Framing and Context Effects in Visual Search Training

Frank C. Lacson
Pacific Science and Engineering Group
San Diego, California

Cleotilde Gonzalez
Carnegie Mellon University
Pittsburgh, Pennsylvania

Poornima Madhavan
Old Dominion University
Norfolk, Virginia

Framed incentive structures and context effects may have training implications for applied visual search tasks such as airline luggage screening. Participants were trained with various incentive structures that focused, or were sensitive to, various signal detection outcomes. Also, participants were trained with different context representations (weapons or produce search). Twenty-four hours later, participants performed a transfer session in which the incentive structure and target set was unknown. Incentive structures that focused on negative outcomes (misses) led to a response bias that was closer to optimal compared to structures that focused on positive outcomes (hits). Task context affected response bias but had mixed effects on sensitivity. Results of this study may better inform the design of training and automated support for airline luggage screening and similar applied visual search tasks.

INTRODUCTION

Visual search involves scanning for a signal item of unknown location and presence amongst a number of noise items. Visual search research is common in variety of domains such as industrial inspection (Parasuraman, 1986; Swets, 1992; Gramopadhye, Melloy, Gopinath, & Budgavi, 1997), medical diagnosis (Pisano, Gatsonis, Hendrick, Yaffe, Baum, Acharyya, et al., 2005), and airline luggage screening (McCarley, Kramer, Wickens, Vidoni, & Boot, 2004).

Research in the luggage screening task has investigated several cognitive and memory factors that influence visual search and detection. For example, Madhavan and Gonzalez (2006) investigated the effects of mapping of stimuli and workload on the effective detection of targets. In agreement with other mapping studies they found that when targets were consistently mapped (kept as targets only and never as distractors) detection was most accurate even under high workload. Another example is the study by Madhavan, Gonzalez and Lacson (2007), where multiple levels of signal probability (*base rate*) were studied on learning in an airline luggage screening task. Their results revealed that participants receiving higher base rates during training obtained higher hit rates at transfer, when base rates were low, compared to participants encountering lower base rates during training.

The current study is an extension of the research summarized above, aiming at constructing a memory model of visual detection. The current study focuses on investigating the framing of incentive structure and the emotional characteristics of the context on detection of visual targets.

Framing and Context Effects

The *framing effect* is observed when the description of options in terms of gains (positive frame) rather than losses (negative frame) elicits systematically different choices (Tversky & Kahneman, 1984; Levin, Schneider, & Gaeth, 1998 provides an extensive review). Specifically, contradictions in a decision-maker's choices were found when faced with consequentially identical decision problems framed positively (in terms of gains) versus negatively (in terms of losses). The framing effect consistently causes risk seeking when one is presented with a negative frame and a risk aversion when one is presented with a positive frame. Applications of framing effect research has been found in aviation multi-tasking (Nygren, 1997), military threat judgment (Perrin, Bennett, Walrath, & Grossman, 2001), and automation use in signal detection (Dzindolet, Pierce, Beck, & Dawe, 2002; Lacson, Wiegmann, & Madhavan, 2005).

In this study, incentive structures in a visual search task were framed to provide a relatively large positive monetary gain for hit outcomes (*hit-sensitive*) or to deliver a relatively large monetary loss for miss outcomes (*miss-sensitive).* Participants trained with a miss-sensitive structure were expected to exhibit different detection behaviors than those presented with a hit-sensitive structure, although optimal detection behaviors of the two structures are identical.

Framing and Context. A recent fMRI study on the framing effect suggested that the differential responses in structurally identical problems may be due to the interaction of emotional and cognitive responses (Gonzalez, Dana, Koshino & Just, 2005). As a result, contexts involving moral and emotional consequences as well as high stakes are expected to strengthen framing effects compared to contexts without emotionally-charged consequences. In this study, framed incentive structures were presented within two task contexts: weapons search (airline luggage screening) and produce search (agriculture inspection).

Purpose of Study

This study explores the effects of framed incentive structure and regulatory contexts on inspection performance and decision strategies. Incentive structure framing is expected to affect detection behavior. Specifically, miss-sensitive (negative) framed structures are expected to lead towards a more optimal decision strategy due to an increased focus on avoiding miss outcomes. Although miss-focused incentive structures will lead to a more optimal decision strategy, reductions in overall detection performance (sensitivity) may occur due to an increase in false alarm rates. Additionally, the luggage screening context is expected to cause similar miss-focused behaviors due to the prevention focus on safety and responsibility relative to the produce inspection context. Finally, an interaction between incentive frame and context is expected; more emotionally-charged contexts such as airline luggage screening would have strengthened incentive framing effects.

METHOD

Participants

Ninety undergraduate students, graduate students, and community members at Carnegie Mellon University participated in this experiment. Ages of the participants ranged from 18 to 45 and all reported to have normal or corrected vision. Upon completion of the 90-minute experiment that was conducted over two days, participants were given $15 plus a performance bonus between $0 and $7.

Visual Search Simulation

A visual search simulation was developed to present a series of complex images within two contexts: weapons and produce. Participants in the weapons context played the role of a trainee airline security inspector and searched for one weapon (e.g. knife or gun) embedded within x-ray images of common travel items. In the produce context, participants played the role of a trainee agriculture inspector and searched for one fruit embedded within images of vegetables.

Images were constructed by putting together individual items in the image, controlling for clutter, size of the individual items, and size of the container (see Figure 1). The images for the weapons context were provided by the Transportation Safety Administration (TSA) and the images for the produce context were obtained from an internet search.

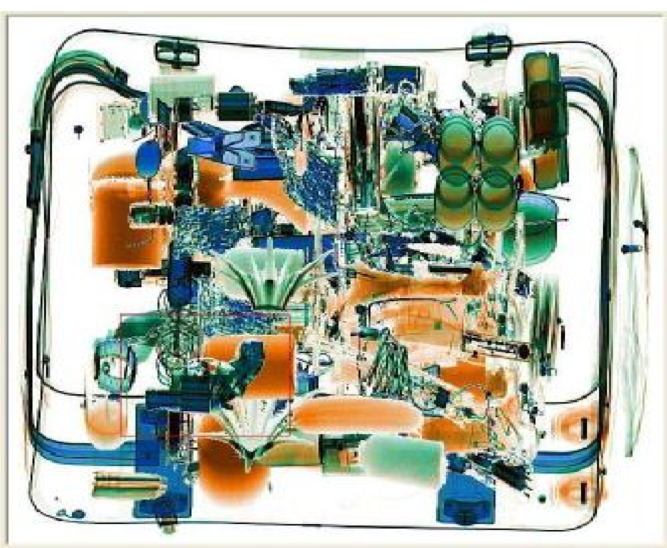


Figure 1. Sample visual search screenshot (weapons context).

Experiment Design

Independent Variables. The experiment consisted of a 3 (framing structure) x 2 (context) between-participants design. The framing structure factor contained three levels: Miss-Sensitive (M-S), Hit-Sensitive (H-S), and Equal. The context factor contained two levels: Weapons and Produce.

Table 1. Incentive structure and optimal response bias.

Framing Structure	**Outcome (Points)**				**Response Bias**
	Hit	**Miss**	**FA**	**CR**	$\ln\beta_{opt}$
Miss-Sensitive (Loss)	+50	-350	-24	+1	-1.39
Hit-Sensitive (Gain)	+350	-50	-24	+1	-1.39
Equal	+1	-1	-1	+1	+1.39

For each outcome, Table 1 (see above) describes the values and costs associated with each outcome. Note that Miss outcomes received the greatest point change (-350) in the M-S structure while Hit outcomes received the greatest point change (+350) in the H-S structure. All outcomes in the Equal group receive either +1 (H and CR) or -1 (M and FA).

The measure of optimal response bias (β_{opt}) (Wickens & Hollands, 2000) was used to ensure an identical expected value between the M-S and H-S frames. The quantities p(N) and p(S) refer to the noise and signal base rate, respectively. Values of correct rejection outcomes and hit outcomes are referred as V(CR) and V(H). Costs of false alarms and misses are referred as C(FA) and C(M). A natural logarithmic transformation was than performed on β_{opt} to account for floor and ceiling effects.

$$\beta_{opt} = [\ p(N) / p(S)\ ] * [\ V(CR) + C(FA) / V(H) + C(M)] \quad (1)$$

The H-S and M-S groups had an identical optimal response bias measure ($\ln\beta_{opt} = -1.39$), meaning that the problems were structurally identical and only differed in the framing of the problem. The Equal group optimal response bias corresponds to a strategy to maximize the number of correct decisions.

Additionally, the incentive structure in Table 1 above was created using a number of assumptions related to a signal base rate below chance (20%) and the representation of the task towards airline luggage screening. First, Hits and Misses were considered more valuable (contained a higher absolute value) than False Alarms and Correct Rejections. Second, False Alarms costs were considered more valuable than Correct Rejections. As a result, it was not feasible to create a FA-Sensitive (FA-S) and CR-Sensitive (CR-S) framing structure that 1) fit the previously mentioned assumptions and 2) contained an identical optimal response bias of the H-S and M-S structures.

Dependent Variables. Dependent variables consisted of sensitivity (d') and response bias ($\ln\beta$) measures.

Procedure

Day 1: Training. Participants were assigned into one of three framing levels: M-S, H-S, and Equal and one of two stimuli types: weapons and fruit. Participants were then given an instruction set containing a role and task description. This instruction set also described the assigned outcome structure. All participants were instructed to maximize their score, thus maximizing their performance bonus.

Participants then performed the visual search task on 360 images, split into 12, 30-trial blocks. A unique four-item target set was shown to the participant before each block. Twenty percent of trials in each block (6 out of 30) contained one item from the signal set. In each trial, an image was presented for 3 seconds; time remaining was graphically shown below the luggage image. After the presentation, the screen cleared and participants were prompted to make a decision on whether to search or to pass the bag. Confidence ratings for each decision were collected on a 5-point Likert scale.

Outcome feedback was provided after the decision and confidence rating declaration. Participants were then presented with a button to proceed to the next trial. A summary screen containing the participant's current performance bonus and the cumulative hit, miss, false alarm, and correct rejection rates appeared after every two trial blocks (60 trials). On average, the training phase lasted 60 minutes.

Day 2: Transfer. Participants returned at the same time slot the following day for a transfer phase that differed from the training phase in four aspects. The target set contained novel items that were not previously seen by the participants. Novel items in the weapons context consisted of non-gun, non-knife dangerous items. In the produce context, novel items consisted of fruit not previously shown to the participant. The four-item target sets were not presented before each trial block. The outcome structure was also hidden and instructions stated that participants should "use previous training to maximize their perceived point total." Participants performed the signal detection task on 6, 30-trial blocks. Twenty percent of the trials (6 out of 30) contained a signal item.

Outcome feedback was provided in a manner similar to the training phase; the stimulus image was shown on the screen and a red rectangular box highlighted the target= if present. The words "Hit," "Miss, "False Alarm," or "Correct Rejection" appeared below the image. However, no outcome score was given for each trial. On average, the 180-trial transfer phase lasted 30 minutes.

The performance bonus was calculated based on the participant's score on each day compared to a normalized distribution of possible scores for that particular incentive structure. After each block of 30 trials, the simulation provided the participant with feedback on their performance bonus.

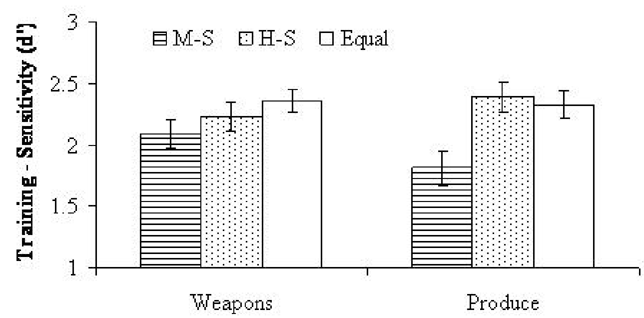
RESULTS

To recall, in agreement with the framing effect, H-S and M-S frames are expected to produce opposite results in performance, despite the equivalent problem structure. Also, stronger framing effects were expected in the Weapons compared to the Produce context. Results were presented using a 3 (Framing Structure) x 2 (Context) ANOVA. A series of Independent Samples t-tests were performed if main or interaction effects were found.

Sensitivity (d')

Training phase. Figure 2 presents the mean and standard error of d' for each of the framing structures and contexts during training and transfer. A main effect for Framing Structure was found $F(2, 84) = 6.51$, $p < .01$. No significant d' differences were found between the H-S and the Equal levels. However, the Equal level ($M = 2.35$, $SD = 0.39$) had a greater d' than the M-S level ($M = 1.94$, $SD = 0.53$), $t(57) = 3.29$, $p < .01$. The H-S level ($M = 2.31$, $SD = 0.47$) had a greater d' than the M-S level, $t(59) = 2.86$, $p < .01$. No significant effects were found for Context and a Framing Structure X Context interaction.

Transfer phase. A main effect for Framing Structure was found at transfer $F(2, 84) = 7.04$, $p < .01$. No significant d' differences were found between the H-S and the Equal levels. The Equal level ($M = 2.28$, $SD = 0.61$) had a greater d' than the M-S level ($M = 1.82$, $SD = 0.74$), $t(57) = 2.59$, $p < .02$. The H-S level ($M = 2.21$, $SD = 0.63$) had a greater d' than the M-S level, $t(59) = 2.20$, $p < .05$. In contrast to the training results, at transfer there was a main effect for Context $F(1, 84) = 50.63$, $p < .01$. The Produce context ($M = 2.49$, $SD = 0.63$) had a greater d' than the Weapon context ($M = 1.70$, $SD = 0.49$), $t(88) = 6.62$, $p < .01$. No significant Frame X Context interaction was found.



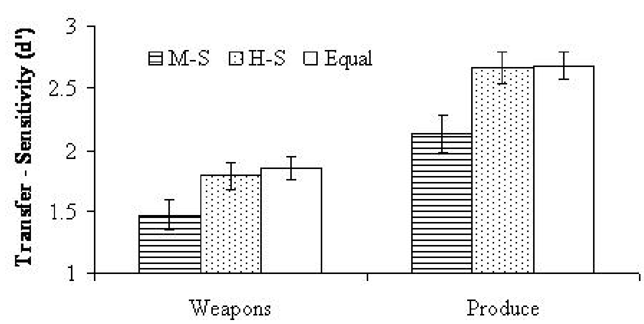


Figure 2. Training and transfer Sensitivity (d').

Response bias (lnβ)

Training phase. Figure 3 presents the mean and standard error of lnβ for each of the framing structures and contexts during training and transfer. A main effect for Framing Structure was found $F(2, 84) = 7.73$, $p < .01$. No significant lnβ differences were found between the H-S and the Equal levels. However, the Equal level ($M = 1.71$, $SD = 0.61$) also had a greater lnβ than the M-S level ($M = 1.03$, $SD = 0.86$), $t(57) = 3.47$, $p < .01$. However, the H-S level ($M = 1.55$, $SD = 0.79$) had a greater lnβ than the M-S level, $t(59) = 2.44$, $p < .02$. A main effect for Context was also found $F(1, 84) = 16.21$, $p < .01$. The Produce context ($M = 1.72$, $SD = 0.75$) had a greater lnβ than the Weapon context ($M = 1.13$, $SD = 0.76$), $t(88) = 3.70$, $p < .01$. No significant Framing Structure X Context interaction was found.

Transfer phase. A main effect for Framing Structure was found $F(2, 84) = 11.00$, $p < .01$. No significant lnβ differences were found between the H-S and the Equal levels. However, the Equal level ($M = 2.16$, $SD = 0.95$) had a greater lnβ than the M-S level ($M = 1.21$, $SD = 1.03$), $t(57) = 3.67$, $p < .01$. The H-S level ($M = 1.94$, $SD = 0.85$) had a greater lnβ than the M-S level, $t(59) = 3.01$, $p < .01$. A main effect for Context was also found $F(1, 84) = 25.63$, $p < .01$. The Produce context ($M = 2.20$, $SD = 1.00$) had a greater lnβ than the Weapon context ($M = 1.32$, $SD = 0.84$), $t(88) = 4.51$, $p < .01$. No significant Framing Structure X Context interaction was found.

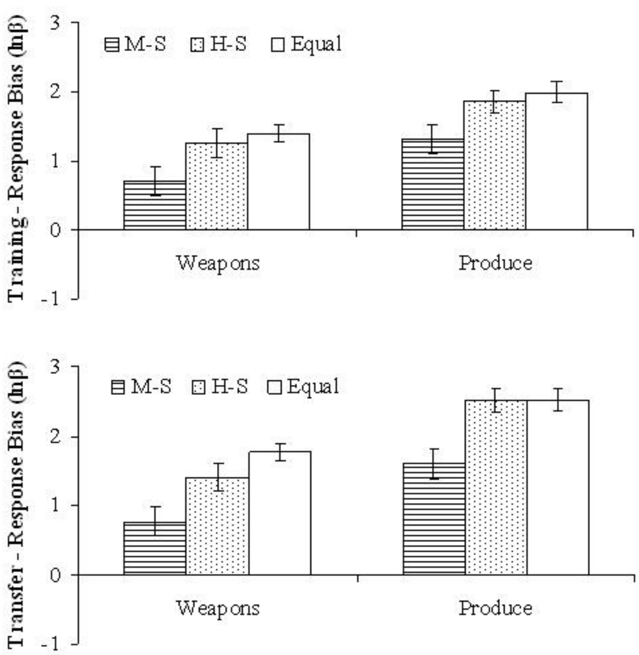


Figure 3. Training and transfer Response Bias (lnβ).

DISCUSSION

As expected, participants with a miss-focused incentive structure consistently exhibited detection behaviors closest to optimal compared to the control incentive structure. Explanations for these consistent miss-sensitive response bias differences can be explained by framing effects. Participants receiving negatively framed information were risk seeking; they indicated signal presence more often to avoid the relatively large penalty for miss outcomes. Participants indicating 'signal present' even when the signal was not visually detected may have adopted this strategy.

However, participants receiving hit-sensitive structures acted in a similar fashion than the control incentive structure. The relatively high value given for a hit outcome may not have been enough to counter consistent effects in visual search such as conservatism in setting response bias (Healy & Kubovy, 1981). Due to a base rate significantly lower than chance, it is often statistically correct to indicate: 'no, signal absent.' when one is unsure of the signal item's presence.

While the miss-focused incentive structure led to a more optimal decision strategy in terms of response bias, it also led to reduced detection sensitivity. The reduced sensitivity was possibly consequence of participants choosing to search the bag even when the base rate statistically supported passing the bag. Visual inspection of the training and transfer results suggest that detection behaviors learned in training were consistent in transfer.

Context effects also supported our expectations; the prevention focus exhibited by the weapons context led to a more optimal detection strategy compared to a produce context. These effects are similar to the ones exhibited by the miss-sensitive framing structure and can be explained in a similar manner. The participant's goal in a prevention focus context such as airline luggage screening invokes obligations such as national security. However, in the unknown-incentive transfer environment, reversals in sensitivity were found in the produce context. Distinguishing between previously unseen fruit and vegetables produce can be seen as much easier than distinguishing between unseen dangerous items and common travel items. In addition, context effects include a focus on desires and positive outcomes (promotion-focus) or responsibilities and negative outcomes (prevention-focus), known as *regulatory focus* (Higgins, 1997).

Framing by context interaction effects were expected, but not shown in this study. A lack of interaction effects may be explained by a lack of establishing the intended context with the chosen participant set. Task context may be established over a long period time as users experience the non-monetary benefits and costs of their decisions. Additional non-search factors such as queuing (Marin, Drury, Batta, & Lin, 2007) may strengthen the task's representativeness to airline luggage screening.

Potential future areas of study with framing and context effects in visual search include the use of measures and techniques of the Cutoff Reinforcement Learning Model (Erev, 1998; Erev, Gopher, Itkin, & Greenshpan, 1995). Additionally, fixing the signal and noise distributions, known as an *external noise paradigm* (Gopher et. al, 1995), may help better isolate specific response bias effects. Exploring more salient incentive structure frames and contexts to match the structure of high-stakes decisions (Kunreuther et al., 2002) is another promising direction. Changing the task from a yes/no signal detection task to a rating task may provide additional insight to how participants behave in unsure conditions.

In summary, considering regulatory and framing effects in the perception of incentive structures can be used to inform the design of training regimes and automated decision support for airline luggage screening or similar applied visual search tasks.

ACKNOWLEDGEMENTS

This research was partially supported by the Multidisciplinary University Research Initiative Program (MURI; N00014-01-1-0677). The experiment reported here was done at Carnegie Mellon University, when all the authors were part of the Dynamic Decision Making Laboratory. We thank Varun Dutt for programming the visual search simulation used in this study. We also thank the Pacific Science and Engineering Group for their guidance and support during the preparation of this manuscript.

REFERENCES

Dzindolet, M.T., Pierce, L.G., Beck, H.P., & Dawe, L.A. (2002). The perceived utility of human and automated aids in a visual detection task. *Human Factors, 44(1),* 79-94.

Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychological Review, 105(2),* 280-298.

Erev, I., Gopher, D., Itkin, R., & Greenshpan, Y. (1995). Toward a generalization of signal detection theory to n-person games: The example of two person safety problem. *Journal of Mathematical Psychology, 39,* 360-375.

Gonzalez, C., Dana, J., Koshino, H., & Just, M. (2005). The framing effect and risky decisions: Examining cognitive functions with fMRI. *Journal of Economic Psychology, 26(1),* 1-20.

Gramopadhye, A.K., Bellow, B.J., Gopinath, M., Budgavi, M. (1997). An evaluation of economic and performance feedback in an inspection task with explicit economic consequences. *International Journal of Industrial Ergonomics, 20(4),* 327-337.

Healy, A.F. & Kubovy, M. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 344-354.

Higgins, E.T. (1997). Beyond pleasure and pain. *American Psychologist, 52,* 1280-1300.

Kunreuther, H., Meyer, R., Zeckhauser, R., Slovic, P., Schwartz, B., Schade, C., Luce, M. F., Lippman, S., Krantz, D., Kahn, B., & Hogarth, R. (2002). High stakes decision making: Normative, descriptive, and perspective considerations. *Marketing Letters, 13(3),* 259-268.

Lacson, F. C., Wiegmann, D.W., and Madhavan, P. (2005). Effects of attribute and goal framing on automation compliance and reliance. *Proceedings from the Human Factors and Ergonomics Society 49th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

Levin, I.P., Schneider, S.L. & Gaeth, G.J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes, 76(2),* 149-188.

Madhavan, P. & Gonzalez, C. (2006). Effects of sensitivity, criterion shifts and subjective confidence on the development of automaticity in airline luggage screening. *Proceedings from the Human Factors and Ergonomics Society 50th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

Madhavan, P., Gonzalez, C., and Lacson, F.C. (2007, October). Differential base rate training influences detection of novel targets in a complex visual inspection task. *Proceedings from the Human Factors and Ergonomics Society 51st Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

Marin, C.V., Drury, C. G., Batta, R., & Lin, L. (2007, October). Human factors contributes to queuing theory: Parkinson's law and security screening. *Proceedings from the Human Factors and Ergonomics Society 51st Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

McCarley, J.S., Kramer, A.F., Wickens, C.D., Vidoni, E.D., & Boot, W. R. (2004). Visual skills in airport-security screening. *Psychological Science 15(5),* 302–306.

Nygren, T.E. (1997). Framing of task performance strategies: Effects of performance in a multiattribute dynamic decision making environment. *Human Factors, 39(3),* 425-437.

Parasuraman, R. (1986). Vigilance, monitoring, and search. In K. Boff, L. Kaufman, & J.Thomas (Eds.), *Handbook of perception and human performance. Vol. 2: Cognitive processes and performance.* New York: Wiley.

Perrin, B.M., Barnett, B.J., Walrath, L., & Grossman, J.D. (2001). Information order and outcome framing: An assessment of judgment bias in a naturalistic decision-making context. Human Factors, 43(2), 227-238.

Pisano, E.D., Gatsonis, C., Hendrick, E., Yaffe, M., Baum, J.K., Acharyya, S., et al. (2005). Diagnostic performance of digital versus film mammography for breast-cancer screening. New England Journal of Medicine, 353, 1773-1783.

Swets, J.A. (1992). The science of choosing the right decision threshold in high-stakes diagnostics. *American Psychologist, 47,* 522-532.

Tversky, A. & Kahneman, D. (1984). The framing of decisions and psychology of choice. *Science, 211(4481),* 453-458.

Wickens, C. D. & Hollands, J. G. (2000*). Engineering psychology and human performance*. Upper Saddle River, NJ: Prentice Hall.