathing apparatus teams were in the building when you arrived T/F"* or *"There was a report that some doors may have been wedged open T/F"*.

As in Session 1, an investigator with senior FRS experience attended the exercises and confirmed good correspondence (i.e. that the probes addressed aspects of the situation relevant to the role of the participant being assessed) between the T/F QASA probes and the specific events in the incident.

As for Session 1, the QASA questionnaire was completed immediately after the exercise, along with details of age and years of FRS experience.

3.2. Results

As for Session 1, there is a high level of ASA (Mean = 59.1; SD 21.94) reflecting the expertise of the participants. The ASA scores are normally distributed: $W = 0.951$, $p = 0.410$. ASA is lower than for Session 1 but not significantly so: $t(18) = 1.78$, $p = .092$, although there is also no significant correlation between ASA scores for the two sessions: $r = -.110$, $p = 0.654$, so ASA as measured here can clearly vary across exercises or incidents for individuals.

Again there is also a high level of PSA or confidence (Mean = 76.63; SD 18.28). The PSA scores are normally distributed: $W = 0.928$, $p = 0.182$. There were only 18 scores as one individual did not complete the confidence ratings in this session. The PSA does not differ significantly from that for the first session: $z = -1.350$, N - *Ties* = 18, $p = 0.177$ and there is a significant positive correlation in PSA scores for the two sessions: $r_s = 0.521$, $p = .027$, confirming consistency of confidence across the sessions.

In this session the PSA scores are significantly higher than the ASA scores: $t\ (17) = 2.168$, $p = 0.045$ and as for Session 1, there is no significant correlation between ASA and PSA scores: $r = -0.335$, $p = 0.174$. In fact, such correlation as exists is negative. These results confirm that there is no necessary correspondence between actual and perceived SA for individuals in this exercise.

As for Session 1, log-transformed scores were used to calculate Bias. As for Session 1, there is a bimodal tendency in the distribution of Bias scores, although the divergence from normality is not as extreme here: $W = 0.917$, $p = .101$. There are 12 positive Bias scores (Mean = 58.03, SD = 26.85) (conservative, narrow bias with a tendency to make *miss* errors), six negative scores (Mean = -48.38, SD = 14.27) (liberal or lax bias with greater risk of *false alarm* errors) and one person showing zero Bias.

As for Session 1, Bias is not significantly correlated with ASA: $r_s = 0.370$, $p = .119$, again confirming that bias tendency cannot be predicted from level of knowledge.

Of most interest here however is the consistency in Bias scores over the two exercises. The Bias scores for the two sessions are not significantly different: $z = -0.327$, N – Ties = 18, $p = 0.744$ and most importantly, are significantly correlated: $r_\tau = 0.335$, $p = .046$ (*Kendall's tau* was used here as there were tied scores). In respect to individual patterns, eight people showed consistently positive scores, three consistently negative scores and the remaining seven showed only non-significant changes: z = -0.676, N – Ties = 7, p = 0.499. In sum, there is a clear trend towards bias consistency across the two exercises.

3.3. Conclusions.

The results for the second session mirror those for the first in showing a high level of ASA and PSA but again with no significant correlation between actual and perceived SA scores. This session also extends those results in two important ways. Firstly, there is no significant correlation between ASA scores for the two sessions, showing that SA as measured here can vary across exercises for the same individuals. That there was no significant difference in the ASA scores across the two sessions suggests that one session was not inherently more ‘difficult’ than the other but that the individuals, in terms of ASA, performed differently in the two sessions, as illustrated by the lack of a significant correlation between the ASA scores across sessions. These data should, however, be interpreted with caution, as the small sample size means that only large effects are likely to generate statistically significant results.

There is, however, correspondence in individual PSA scores across the exercises, indicating that confidence was consistent even if actual SA was not. This discrepancy in ASA and PSA may lead to faulty and risky decision-making, either from undue confidence or a lack of conviction.

Also of key interest in regard to errors in decision-making are the results for Bias. Again there is no significant correspondence between ASA and Bias, suggesting that bias is not a simple function of level of understanding. Comparable numbers of individuals in both sessions showed either positive Bias scores reflecting conservative bias with *miss* errors or negative Bias scores indicating liberal bias with *false alarm* errors. Of particular interest here however is the significant positive correlation in Bias scores over the two exercises. This outcome clearly establishes that there was a tendency towards consistency in bias patterns across the different scenarios.

4. Discussion.

The current findings resemble past results [4, 5, 7, 9, 14] by showing relatively high actual and perceived SA as expected for well-trained FRS personnel, but nevertheless indicating several reasons why such professionals are susceptible to decision-making errors.

Firstly, it is concerning that there was no significant correlation between actual SA and PSA (confidence) in either session (the raw scores can be seen in Appendix 3). This discrepancy could produce either unwarranted over-confidence or alternately undue low confidence, leading respectively to either hasty or hesitant decision-making, both of which could carry risk on the fireground. PSA was correlated across the sessions, possibly reflecting core personality characteristics that could influence FRS decision-making and this could be a useful focus for further research.

Secondly, the bias patterns evident in past studies [4, 5, 7, 9, 14] were also apparent here. Both conservative (strict, narrow) and liberal (lax, broad) bias tendencies were displayed, with only one person showing no bias in one exercise. As noted, conservative bias (the most common bias found) is linked to *miss* errors and liberal bias to *false alarms*. Having a narrow or alternately a broad cognitive, attentional or perceptual focus may be necessary to acquire information at different phases of an incident, but bias reflects how any such information is being *accepted for decision-making* and can explain errors of judgment even for well-trained FRS professionals.

Finally, of key importance is the trend to consistency in bias across two different exercises in two different venues in the current investigation. This result could mean that bias is driven by factors inherent to the individual. The finding of consistency here, but not in prior studies [14] however, also implicates situational factors as well. As argued above, the critical factor may be the level of situational pressure and personal stress. An individual may revert to a “resting” or habitual bias if excessive cognitive workload, anxiety and stress overwhelm high-end analytic and adaptive processing, allowing less cognitively-demanding customary responses to prevail [39, 40, 42-44]. FRS personnel operating in physically and cognitively overwhelming conditions such as in the Storm King Mountain wildfire [14, 23] may therefore be more likely to rely on habitual or resting bias dispositions. There may in fact be little available mental or situational capacity to do otherwise in such circumstances.

As noted, prior studies did not find such individual consistency [14], but the conditions may have not been as stressful as in the current study, given the highly punitive real-life outcomes of poor performance in the latter. The pressure was noted by participants during the debriefing sessions with comments such as "*You are not under that much pressure at an incident that often*" and "*It is easy to get [only] part information because there is a lot going on and you fill in those gaps which don't always ring true"*.

In sum, the bias consistency here may be due to the interface between the high demands of the situation and personal characteristics such as anxiety or capacity to handle cognitive complexity. For example, anxiety produces a narrow attentional focus [28] and negative affect a positive (conservative) bias [26]. An anxious individual may therefore tend towards habitual conservative bias under pressure [31], but not necessarily do so under more relaxed circumstances when high-end processing may override that induced by anxiety. Individual bias tendencies may therefore be most apparent and predictable under high pressure conditions, when there is less time and capacity for more evaluative, creative and analytic processing that could elicit more variable bias patterns. To use the "*pin and spring*" metaphor proposed above, an individual's bias may therefore be moderated by both internal and external factors that can induce habitual or resting bias patterns or alternately can move towards less predictable tendencies. The *"pin-and-spring"* metaphor is, we believe, apt as it does not suggest that individuals will always display the same level of bias, or even in the same direction (the data in Appendix 3 do show that some individuals changed the 'direction' or their bias across sections), only that there is a 'pull' to a particular bias and the spring can be extended in either direction (towards a positive or negative bias).

The “pin-and-spring’ metaphor also encapsulates the notion that there may be multiple influences acting on individual bias level. Bias is, we believe, a state internal to the individual (a tendency to accept or reject information) but external factors, such as pressure can moderate that bias. Similarly, although bias may be an important element in building ASA and performance based on that ASA, it is unlikely to be the only factor.

Even if the pattern of personal bias does vary across situations, the knowledge that bias operates *at all* even with good SA, may be critical to improving safety in FRS decision-making. To this end, the QASA method is incorporated in online training tools *FireMind* [4; https://uniofglos.blog/firemind/] and *FireFront* [51; https://firefront.eu/] (European Union *Erasmus+* projects) to provide accessible platforms for training self-awareness of bias tendencies with the ultimate goal of diminishing risk in FRS decision-making and operations.

5. Acknowledgements:

The research in this report formed part of the requirements for the doctoral degree of the first author at the University of Gloucestershire. The authors are grateful to the Gloucestershire Fire and Rescue Service for their assistance in this project.

6. Funding:

The background research and theory development associated with this report has been funded by the *Erasmus+* scheme of the European Union in the *FireMind* [grant number 2014-1-UK01-KA202-001637]and *FireFront* [grant number 2018-1-UK01-KA202-048113] projects.

7. References

1. Fire Brigade Union, *In the line of duty.* Report by the Fire Brigade Union, UK, 2008.
2. Endsley, M.R., *Toward a theory of situation awareness in dynamic systems.* Human Factors, 1995. **37**(1): p. 32-64.
3. Endsley, M.R., *Theoretical Underpinnings of Situation Awareness: A Critical Review*, in *Situation Awareness Analysis and Measurement*, M.R. Endsley and D.J. Garland, Editors. 2000, Lawrence Erlbaum Associates: Mahwah, NJ. p. 3-32.
4. Arendtsen, B., et al., *Firemind: Trialling a new tool for training fire and rescue service decision-making*, in *International Fire Professional*. 2016. p. 14-17.
5. Catherwood, D., et al., *Fire Alarm or False Alarm?! Decision-making "Bias" of Firefighters in Training Exercises.* International Journal Emergency Services, 2012. **1**(2): p. 135-158.
6. Catherwood, D., et al., *Scoping the Fireground: the range and bias of information used in decision-making in simulated fireground exercises*, in *27th International Congress Applied Psychology*. 2010: Melbourne, Australia.
7. Catherwood, D., et al., *Understanding fireground 'situation awareness'.* Fire, 2011. **103**(1333): p. 27-29.
8. Edgar, G.K., et al., *Quantitative Analysis of Situation Awareness (QASA): Modelling and Measuring Situation Awareness using Signal Detection Theory.* Ergonomics, 2018. **61**(6): p. 762-777.
9. Edgar, G.K., et al., *"I always know what's going on." Assessing the Relationship between Perceived and Actual Situation Awareness across Different Scenarios.* World Academy of Science, Engineering and Technology, 2012. **71**: p. 1480-1481.
10. Gasaway, R.B., *Making intuitive decisions under stress: understanding fireground incident command decision-making.* International Fire Service Journal of Leadership and Management, 2007. **1**(1): p. 8-18.
11. Klein, G., R. Calderwood, and A. Clinton-Cirocco, *Rapid Decision Making on the Fire Ground: The Original Study Plus a Postscript.* Journal of Cognitive Engineering and Decision Making, 2010. **4**(3): p. 186–209.
12. Omodei, M.M., et al., *"More is better?": A bias toward overuse of resources in naturalistic decision-making settings*, in *How professionals make decisions*, H. Montgomery, R. Lipshitz, and B. Brehmer, Editors. 2005, Lawrence Erlbaum: Mahwah, New Jersey. p. 29-41.
13. Perry, N., et al. *Decision support for competent and expert firefighters*. in *Proceedings of the 8th Industrial and Organisational Psychology Conference (IOP 2009)*. 2009. Australian Psychological Society.
14. Sallis, G., et al., *The human brain in fireground decision-making: trustworthy firefighting equipment?* International Fire Professional, 2013(5): p. 21-24.
15. Edgar, G.K. and H.E. Edgar, *Using Signal Detection Theory to Measure Situation Awareness: The Technique, The Tool (QUASA), the TEST, the Way Forward*, in *Decision making in complex environments*, M. Cook, J. Noyes, and Y. Masakowski, Editors. 2007, Ashgate: Aldershot, UK. p. 373-385.
16. Stanislaw, H. and N. Todorov, *Calculation of signal detection theory measures.* Behaviour Research Methods, Instruments, & Computers, 1999. **31**(1): p. 137-149.
17. Kemeny, M.E., *The psychobiology of stress.* Current Directions in Psychological Science, 2003. **12**: p. 124-129.
18. Purves, D., et al., *Principles of Cognitive Neuroscience*. 2008, Sunderland, MA: Sinauer.

19. Shields, G.S., M.A. Sazma, and A.P. Yonelinas, *The effects of acute stress on core executive functions: a meta-analysis and comparison with cortisol.* Neuroscience & Biobehavioral Reviews, 2016. **68**: p. 651-668.
20. Staal, M.A., *Stress, cognition and human performance: A literature review and conceptual framework*. 2004, NASA Ames Research Centre: Moffett Field, CA.
21. Wirth, M.M., *Hormones, stress, and cognition: the effects of glucocorticoids and oxytocin on memory.* Adaptive Human Behaviour and Physiology, 2015. **1**(177-201).
22. Endsley, M.R. and M.D. Rodgers, *Distribution of attention, situation awarenes, and workload in a passive air traffic control task: implications for operational errors and automation.* Air Traffic Control Quarterly, 1998. **6**(1): p. 21-44.
23. Useem, M., J. Cook, and L. Sutton, *Developing leaders for decision making under stress: wildland firefighers in the south canyon fire and its aftermath.* Academy of Management Education and Learning, 2005. **4**: p. 461-485.
24. Forster, S. and N. Lavie, *Attentional capture by entirely irrelevant distractors.* Visual Cognition, 2008(16): p. 2 & 3.
25. Franco-Watkins, A.M., T.C. Pashler, and H. Rickard, *Taxing executive processes does not necessarily increase impulsive decision making.* Experimental Psychology, 2010. **57**: p. 193–201.
26. Nikolla, D., et al., *Can bottom-up processes of attention be a source of "interference" in situations where top-down control of attention is crucial? .* British Journal of Psychology, 2018. **109**(1): p. 85-98.
27. Rees, G., C.D. Frith, and N. Lavie, *Modulating irrelevant motion perception by varying attentional load in an unrelated task.* Science, 1997. **278**: p. 1616–1619.
28. Derryberry, D. and D.M. Tucker, *Motivating the focus of attention*, in *The heart's eye: Emotional influences in perception and attention*, P.M. Neidenthal and S. Kitayama, Editors. 1994, Academic Press: San Diego, CA. p. 167-196.
29. Isen, A.M. and A.A. Labroo, *Some ways in which positive affect facilitates decision making and judgment*, in *Emerging perspectives in judgment and decision making*, S.L. Schneider and J. Shanteau, Editors. 2003, Cambridge University Press: Cambridge. p. 365-393.
30. Burke, E., *Psychological research and development in the London Fire Brigade*, in *Decision making under stress: emerging themes and implications*, R. Flin, et al., Editors. 1997, Ashgate: Aldershot, UK. p. 116-125.
31. Li, C.R., H.H.-A. Chao, and T.-W. Li, *Neural correlates of speeded as compared with delayed responses in a Stop Signal Task: An indirect analog of risk taking and association with an anxiety trait.* Cerebral Cortex, 2009. **19**: p. 839-848.
32. Damasio, A., *The feeling of what happens: body, emotion and the making of consciousness*. 2000, London: Vintage.
33. Baumann, M.R., C.L. Gohm, and B. Bonner, *Phased training for high-reliability occupations: live-fire exercises for civilian firefighters.* Human Factors, 2011. **53**: p. 548-557.
34. Keller, N., et al., *Naturalistic Heuristics for Decision Making.* Journal of Cognitive Engineering and Decision Making, 2010. **4**: p. 256-274.
35. Klein, G., *Making decisions in natural environments*. 1997, Klein Associates: Yellow Springs, OH.
36. Klein, G., *Naturalistic Decision Making.* Human Factors, 2008. **50**(3): p. 456–460.
37. Klein, G., *The recognition-primed decision (RPD) model: Looking back, looking forward*, in *Naturalistic Decision Making*, C.D. Zsambok and G. Klein, Editors. 1997, Lawrence Erlbaum Associates: Mahwah, NJ. p. 285-292.
38. Crundall, D. and V. Kroll, *Prediction and perception of hazards in professional drivers: Does hazard perception skill differ between safe and less-safe fire-appliance drivers?* Accident Analysis & Prevention, 2018. **121**: p. 335-346.

39. Arnsten, A.F., *Stress signalling pathways that impair prefrontal cortex structure and function.* Nature Reviews Neuroscience, 2009. **10**: p. 410-422.
40. Schwabe, L. and O.T. Wolf, *Stress-induced modulation of instrumental behaviour: from goal-directed to habitual control of action.* Behavioral Brain Research, 2011. **219**(321-328).
41. Catherwood, D., et al., *Mapping brain activity during loss of situation awareness: an EEG investigation of a basis for top-down influence on perception.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 2014. **56**(8): p. 1428-1452.
42. McEwen, B.S. and J.H. Morrison, *The brain on stress: Vulnerability and plasticity of the prefrontal cortex over the life course.* Neuron, 2013. **79**: p. 16-29.
43. Balleine, B.W. and J.P. O'Doherty, *Human and rodent homologies inaction control: Corticostriatal determinants of goal-directed and habitual action.* Neuropsychopharmacology, 2010. **35**.
44. Valentin, V.V., A. Dickinson, and J.P. O'Doherty, *Determining the neural substrates of goal-directed learning in the human brain.* Journal of Neuroscience, 2007. **27**: p. 4019-4026.
45. Edgar, G., et al., *Situation awareness, information use, and EEG. It's not what you have, it's what you do with it.* British Association Cognitive Neuroscience Annual Conference, University of the West of England, Bristol, 2010.
46. Rousseau, R., et al., *The role of metacognition in the relationship between objective and subjective measures of situation awareness.* Theoretical Issues in Ergonomics Science, 2010. **11**(1/2): p. 119-130.
47. Wickens, C.D., *Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 2008. **50**(3): p. 397-403.
48. Macmillan, N.A. and C.D. Creelman, *Response bias: Characteristics of detection theory, threshold theory, and" nonparametric" indexes.* Psychological Bulletin, 1990. **107**(3): p. 401.
49. Perneger, T.V., *What's wrong with Bonferroni adjustments.* BMJ: British Medical Journal, 1998. **316**(7139): p. 1236.
50. Bender, R. and S. Lange, *What's wrong with Bonferroni adjustments.* British Medical Journal, 1998. **316**: p. 1236-1238.
51. Thoelen, F., et al., *FireFront: A new tool to support training in Fireground Situation Awareness, Situation Understanding and Bias.* International Fire Professional, 2020. **34, Names alphabetically by country and then surname**.

Appendix 1

Full details about the QASA scores are provided in the Results above (and see Edgar & Edgar, 2007; Edgar et al., 2018), but the essential computations can be represented as (after Stanislaw and Todorov, 1999):

$$A' = 0.5 + \left(\text{sign}(H-F) \frac{(H-F)^2 + |H-F|}{4\max(H,F) - 4HF} \right)$$

$$B'' = \text{sign}(H-F) \frac{H(1-H) - F(1-F)}{H(1-H) + F(1-F)}$$

(where *H*= hit rate and *F* = false alarm rate and *max (H,F*) = either *H* or *F*, whichever is the greater).

Appendix 2

Information in the Session Briefs

Information for the Brief in Session1

The brief stated the following information:

The FRS has been called to a fire at a nearby school. The time is 1000 Monday, with a light rain, a brisk northerly wind and temperatures around 4˚C. Two appliances from local FRS stations and a Station Commander (SC) have been mobilised. On receipt of a Make Pumps 5 message from the current Incident Commander (IC) at the school, Operations Control has mobilised more resources to the Incident in line with the FRS procedures, including two further Station Commanders (one to manage the Incident Command Unit (ICU) and the other participant to support the Incident Command System).

The brief also included an incoming informative from the initial incident commander who was at Watch Commander (WC) level on the first appliance to Operations Control which provided the following information:

The 10M x 80M school building is well alight. There are five people confirmed unaccounted for and search & rescue and firefighting operations are underway, with one pump at the front, one at the back, and one securing a water supply at the entrance. The ICU is setting up and a SC is there to run command support but has not received a brief as yet. There is a great deal of smoke with a thick black plume from the roof and the smell indicated that plastics are involved. The teachers and pupils are self-evacuating but difficult to manage. The fire development has increased over the last 10 minutes and it is the intention to flood the building with BA teams to undertake search and rescue for the identified missing people as well as fight the fire. There are four pumps in attendance and all personnel are presently employed.

Information for the Brief in Session 2

The mobilising message for the brief was as follows:

It is 16.00 and there is a serious fire in a palliative care centre for neurological patients, all with mobility issues and other complications. The participant is the nearest available Flexi-Duty Officer and has to take over operational command of the incident. Five pumps are already in attendance. The care centre has treatment rooms, physiotherapy facilities, a pharmacy and 13 bedrooms, 5 with en-suite facilities and with oxygen cylinders in some rooms (unknown location). The fire began in the kitchen area but there is heavy smoke-logging and it has rapidly spread to other rooms in the sub-ground floor level and the lift shaft, moving up to the ground floor where it is currently burning in the reception area. If the fire is not checked it will easily spread to the first floor. Four people are unaccounted for but twenty people have been evacuated and are collected in front of the premises. The incident has already been sectorised, with four BA committed through the side entrance with a 45mm jet. A team of two BA is about to be committed through the front door with a hose-reel to prevent the fire coming up the internal stairs. A further team of two BA are rigged but not committed to search the residential wing. The mobilised SM was to then take over the incident and respond to control formally "*that they were taking command of the incident*".

Appendix 3

Raw scores across the two sessions

	ASA		PSA		Bias	
Participant	Session 1	Session 2	Session 1	Session 2	Session 1	Session 2
1	68.65	80.31	64.44	61.91	79.9465401	83.7585956
2	80	92.86	55.56	33.33	100	100
3	62.9	71.39	60.92	80.25	50.9700056	-28.054027
4	58.33	50.73	88.89	*	65.9684276	65.1590059
5	62.89	78.27	60	47.62	48.5975411	86.2354198
6	61.03	77.25	84.44	100	54.1653301	-64.206537
7	69.81	44.44	88.89	52.38	40.1542731	0
8	74.18	41.16	91.11	80.25	-72.736933	-51.599872
9	70	83.33	42.22	76.19	-100	21.5216866
10	41.67	66.82	37.78	66.67	-59.092693	-51.134313
11	59.39	52.78	77.78	97.62	-77.067882	41.7891287
12	75	35.42	77.78	71.43	-21.840338	31.249498
13	66.7	68.82	88.89	95.24	-51.957371	51.1343128
14	85.77	83.33	75.56	83.33	77.258681	21.5216866
15	74.92	25	70.12	83.33	74.5536014	41.1174498
16	83.33	56.55	44.44	78.57	100	77.1911701
17	74.88	65.96	81.61	92.59	38.8236767	75.6368839
18	77.14	29.09	88.89	88.1	-87.699284	-60.598014
19	72.35	19.43	82.72	90.48	68.0429777	-34.69911

* There is one missing data point where a participant did not complete the confidence ratings.