Evidence for a Hierarchical Structure Underlying Avoidance Behavior

Mieke Declercq and Jan De Houwer
Ghent University

In studies on avoidance learning, a warning signal is followed by an aversive unconditioned stimulus unless the participant performs a designated response. The authors examined whether avoidance behavior can be based on hierarchical knowledge, that is, knowledge about the conditions under which certain relations hold. In the present study, a single avoidance response had different effects depending on the nature of the warning signal. Results showed that participants acquired this hierarchical knowledge and used it to avoid negative outcomes. The results are in line with an occasion setting account of avoidance learning and can be explained also by a modified version of Lovibond’s (2006) account of avoidance learning.

Keywords: avoidance, hierarchical structure, occasion setting, humans, instrumental

A typical avoidance learning procedure involves three elements: a warning signal (also called the discriminative stimulus or Sd) is followed by an unconditioned aversive stimulus (US) unless a designated response (R) is executed. For example, in the classic studies of Solomon and Wynne (1953), dogs were given an electric shock (US) after a light was turned on (Sd) unless they jumped over a hurdle (R).

The study of avoidance behavior has been guided by different theoretical accounts. One of the most influential accounts is the two-factor theory of Mowrer (1947). This theory suggests that an association is formed between two elements of the learning procedure, namely the Sd and the avoidance response (stimulus–response [S–R] association). The reinforcement that drives this operant S–R learning is assumed to originate from Pavlovian learning. Because of the repeated pairings of the Sd with the US, the Sd will start to elicit a conditioned fear response. The reduction of this conditioned fear that occurs when the avoidance response is emitted, serves as a reinforcer for this response. In this two-factor theory, the US no longer has a role once the conditioned fear for the Sd is established. Despite its popularity, the theory has not escaped criticisms. As a result, new accounts of avoidance learning were developed.

One of these new accounts was proposed by De Houwer, Crombez, and Baeyens (2005; for other theories of avoidance learning, see Gray, 1987; Seligman & Johnston, 1973, and Lovibond, 2006). De Houwer et al. (2005) postulated that an avoidance response can function as a negative occasion setter. From a structural point of view, the similarity between negative occasion setting and avoidance behavior is very clear: A negative occasion setter signals when another stimulus (also called the target) will be followed by the US. If the negative occasion setter accompanies the target, then the US will not follow. If the occasion setter is not presented, then the target will be followed by the US (for a review of the literature on occasion setting, see Holland, 1992, and Swartzentruber, 1995). Avoidance behavior also signals when the target (i.e., Sd) will be followed by the US: If an avoidance response is performed, then the target will not be followed by the US. If the avoidance response is not performed, the target will be followed by the US. De Houwer et al. (2005) and Declercq and De Houwer (in press) reported evidence that a negative occasion setter and an avoidance response are similar at a functional level as well. According to Holland (1992), a negative occasion setter possesses two unique properties. The first and most crucial property is that the modulatory function of a negative occasion setter is resistant to counterconditioning. This means that the occasion setter’s power to modulate responding to a target is unaffected by trials on which the occasion setter on its own is followed by the US. Modulation thus appears to be independent of the direct relation between the occasion setter and the US. The second unique property is selective transfer. This refers to the fact that the modulatory power of an occasion setter will transfer to other targets in a selective way: It will transfer especially to targets whose relation with the US was previously modulated by another occasion setter. Although these two properties are the defining properties of negative occasion setting, it should be noted that the second is less well understood and observed less consistently than the first (e.g., Skinner, Goddard, & Holland, 1998). De Houwer et al. (2005) and Declercq and De Houwer (in press) established that the modulatory power of an avoidance response was resistant to counterconditioning and transferred selectively to other Sd–US relations that had been involved in occasion setting training.

The occasion setting account of De Houwer et al. (2005) differs from other theories of avoidance in that it assumes that a hierarchical structure underlies avoidance learning: It postulates that the avoidance response becomes a signal for the presence of the Sd–US relation rather than for the presence of a specific stimulus (Sd or US). We set out to test this hypothesis by adapting a procedure that Colwill and Rescorla (1990) had used to establish that hierarchical associations can be involved in appetitive instrumental learning. In one of their experiments (Colwill & Rescorla, 1990; Experiment 2) rats could make two different responses.
(R1 and R2) during the presence of a stimulus S1. Performing R1 resulted in the delivery of one positive reinforcer, US1, whereas performing R2 resulted in the delivery of a different positive reinforcer, US2. During the presence of another stimulus S2, performing R1 resulted in the delivery of the US2 and performing R2 resulted in the delivery of US1. In a second learning phase, either US1 or US2 was paired with a toxin to reduce its value. During a subsequent extinction phase, both responses were available and several presentations of S1 and S2 were delivered. The results indicated that a response was performed more frequently during the presence of a stimulus that signaled that the response would be followed by a valued US than during the presence of a stimulus that signaled that the response would be followed by a devalued US. Colwill and Rescorla (1990) argued that their results could not be explained by simple stimulus-reinforcer (S-S) associations because each reinforcer was presented with the same frequency in the presence of both stimuli. Furthermore, their results could not be explained by simple response-reinforcer (R-S) associations because the reinforcer was presented with the same frequency in the presence of both responses. A S-R theory is also not able to explain the obtained results. In these theories, it is assumed that presentations of the stimulus coupled with the reinforcer will establish a S-R relation, enabling the stimulus to generate a response. Revaluation of the USs should not affect these S-R relations. Colwill and Rescorla (1990) concluded that the observed effect was based on more than individual S-R, S-S, or R-S associations or combinations of these associations. The results can be explained only if one assumes the presence of a hierarchical knowledge structure in which the stimulus signals the relation between the response and the reinforcer.

We adapted the design of Colwill and Rescorla (1990) in the following manner (see Table 1): Participants were first familiarized with two different negative USs (US1 and US2). In a second learning phase, two different Sds were presented. Each was followed by both US1 and US2. In a third learning phase, two different avoidance responses (R1 and R2) were introduced. After Sd A, performing R1 resulted in the omission of US1 and performing R2 resulted in the omission of US2. After Sd B, this was reversed: Performing R1 resulted in the omission of US2 and performing R2 resulted in the omission of US1. In a subsequent revaluation phase, one of the two USs was revalued. In a final test phase, each Sd was presented repeatedly while both avoidance responses were available. Participants were asked to select one of the avoidance responses. We expected that participants would perform more frequently the avoidance response that avoided the nonrevalued US in the context of the presented Sd. That is, if US1 was made positive during the revaluation phase, participants should select R2 in the presence of Sd A and R1 in the presence of Sd B. The reverse should be true if US2 was revalued during the revaluation phase. Such a result cannot be explained in terms of simple associations between two elements, but rather requires an explanation in terms of a hierarchical structure.

Method

Participants
Twenty-four undergraduate students (16 women) at Ghent University participated in the experiment in exchange for course credits.

Stimuli and Materials
The experiment was run on a HP Compaq portable computer with a 17-inch screen and was programmed in Affect (version 4) software (Hermans, Clarysse, Bueyens, & Spruyt, 2005). A white square (2 cm × 2 cm) and a white triangle (base = 2 cm and height = 2 cm) served as Sds. As US1, we used a red letter X and as US2, we used a red letter Y. Both were 2 cm wide and 2 cm high. The stimuli were presented in a main frame that was 20 cm wide and 13 cm high. At the top of this main frame a smaller frame (9 cm wide and 1.5 cm high) was drawn, in which a message could appear if a certain key was available. At the bottom of this main frame, there was a smaller frame that was 5 cm wide and 1.5 cm high. In this frame a confirmation bar was presented when participants had given a valid response. This confirmation bar was a blue or green horizontal rectangle of 2 cm × 0.5 cm. Participants could perform an avoidance response by pressing a green or blue key, which were respectively the D and K key on the AZERTY keyboard. The situations in the rating phases were presented in Arial, 18. Participants could give their rating by moving with the arrow buttons over a scale from 0 to 100, confirming their choice with the enter button. All messages and instructions appeared in Dutch.

Procedure
All participants took part individually in a darkened room. After signing the informed consent form, they were seated at a distance of approximately 60 cm from the computer screen. Instructions

Table 1
Summary of the Design

<table>
<thead>
<tr>
<th>Learning phase 1</th>
<th>Learning phase 2</th>
<th>Rating phase 1</th>
<th>Learning phase 3</th>
<th>Rating phase 2</th>
<th>Revaluation phase</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>US1 = −0.25</td>
<td>A-US1US2</td>
<td>A?</td>
<td>A-R1-noUS1US2</td>
<td>A?</td>
<td>US1 = +0.25</td>
<td>A-R1R2?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B-R1-US1noUS2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B-R2-noUS1US2</td>
<td>AR1?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A-US1US2</td>
<td>BR1?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B-US1US2</td>
<td>BR2?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. US = unconditioned aversive stimulus; R = avoidance response.
* The particular US was absent only when participants pressed the appropriate key.
appeared on the screen which informed participants that they would receive €16 at the beginning of the experiment and that they would lose or gain €0.25 every time a red X or a red Y appears on the screen. Their main task was to take home as much money as possible. They were told that the value of X and Y could change during the experiment and that sometimes a green or blue key would be available. Pressing those keys might have an influence on the presentation of the red X and red Y. They were also informed that it was important to learn as much as possible during the experiment. After participants indicated that they had read the instructions, a number of practice trials were presented to show participants how they could give a valid response. On these practice trials, no Sds or USs were presented. The only available key was the space bar. After these practice trials, the first learning phase started.

In this first learning phase, US1 and US2 were each presented four times. The aim of this phase was to familiarize participants with the value of the USs. As was the case in all other phases, all trials in this phase were presented in a random order. The US was presented 1,500 ms after the onset of the frame and remained on the screen for 1,500 ms. US1 was presented in the left part of the frame and US2 was presented in the right part of the frame. Together with the US presentation, a message that informed participants that their credit diminished with €0.25 was presented at the top of the screen. The intertrial interval was 5,000 ms.

A second learning phase started immediately after the first one. In this phase, four trials in which Sd A was followed by both US1 and US2 (A-US1US2) and four trials in which Sd B was followed by both US1 and US2 (B-US1US2) were presented. The assignment of the shape (square or triangle) to Sd A and B was counterbalanced. On all trials, the Sd was presented 1,500 ms after the onset of the frame and remained on the screen for 2,000 ms. After 3,000 ms, both USs were presented on the screen (US1 in the left side and US2 in the right side of the frame). The USs remained on the screen for 1,500 ms and were accompanied by a message that informed participants that their credit diminished with two times €0.25.

After this second learning phase, participants were told that a rating phase would follow during which they had to judge the likelihood of the USs in several situations. They were also told that dependent on the accuracy of their ratings, the experimenter would decide whether they could proceed to the next learning phase. The aim of the rating phase was to verify that participants had correctly learned that both USs always followed each of the two Sds. Each situation was presented two times. On one of these occasions, participants indicated how likely it was that the US1 would be presented and on the other occasion they indicated how likely it was that the US2 would be presented. This means that in this rating phase, participants were asked to give four separate ratings: one for the likelihood of US1 being presented after warning signal A; one for US2 after warning signal A; one for US1 after warning signal B; and one for US2 after warning signal B. The description of these situations read as follows: “If you see the [name of shape], how likely is it that the [red X/red Y] will be presented?” Participants could give their ratings by moving a red dot over a scale from 0 (very unlikely) to 100 (very likely) using the left and right arrow buttons. The selected rating could be confirmed by pressing the enter key. The description disappeared after participants pressed the enter button. After 500 ms, a new description was presented. After this rating phase, the experimenter started a new learning phase only if participants had answered correctly to all described situations. As a correct answer, we considered all US expectancies that deviated less than 15 points from the objective contingency (e.g., if contingency was perfect, we considered as correct ratings between 85 and 100). If participants had not answered correctly to the described situations, the previous learning phase was presented again.

In a third learning phase, two A-US1US2 and two B-US1US2 trials were presented. Furthermore, participants received four trials in which Sd A was present and R1 was possible (A-R1-US1US2), four trials in which Sd A was presented and R2 was possible (A-R2-US1noUS2), four trials in which Sd B was presented and R1 was possible (B-R1-US1noUS2), and four trials in which Sd B was presented and R2 was possible (B-R2-US1US2). Whether the green (D) or the blue (K) key functioned as R1 or R2 was counterbalanced. The trials were identical to the A-US1US2 and B-US1US2 trials except on following points: Immediately after the Sd was presented, the message “The green key is available” or “The blue key is available” was presented for 2,000 ms in the top frame. If participants pressed the available key during the presentation of the message, a registration bar in the corresponding color of the key was presented for 1,000 ms. If participants pressed key R1 after Sd A, US1 was absent and US2 was presented. If participants pressed key R2 after Sd A, US1 was present and US2 was absent. For Sd B, the reverse was true. If participants pressed neither R1 nor R2 or pressed a key that was not available, both USs were presented. On each trial, the presentation of the USs was accompanied by a message that informed participants how much money they lost. If one US was present, they lost €0.25. If two USs were present, they lost €0.50.

After this third learning phase, we again inserted a rating phase to check whether participants had correctly learned the events presented during this phase. As in the previous rating phase, each situation (A, B, AR1, AR2, BR1, and BR2) appeared two times on the screen; once for indicating how likely it was that the US1 would be presented and once for indicating how likely it was that US2 would be presented. In total, participants had to give 12 ratings. For situations in which only a shape was presented, the description read as follows: “If you see the [name of shape] and you do not press a key, how likely is it that the [red X/red Y] would be presented?” For situations in which a shape and a color key were presented, the description read as follows: “If you see the [name of shape] and you press the [color of key] key, how likely is it that the [red X/red Y] would be presented?”

The revaluation phase started after participants gave a correct rating (see above) in response to each of the 12 questions. In this revaluation phase, both US1 and US2 were presented four times. These trials were identical to the trials of the first learning phase, except for the message that accompanied the revalued US. Now, the message told participants that their credit increased with €0.25 every time that the revalued US appeared on the screen. Whether US1 or US2 was revalued, was counterbalanced.

After the revaluation phase, a test phase was presented. Participants were told that from now on both keys would be available but that only one key could be pressed on any given trial. Again, participants were encouraged to take home as much money as possible. They were also told that they would not receive feedback about the correctness of their choice (i.e., no USs would be
presented). This was done to ensure that the outcome of the choices on the first test trial would not influence the choices on the other test trials. In the test phase, two A-R1R2 and two B-R1R2 trials were presented. On these trials, the Sd appeared on the screen 1,500 ms after the onset of the frame and remained on the screen for 2,000 ms. Immediately after the Sd disappeared, the message “The green key and the blue key are available” appeared on the screen. This message remained on the screen until participants pressed one of both keys. After participants pressed one of the keys, the message disappeared and a confirmation bar in the corresponding color of the pressed key was presented for 1,000 ms. The message “No feedback” appeared on the screen for 1,500 ms, 1,500 ms after the confirmation bar disappeared. The intertrial interval was 5,000 ms. For half of the participants, the test phase started with a A-R1R2 trial. For the other participants, it started with a B-R1R2 trial. The other three trials were presented in a random order.

Results

The second learning phase was presented only once for all participants because all participants gave correct ratings after the first presentation of this phase. After the third learning phase, 46% of the participants needed a second presentation of this phase and 8% needed a third presentation before they gave all ratings correctly. We analyzed only the data of the first test trial because the choices on the subsequent test trials are in all likelihood not independent from the first choice. It is therefore not clear what the chance level is for selecting the correct response on multiple trials. The correct response is defined as the response that leads to highest monetary outcome. The probability that participants selected the correct avoidance response on the basis of chance on the first test trial is 50%. From the 24 participants, 19 (79%) selected the correct avoidance response. A binomial test indicated that this was significantly different from chance level ($p < .01$). Note that all participants who chose the correct avoidance response on the first test trial also selected the correct avoidance response on the next three test trials. This observation suggests that the first choice was not based on chance, but rather increases confidence that the choice on the first test trials was well-reasoned. Three participants who did not select the appropriate avoidance response on the first test trial, did select the correct response on the next three test trials. The other two participants who selected the wrong avoidance response on the first test trial, also chose the wrong avoidance responses on the next three test trials.

From the participants who received A-R1R2 as the first test trial, 10 of 12 (83%) performed the correct avoidance response whereas nine out of 12 (75%) performed the correct avoidance response when B-R1R2 was presented as the first test trial. These proportions do not differ significantly ($p = .63$).

General Discussion

In all previous studies on avoidance behavior, the behavior could be based either on knowledge about simple S-S, R-S, S-R relations, or on hierarchical knowledge about the way in which the avoidance response modulates the Sd-US relation. Based on an adaptation of a design introduced by Colwill and Rescorla (1990), we created a situation in which the adaptive avoidance behavior could be based only on a hierarchical knowledge structure. The fact that participants did select the appropriate avoidance response above chance thus supports the idea that avoidance behavior can be based on a hierarchical knowledge structure. At present, this result is compatible only with the occasion setting account of avoidance learning (De Houwer et al., 2005) because this is the only account that postulates that avoidance behavior is based on a hierarchical knowledge structure.

There is, however, one alternative explanation of our results that does not involve a hierarchical knowledge structure. Participants might have formed a separate configurational representation for each specific combination of a warning signal and a response. Each of these configurational representations could have been associated with the representations of (the absence of) specific USs, allowing participants to anticipate the USs that occur when a specific avoidance response is emitted in the context of a specific warning signal. Note that configurational learning also provides a possible explanation for occasion setting and could thus also explain why avoidance behavior can have the functional properties of an occasion setter (see Declercq & De Houwer, in press). As Colwill and Rescorla (1990, p. 80) pointed out, it is very difficult to distinguish between an account in terms of configural learning and an account in terms of hierarchical knowledge. Nevertheless, we believe that it is unlikely that our results are attributable to configural learning.

Whereas one can imagine that the presentation of two stimuli can somehow be encoded as one unique configural stimulus that is more than the sum of the two stimuli, it seems unlikely that a warning stimulus and an avoidance response could be coded in terms of a single configural representation. An explanation of our results in terms of hierarchical knowledge thus seems to be the most plausible one.

Regardless of whether our results are due to hierarchical knowledge or configural learning, they are problematic for several models of avoidance behavior. First, our results cannot be explained by the two-factor theory of Mowrer (1947). The major problem for the two-factor theory concerns the revaluation of the US (for another demonstration of US revaluation in avoidance learning, see Declercq & De Houwer, 2008). US revaluation requires that there is a mental representation of the US. Yet, the two-factor theory assumes that the avoidance response is not performed to avoid the presentation of the US, but rather postulates that it is elicited automatically by the Sd as a result of reinforcement by the reduction of conditioned fear. Therefore, according to the two-factor theory, US revaluation should not have influenced the behavior of the participants. A second problem also relates to the fact that the US representation is not assumed to mediate avoidance behavior. From the perspective of the two-factor theory, only the aversiveness of the US matters (because it drives conditioned fear) but not the specific identity of the US. The results of our study show that avoidance behavior during the test phase depended on which US was most likely in which situation.

In other versions of the two-factor theory, the mental representation of the US does play a role. For example, the safety signal theory (Gray, 1987) presumes that stimuli that accompany the avoidance response become safety signals because they signal the absence of an expected US. These safety signals acquire Pavlovian conditioned inhibitory properties. The theory assumes that the avoidance response is elicited by the Sd and reinforced by these safety signals. Revaluation of the US can have an influence on
avoidance behavior only by changing the reinforcing value of the safety signals. However, because the avoidance response is elicited automatically by the Sd and safety signals are present only after the avoidance response, US revaluation can have an effect on avoidance behavior only after the first trial following US revaluation. The fact that we observed an effect of US revaluation on the first avoidance response after US revaluation is thus incompatible with the safety signal theory.

The theory of Seligman and Johnston (1973) is also not able to account for our results because it postulates that only R-S knowledge underlies avoidance behavior. In their theory, Seligman and Johnston (1973) argue that participants base their avoidance response on propositional knowledge about how a response affects the presentation of the US. In essence, participants are assumed to emit an avoidance response because they expect a US after not performing the response and expect the absence of the US after performing the avoidance response. Yet, performing the correct avoidance response in our study cannot be due only to knowledge about the R-US relation because the nature of the R-US relation depended on the presented Sd.

Another, recently developed theory that cannot explain our results is the expectancy-based theory of Lovibond (2006). This is also a two-factor theory in that it is assumed that both a Pavlovian and an instrumental component underlie avoidance behavior. Yet, Lovibond (2006) has a clear cognitive view on both of these components. First of all, Pavlovian learning is assumed to involve the construction of conscious expectancies about the presence of the US after the presentation of the Sd (S-S knowledge). Second, instrumental learning is regarded as the construction of expectancies about the absence of an expected US after performing an avoidance response (R-S knowledge). Lovibond (2006) assumes that avoidance behavior results from the combination of both pieces of knowledge. In the present study, however, no warning signal or avoidance response was uniquely correlated with (the absence of) a US. Rather, which US was presented depended on the specific combination of a warning signal and an avoidance response. It was therefore impossible to select the appropriate avoidance response during test merely on the basis of knowledge about S-S and R-S relations.

The model of Lovibond (2006) can explain our results if it is extended with the assumption that the R-S knowledge is context specific. That is, the Sd can be seen as an occasion setting stimulus that signals when a particular R will have a particular result (i.e., absence of US1 or US2). Consider the case in which US1 is absent when R1 is performed after Sd A and US2 is absent when R1 is performed after Sd B. Participants would then learn that R1 prevents US1 in the context of Sd A and that R1 prevents US2 in the context of Sd B. When Sd A is presented during test, they will expect that both US1 and US2 will be present and that they can prevent US1 by performing R1. They will use this knowledge to avoid the presence of the negative US. Like the occasion setting account, avoidance behavior is thus assumed to be based on a hierarchical knowledge structure. The only difference is that in the modified account of Lovibond, the Sds are seen as the modulators of the R-S relations whereas in the occasion setting account, the Rs are seen as the modulators of the S-S relations.

In this context, it is important to note that the occasion setting account as formulated by De Houwer et al. (2005) does not include assumptions about how hierarchical knowledge is represented in memory (also see Declercq & De Houwer, in press). One possibility is that the representation of the avoidance response is linked to the association between the CS and US representations rather than to the CS or US representations itself (see Holland, 1983, and Ross & Holland, 1981, for a similar account of occasion setting). Another possibility is that hierarchical knowledge is represented in propositional format. This implies that during avoidance learning, participants construct two propositional beliefs that they judge to be true: (1) When the avoidance response is not emitted, the warning signal will be followed by the US; and (2) When the avoidance response is emitted, the warning signal will not be followed by the US. This knowledge is then used to avoid the delivery of the negative US. From this point of view, the occasion setting model differs from the model of Lovibond only with regard to the content of the propositions that are thought to drive avoidance behavior.

One way to distinguish between the adaptation of the Lovibond model (2006) and the occasion setting account of De Houwer et al. (2005) is to examine whether the Sds or the Rs have the properties of an occasion setter. Previous studies have shown that avoidance Rs can have the properties of an occasion setter (Declercq & De Houwer, in press; De Houwer et al., 2005). It remains to be seen, however, whether Sds or Rs function as an occasion setter when, as in the present design, avoidance behavior can be based only on hierarchical knowledge.

The reported results are important not only theoretically but they also have some important clinical implications. Avoidance behavior is present in many forms of psychopathology. If a hierarchical structure underlies avoidance behavior, then therapy should focus not on changing simple associations between two elements but instead, all three elements of avoidance learning should be incorporated in therapy.

One possible criticism of this study might be that loss of a small amount of money was not a very aversive stimulus. The reason for using it was that it was easy to revalue. It is possible, however, that with more fear-evoking stimuli, other processes might be involved in avoidance learning. Although it is for ethical reasons impossible to use really fear evoking stimuli in the lab, it is clear that further research on avoidance learning is essential.

References

(Eds.), Quantitative analyses of behavior: Discrimination processes (pp. 183–206). New York: Ballinger.


