A Cognitive Basis for Friend-Foe Misidentification in Combat

Joseph R. Keebler, Lee W. Sciarini, Florian Jentsch, and Denise Nicholson

*University of Central Florida, Orlando*
A Cognitive Basis for Friend-Foe Misidentification in Combat

It is important for members of any modern military to quickly and reliably discriminate between friendly and enemy vehicles. Otherwise, friendly fire incidences occur, also known as “fratricide.” During World War II and the Vietnam War, 2% to 3% of U.S. casualties were due to fratricide. Forty-five years later, in Operation “Desert Storm,” the first Iraq war saw rising rates, with overall casualties due to fratricide near 24% (Koehler, 1992). In fact, fratricide accounted for 77% of the combat damaged sustained by U.S. armored vehicles during this conflict (Koehler).

According to Briggs and Goldberg (1995), the potential for fratricide from inaccurate recognition is immense even under favorable environmental conditions. If inaccurate recognition can lead to higher levels of false alarms (i.e., what fratricide is when placed in a signal detection framework), then the opposite situation, i.e., that of misses, or mistaking an enemy vehicle for a friendly vehicle, should also be a relevant problem. We will refer to both types of errors in target identification throughout this paper under the global term misidentification.

Factors Involved in Misidentification

Misidentification may be a result of many different factors: inadequate training, confusion during battle, decision making under pressure, and stimulus similarities among friendly and enemy vehicles. There is also the problem of bias on the battlefield. Specifically, it has been shown that during combat, soldiers are more likely to judge an unknown vehicle as a foe (Briggs & Goldberg, 1995). Although humans are capable of recognizing objects of many shapes, forms, and colors under various conditions, battlefields are complex environments that can mask targets and confuse soldiers, making identification of potential threats much harder.
Decisions made under the different pressures exerted on the battlefield therefore can lead to more errors in detecting and identifying targets.

*The “fog of war.”* One factor related to target misidentification is what is known as the “fog of war”. This term is used to describe the level of ambiguity in situation awareness experienced by those conducting military operations (Joint Service Command and Staff College, 2001). The fog of war can be described as a combination of factors: featureless terrain; large complex and fast-moving formations; fighting in precipitation, darkness or low visibility; and the ability to engage targets from long distances. These factors all contribute to friendly fire incidents (Koehler, 1992). Training, attention, similarity between targets and non-targets, and timed decision making, all play a vital role in making correct judgments. Although the fog of war will always be present (Koehler), we believe that through technology and training, some of these negative outcomes can be alleviated, and the ability to correctly identify military vehicles can be increased and enhanced (Biederman & Shiffrar, 1987).

*Stimulus similarities.* Similarities among vehicles from different militaries are yet another problem that can lead to misidentification. If a vehicle is similar enough to another one, it can easily be mistaken for a friend or an enemy, when the opposite is true. Frontal views of vehicles are often associated with more mistakes in identification as they show fewer of the necessary and salient cues (Briggs et al., 1995). Tanks, for example, share fundamentally the same basic shapes and construction, especially when seen head-on (i.e., a turret on top of a two-track undercarriage; Keebler et al., 2007) and thus provide a strong example of pan-military vehicle similarity.

*Training.* One of the traditional “antidotes” to combat misidentification has been vehicle identification training. Generally, having high vehicle image fidelity is not important for proper
training, provided that the critical cues (e.g., chassis, turret shape, mortar tubes, etc.) are visible (Briggs et al., 1995). Nonetheless, researchers have questioned whether using 2D media (cards, images in a computer simulation, etc.) is effective for familiarizing soldiers with the vehicles that need to be identified and discriminated in a 3D environment (Briggs et al.). In fact, as critical cues become less salient, the difficulty of the identification task rises drastically. Further, 2D training aids usually show only a limited number of views, whereas vehicles are seen in the real world from any angle and orientation. This suggests that using 2D training aids may not be effective at facilitating accurate vehicle identification. Later, we will refer to research that we have conducted that may lend support to the idea of using 3D objects as superior training media.

**Purpose of this Paper**

To help explain some types of misidentification problems, we will look to Treisman’s Feature Integration Theory (FIT) (Treisman & Gelade 1980). This theory can lend support towards a better understanding of what is happening to a soldier’s visual perception and processing on the battlefield due to attention distractions, and specifically how and why a soldier would mistake the identity of one vehicle for another when placed under the stress of being in the midst of a battle.

We will also look at memory, and the model for memory we will be discussing in this chapter will be Baddeley’s Working Memory Model (Baddeley & Wilson 2002). Baddeley’s model treats the brain as if it were a computer. Much like a computer has a hard disk to store information and a microprocessor to make calculations, we too as human beings have analogous processes in our perceptual systems.

Finally, to explain the possible advantage of using 3D models as a novel, and potentially better method for training, we will turn to Biederman’s Recognition By Components Theory
According to RBC Theory, the ability to recognize an object comes from organizing a limited number of constituent “parts,” referred to as geons, into an “object model”, which is a mental representation of the perceived object. Together, the three theories (i.e., FIT, WM, and RBC) provide a strong theoretical foundation that can be used to conceptualize combat identification and to make predictions about ways in which misidentification can be reduced.

The “Fog of War” and Feature Integration Theory (FIT)

Anne Treisman developed a cognitive theory that gives insight into what happens to the human perceptual system when making an identification decision during a battle (Treisman & Gelade 1980). According to Treisman’s model, **attention** plays an integral role in perception (see Figure 1). Other research has even argued that perception is impossible without attention (Briand et al. 1987). Analogous to a spotlight shining into the darkness to illuminate an area, attention can only focus on a small part of the entire visual field. When focused, this “spotlight” allows for temporary object representations to form in the observer’s mind which then is compare to stored memories (Treismann, 1986). When unfocused, however, these temporary object representations can become chaotic, and information can be mixed up within and between objects. This is known as “**illusory conjunction**” and is one of the integral findings in Treismann’s research. An illusory conjunction is an occurrence of misperception, where an object appears to have a different shape, size, or color then it actually has in reality (Treismann). For example, during one of Treisman’s experiments, different colored letters (X’s & O’s) placed in the visual periphery were misperceived (a red X would be shown with a yellow O, but participants would report a yellow X and a red O) significantly more than chance, while participants were involved in another task that consumed most or all of their attention.
If we subscribe to this theory, then it could be argued that a soldier, when under the pressure of combat, must focus attention to properly make decisions. If the soldier cannot provide adequate attention to a vehicle to allow for a temporary object representation to form, then they will be unable to make correct judgments as to the identity and alliance of that vehicle. This leads to a critical need for those soldiers performing identification tasks to attend to that task alone. Any other task involving even minimal amounts of attention could diminish perceptual resources needed to make correct identifications.

*Training Media and Recognition by Components Theory (RBC Theory)*

In addition to Treisman’s FIT, Biederman’s (1987) Recognition by Components (RBC) theory is helpful in providing information about instances of combat misidentification. As we will show as well, RBC Theory can also explain why 3D models could be a more useful training tool.
tool than line drawing equivalents or simulation vehicles. Biederman’s cognitive model suggests that objects can be broken down into a limited number of constituent “parts” referred to as **geons** (geometric icon), the basic volumetric form of our object perception. Geons are represented by many objects such as cubes, cones and spheres. The model further suggests that we recognize objects by initially firing “feature” detectors, and that these feature detectors locate basic geons in the objects we perceive. Our ability to recognize an object comes from organizing present geons into an “object model”, which is a mental representation of the perceived object (Biederman et al., 1991). These geons may be very similar to the object representations explained in FIT theory, FIT theory, however, does not include a descriptive alphabet of the actual objects used in perception. This leads to a need to integrate Biederman’s model alongside Treisman’s.

When perceived geons are properly matched to those in memory, observers can reliably identify the object. If the object is physically or perceptually obscured, however, the observer may not perceive all of its components. In this case, improper identification of vehicles can lead to misidentification. As in FIT, identifying critical components and cues is pertinent to lowering error rates and increasing accurate identification. Targets are often misidentified since battlefield observers are rarely able to see an entire vehicle because of obscuration by terrain, dust, smoke, camouflage and poor illumination (Briggs et al. 1995). This could be interpreted to mean that observers are missing critical cues of the vehicle that allow for proper identification, and therefore make decisions based on the available componential information.

When identifying and discriminating between military vehicles, the RBC model suggests that an observer would first notice the basic outline and structure of the vehicle’s geons: the
turret, the main gun and the body. Many tanks and armored vehicles throughout the world have geons that are arranged in the exact same fashion.

The similarity in the arrangement of these vehicle geons could make it difficult to discriminate between a friendly vehicle and an enemy one. From certain views, especially the frontal view, it is almost impossible to tell many of these vehicles apart from one another (see Figure 2). Researchers have suggested that frontal views of tanks often force rapid decision making, therefore presenting a potentially dangerous situation (Briggs et al., 1995).

Figure 2. Similar frontal views of a Russian T-72 and a German Leopard II.

We believe that this theory supports the idea that there may be a considerable benefit in using models of military vehicles as training devices. As mentioned earlier, 2D training media may not be the most effective medium for training military vehicle identification (Keebler et al 2007). Chang (2003) suggested that objects in a 2D space present a problem due to the restriction in the way the object can be viewed or posed. In other words, 2D representations poorly depict the full range of angles that a 3D object presents. According to RBC, 2D representations contain only a few surfaces of the present geons, while 3D models present the geons fully intact.

Pierno et al. (2004) suggested that 3D objects also provide a perspective-based visual cue that may be easier to utilize than 2D visual cues. The argument for 3D training has a strong
foundation, and should be introduced in future investigations, given that a 3D cue is more easily remapped onto real-world coordinates. If 3D cues are more readily mapped then specific training on cues should aid in enhancing correct judgments (Briggs et al., 1995). Further, Pierno et al. found that experiencing a 3D based object rather than a 2D based object produced faster target acquisition times. Friedman, Spetch, and Ferrey (2005) also argued that the depth cues available in 3D objects may facilitate tank identification. These authors also suggested that 3D objects provide a more detailed account of an object. Accordingly, 3D data should yield better tank identification and recognition performance (see Chang et al., 2003). In addition, past studies have indicated that 3D training of plate tectonics led to significantly better performance than a similar 2D training method (Kim, 2006), and that training on 3D models led to significantly higher scores on a tank identification task (Keebler et al., 2007).

If training can be altered in such a way as to improve the detail of the object model that will be stored in the trainee’s memory, we can potentially mitigate some of the problems which contribute to misidentification. When attention can be focused enough to perceive the components of potential threat vehicles, then training must be powerful enough to make sure those components are matched to their proper memory store. If this occurs, there can be an increase in correct identification decisions being made. Biederman’s model thus supports potential training enhancements with the introduction of 3D training media.

**Shape Memory (ETS MV-1) and Baddeley’s Working Memory Model**

According to Baddeley’s working memory model, the mind functions much like a computer system (Baddeley & Wilson, 2002). There are three main resources in this model: two “slave memory” areas that allow observers to store (a) visual information (i.e., the visual-spatial
sketchpad) and (b) auditory information (i.e., the phonological loop), and a third processing area called the Central Executive (CE). The CE takes information from the two perceptual slave-memory areas, and processes it, comparing it to old memories, expectancies, and schemata, so as to make decisions about the world. Baddeley’s model is integral to this chapter in the sense that an identification task requires comparing the temporary storage of what is being perceived to a long term memory version of the same object. Correctly aligning these two objects should lead to correct judgments in identification. However, to test this notion, it is very important to somehow measure the amount of space or level of detail in the visual-spatial sketchpad. We propose here that the shape memory test (ETS MV-1) may be a measure which quantifies the quality or amount of storage capacity in the sketchpad of an individual. Indeed, we theorize that the Shape Memory Test should be predictive of performance on identification tasks, demonstrating that the more information an individual can temporarily store on their sketchpad, the better they will perform when asked to recall and compare that information during a performance task.

Below, we will reference two studies conducted at Team Performance Laboratory (TPL) in 2007-2008. We believe that this research can shed some light on the theoretical issues discussed above.

Experimental Evidence

The following studies were designed to compare the utility of using 3D models over a current form of training, namely military issued 2D cards containing line drawings. Both studies used the MV-1 Shape memory test as a pre-measure of memory capacity. Also, we introduced a simulated observation post to the second study, to investigate effects due to immersion. As previously discussed, 3D models provide a richer training media, and therefore should lead to
higher performance scores on identification tasks (Keebler et al., 2007). The mental “object model” referred to in both Biederman’s and Treismann’s theories was possibly developed more thoroughly by the 3D training media when compared to 2D media.

Experiment 1

To test our hypotheses, we designed a $2 \times (2 \times 3)$ mixed-model factorial design where training (2D vs. 3D) was the between subjects variable, test media (2D vs. 3D) was the first within subjects variable, with test knowledge area (identification, recognition, and friend/foe differentiation) as the other within-subjects variable. To help clarify the terminology for these studies, we used recognition in the sense of whether someone recognized if they had seen the vehicle before or not. Identification was the explicit naming of the test vehicle, while friend/foe differentiation consisted of knowing whether the vehicle was friendly or a foe.

Participants

Twenty student volunteers participated in this study. The participants’ ages ranged from 18-32, with a mean age of 24. Fourteen of the participants were female, and six were male. There were no participants with prior military experience.

Materials and Procedure

Typical pre-experimental materials included informed consents and demographics forms. The participants then completed the ETS MV-1 (Shape Memory Test). Participants were randomly assigned to one of two conditions: training with cards or training with models. Those in the card training condition viewed six military issue Armored Vehicle Recognition cards. In contrast, those in the model training condition viewed 1:35 scaled models of the same six vehicles (see Figure 4). The participants were each given six minutes to study the vehicles.
The order in which the tests were administered was counterbalanced across participants. The identification test required participants to recall the name of the vehicle; the recognition test required the participants to write “yes” or “no” to the question of whether they had seen a specific vehicle during the exposure training session. For the friend/foe test, the participant was asked to write down whether a vehicle was either a friend or foe.

Testing Conditions

Card testing. For the card test, the cards were all copied onto a sheet of paper. The participants viewed the same graphic cards presented during the exposure training in a different order than was presented in training, and with two extra vehicles included as distracters. Participants answered the identification, recognition, and friend/foe questions on blanks provided on the test sheet.

Model testing. The same 1:35 scale vehicles were used for the model tests as in the training. The vehicles were reorganized on a table, and two extra vehicles were added as distracters. Participants were asked to answer the identification, recognition, and friend/foe questions.

Results

A repeated-measures multivariate analysis of variance (MANOVA) was used to test whether the participants in the model exposure training group outperformed the participants in the card exposure training group on three dependent measures of performance (identification, recognition, and friend/foe) using \( p < .05 \). Wilk’s Lambda multivariate test indicated a statistically significant main effect for testing \( (F(5, 14) = 5.09, p < .01, \text{partial } \eta^2 = .64) \). Furthermore, Wilk’s Lambda indicated a statistically significant interaction for Training x
Testing \((F(5,14) = 5.61, p = .005, partial \eta^2 = .67)\). A one-tailed Duncan’s post hoc analysis was chosen to further analyze the interaction effects.

**Card Training vs. Model Training for Identification.** At \((p < .01)\), the post hoc analysis indicated that the card training group performed better on the card identification test \((M = .63, SD = .37)\) than on the model identification test \((M = .35, SD = .28)\). The model training condition performed better on the model identification test \((M = .57, SD = .35)\) than the card identification test \((M = .40, SD = .34)\) \((p = .04)\).

**Card Training vs. Model Training for Recognition.** Participants in the card training condition performed significantly better on the card recognition test \((M = .89, SD = .11)\) than on the model recognition test \((M = .63, SD = .12; p < .01)\). There was no significant difference between the cards and models on the recognition test for the model training group.

**Card Training vs. Model Training for Friend/Foe.** The card training condition for both the card friend/foe test and the model friend/foe test had means of \(M_{\text{Cards}} = .50\) and \(M_{\text{Models}} = .52\). These means were not compared to the model training condition because they indicated that the participants were doing no better than chance. There was no significant difference between card \((M = .65, SD = .19)\) and model \((M = .67, SD = .34)\) friend/foes scores for the model training group.

**Shape Memory X Training Interaction.** Using multiple regression analysis, a significant interaction was found between shape memory and training condition. At step 1, our \(R^2 = .141, F(2, 17)=1.392, p = .275\). With the addition of the interaction term at step 2, the amount of variance explained rose to \(R^2 = .465, F(1,16) =9.707, p < .007\). This suggests that the combination of model training materials and shape memory was predictive of performance on a military vehicle identification task.
Experiment 2

Participants

Fifty eight volunteer undergraduate students participated in this study (39 males, 19 females). Participants all received credit in an undergraduate psychology course for participation in this study.

Materials and Procedure

The pre-experimental materials included the typical informed consent and demographics form, and post-experimental debriefings were also administered. Before viewing the military vehicles, participants were administered the Shape Memory Test. After the Shape Memory Test, participants were randomly assigned to one of two conditions: card training or model training. Those in the card training condition viewed nine military issued Armored Vehicle Recognition cards. Those participants in the model training condition viewed nine 1:35 scaled models of the same tanks. The participants were each given one minute per vehicle to study the tanks (for a total of nine minutes).

The participants were brought into a simulated foxhole for testing. The foxhole consisted of a sandbag emplacement, covered by a camouflage net. The foxhole was in a darkened room, and projectors on the ceiling projected images of the surrounding terrain onto a 180-degree wide projection area in front of the foxhole. A vehicle would be placed into the simulation, at which time it would appear as a target on the simulation screen, appearing to be approximately 300 meters away. The participants were asked if they could see the tank in the simulation, and then
told to look through simulated binoculars, which had a “zoomed” image equivalent to a distance of about 50 meters. After ten seconds had passed, the vehicle was removed from sight.

Results

To test the hypotheses, we designed a 2 x 3 mixed model factorial study where training (cards vs. models) was the between-subjects variable. Test format (identification, recognition, and friend/foe) was the within-subjects variable.

Card Training vs. Model Training for Identification. An independent samples t-test was executed to investigate if there was significant difference between the means of the two conditions. The test results showed no significant differences between the two conditions, \( p > .05 \).

Shape Memory and Target Identification. A multiple regression correlation was used to test our hypothesis concerning shape memory. Controlling for knowledge of military vehicles and experimenter ratings of participant interest level, it was found that shape memory was predictive of target identification performance, although an interaction was not present this time. Step 1, which included familiarity with military vehicles and interest level, explained about 35% of the variance in performance, i.e., \( R^2 = .348, F(2,58)=14.4, p<.0005 \). Step 2, which added shape memory to the equation, was significant and explained about 12 percentage points more variance, \( R^2 = .477, F(2,57)=16.06, p<.0005 \).

Discussion

2D and 3D identification. Based on the results of the study, our hypothesis that 3D training would lead to better performance on a target identification tasks was not fully supported. Instead, when asked to recall the name of a tank, the participants did better on the tests that matched the media they were exposed to during training. Although we believe that 3D models
have stronger external validity and may provide more realistic training than line drawings on cards, this effect did not reach significance. In the 2D training the participants were only receiving partial views of the vehicles, while in 3D they were able to see the entire vehicle. This finding is consistent with a study by Hah, Reisweber, Picart, and Zwick (1997) in which participants were trained with whole and partial views. They found training with whole views outperformed partial view training. This lends support to Biederman’s theory in that 3D training provides a higher number of complete geons, and therefore, a more enriching training medium.

2D and 3D recognition. When participants were asked whether they had seen the vehicles that they were previously exposed to, the 2D training group did better on the 2D test than the 3D. This suggests that when viewing a 2D object, such as an x-ray, recognition may be better if the viewer has been trained using 2D training aids rather than 3D training aids. The remaining mean test scores for recognition were consistent. This suggests that the 3D recognition task may be trained using either 2D or 3D training aids. Further investigations should be conducted to find the effects of combining both 2D and 3D media in training.

2D and 3D friend/foe accuracy. The Friend/Foe test scores indicated that the participants in the 2D training condition were doing no better than chance when asked to identify a friendly tank from an enemy one. On the other hand, the 3D training group produced consistent Friend/Foe scores for both 2D and 3D testing. 3D training showed a high consistency, and this may have something to do with the amount of information gathered from the training. Specifically, this may be due to the larger amount of geon information provided by the 3D training media. These findings suggest that 3D media may be better for training of discrimination tasks.
Based on the results of this study, we concluded that the addition of 3D training media may have important implications. We also concluded that improving vehicle recognition tasks may require both approaches to training. For example, 3D training may be useful for naming and discriminating tanks and other vehicles, while 2D training can be used for recognition tasks. As a result, we suggest that 3D training media should be incorporated into training programs for 3D object recognition tasks.

In sum, the findings from Experiment 1 suggested that a more multi-faceted training approach should be considered. Training needs to take into account the various aspects of the tank recognition task, and match the training modality appropriately. In addition, this study approached vehicle recognition from only the perceptual aspect of the task. Other cognitive skills, such as decision making are required for the task. Both the perceptual aspects and the cognitive aspects of the tasks need to be studied in conjunction.

**Shape memory X training interaction.** A statistically significant interaction was found between the Shape Memory Test and training media. Specifically, those trained on the 3D models significantly outperformed the rest of the sample, with improved results as shape memory increased. This is support for our theoretical framework, as far as Biederman’s and Baddeley’s theories are concerned. With the addition of an immersive environment in Experiment 2, we found that this interaction disappeared, only to be replaced by a main effect for Shape Memory. This finding shows that visual memory, above and beyond training, can have important implications for identification performance under high stress conditions.

**Conclusions**

The main focus of this chapter is to discuss some of the major cognitive theories involved in the process of identification. Designing training applications and systems around this
theoretical framework should alleviate some of the issues associated with misidentification.

Whether from the sensor of an unmanned vehicle, a photograph from a reconnaissance plane, or a soldier’s direct line of sight in the field; many factors influence the ability of correctly identifying vehicles. The first implication we can take away from our discussion is that individuals tasked with vehicle identification must pay attention to the task. While the possibility of having someone solely dedicated to the task of identifying vehicles is improbable, the importance of applying full attention cannot be stressed enough. Without attention, there is no perception.

The next point is the importance of individual differences in working memory. Through the experiments discussed above, we have found that having higher scores on tests of visual shape memory can lead to higher performance in identification tasks. Making sure that our soldiers and/or operators have high visual memory is critical for them to properly perceive and identify the vehicles that they may encounter. Screening for this capacity should be an important process in the selection of soldiers and/or operators who will be involved in the identification of vehicles.

The addition of scale models to training programs will be the final consideration of this chapter. Both studies mentioned above have introduced a novel training method, namely that of utilizing 1:35 scale models as training devices. As discussed in Biederman’s theory, these highly detailed, realistic reproductions of actual vehicles can provide a richer representation of the geons needed to perceive, remember, and identify the proper military vehicles. Adding scale models to current forms of training (cards, simulations, etc.), according to our data, should be beneficial. The more realistic and detailed an object is during training, the more realistic and detailed the mental representation. This mental representation may be the only piece of
information that a soldier can rely on when they are called upon to make critical decisions. Making sure these mental representations are as reliable, detailed, and useful as possible is the purpose of effective training in identification.

References


