EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2002, *14* (1), 127–153

Intentional fixation of behavioural learning, or how R–O learning blocks S–R learning

Armin Stock and Joachim Hoffmann

Institute for Psychology, University of Würzburg, Germany

Four experiments deal with the acquisition of knowledge for the control of voluntary behaviour. Subjects had to accomplish a computer-controlled learning task that required them to learn which one of four actions (R) had to be performed in order to attain a certain one of four outcomes (O) in the presence of one of four situational contexts (S). Reinforcements were assigned to actions according to two schedules: In the S–R condition each of the four actions was consistently reinforced in the presence of a certain situation, whereas in the R–O condition each of the actions was either always (Experiments 1–3) or mostly (Experiment 4) reinforced if a certain outcome was required. Furthermore, the type of feedback was varied. In Experiment 1, only positive or negative reinforcements were fed back, whereas in Experiments 2–4 the actions resulted in outcomes that had to be compared with the required outcomes in order to determine successes and failures. The results indicate a preference for learning R–O contingencies over learning S–R contingencies. Most subjects were so fixated on learning R–O relations that they remained completely blind to the consistent reinforcement of S–R mappings. The data suggest that, in line with the ideomotor principle, the acquisition of behavioural competence is based primarily on the formation of bidirectional action–outcome relations. Specifications of the underlying learning mechanisms are discussed.

This paper deals with a simple and ubiquitous learning task, namely to learn what to do in order to attain a certain goal. Humans accomplish this learning task rather effortlessly under everyday circumstances. People easily learn for example what to do in order to switch on a radio, to open a bottle, to enter a multi-storey car park and so on, for millions of simple goals. Moreover, folk psychology gives an obvious account on how people acquire such behavioural

Requests for reprints should be addressed to A. Stock, Institute for Psychology, University of Würzburg, Röntgenring 11, D-97070 Würzburg, Germany.

Email: stock@psychologie.uni-wuerzburg.de

We would like to thank Axel Cleeremans, Luis Jiménez, and Peter A. Frensch for their critical and stimulating comments on earlier versions of this paper and we would also like to thank the psychology students who served as subjects, and Julie Barthel and Christian Stöcker for checking and improving the English.

http://www.tandf.co.uk/journals/pp/09541446.html DOI:10.1080/09541440042000340

competence: If one notices that a certain action results in a certain outcome, that action is stored as instrumental for producing the outcome. Then, if later the same outcome is desired, that action that has been experienced to produce it is addressed and performed. For example, if one wants to listen to the radio, one performs the action that has been experienced to switch it on.

The notion that the acquisition of behavioural competence is based on the formation of bidirectional relations between actions and their consistent outcomes is the essence of the ideomotor hypothesis. A precise formulation of the hypothesis was given as early as 1825 by J.F. Herbart (p. 464, our translation):

> Right after the birth of a human being or an animal, certain movements in the joints develop, for merely organic reasons and independently of the soul and each of these movements elicits a certain feeling in the soul. In the same instance, the outside senses perceive what change has come about.... At a later time, a desire for the change observed earlier arises. Thus, the feeling associated with the observation reproduces itself. This feeling is a self-preservation (''Selbsterhaltung'') of the soul, which corresponds to all the inner and outer states in nerves and muscles through which the intended change in the sphere of sensual perception can be brought about. Hence, what has been desired actually happens; and the success is perceived. Through this, the association is reinforced: the action, once performed, makes the following one easier and so on.

Although the ideomotor hypothesis was widely acknowledged at the end of the 19th century (Harleß, 1861; James, 1890/1981; Lotze, 1852; Münsterberg, 1889; Wundt, 1893; cf., Hoffmann, 1993; Hommel, 1998; Prinz, 1997, 1998 for reviews of these early accounts), it did not receive any serious consideration in the early research on animal learning. For the dominating behaviourism at that time, the notion that instrumental behaviour might be determined by mental states like a ''desire for a change with associated feelings'' was in no way scientifically respectable (Thorndike, 1913; Watson, 1930; cf., Greenwald, 1970). Instrumental behaviour was attributed to stimulus–response (S–R) instead of response–outcome (R–O) associations, whereby, according to Thorndike's ''Law of Effect'', S–R associations were assumed to result from a reinforcement of the response in the presence of the stimulus. Dickinson (1994, p. 48) recently correctly noted that ''the most perverse feature of such stimulus–response theories has always been the claim that knowledge about the instrumental contingency between action and outcome plays no role in the performance of the action''. According to S–R theorists, instrumental behaviour is a habitual reflex to the training situation instead of a purposefully willed action. Although the formation of habits surely plays some role in instrumental learning, it is hard to believe that the behaviour of animals and more so that of humans is not also determined by the intended goals.

It has been the merit of Tolman (1932, 1949) to have called the attention of learning research to the purposive character of instrumental behaviour. Tolman argued that animals not only acquire knowledge about reinforcements of actions in situations but also about their outcomes, which he identified as ''means–end readiness''. If, according to his reasoning, an outcome is expected that currently is of some value for the animal, the action is addressed that has been stored as being a means to the end of the ''desired'' outcome. The obvious contradiction of this approach to the traditional S–R theories provoked a more thorough analysis of the extent to which the various associations between the situation (S), the response (R), and the outcome (O) contribute to instrumental behaviour.

If animals are trained with only one reinforcer (O), it is difficult to distinguish the contribution of S–R and R–O associations on the resulting instrumental behaviour. If, for example, a rat is trained to press a lever for food, it is equally possible that lever pressing is triggered by the situation as by a desire for the food pellets. Thus, in order to disentangle the influence of S–R and R–O associations, multiple R–O contingencies have been trained and subsequently the value of a certain outcome has been changed. If the initial training established only S–R habits, the propensity to perform the formerly reinforced responses should not be changed by the manipulation of outcome values. By contrast, if training has led also to the formation of R–O associations, the propensity to perform the different responses should be altered in correspondence to the changed value of their outcomes.

Consider, for example, a study by Colwill and Rescorla (1985): Rats were first contingently reinforced with food pellets after performing R_1 and with a sucrose solution after R_2. Once instrumental training had occurred, one of the two reinforcers (outcomes) was devalued by associating it with a mild nausea. Finally, the rats were again given a choice between the two responses, but with all outcomes omitted. In this test-phase the rats showed a clear suppression of performing the response the outcome of which had been devalued. Obviously, the rats had not only associated two responses with a situation wherein these were reinforced (S–R_1 and S–R_2), but they had also learned which response leads to which outcome (R_1–food pellets, R_2–sucrose solution), so that they were able to selectively avoid performing the response of that outcome which had been devalued.

There are numerous studies which have used the devaluation technique under yet more sophisticated settings (see Dickinson, 1994; Pearce, 1997; Rescorla, 1990, 1991, 1995 for recent reviews). Although there is evidence for the contribution of S–R and S–O associations on the resulting behaviour, the evidence for the formation of response–outcome associations is by far the strongest. Thus, it seems appropriate to conclude that associations between responses and their contingent outcomes are a central part of the representations that are formed by animals in instrumental learning.

Since the late 1970s response–outcome learning has also been established as an experimental paradigm in human learning (Allan, 1980; Allan & Jenkins, 1980; Alloy & Abramson, 1979; Dickinson, Shanks, & Evenden, 1984; Shanks, 1985; Shanks & Dickinson, 1987; Wasserman, Chatlosh, & Neunaber, 1983; Wasserman, Elek, Chatlosh, & Baker, 1993). Typically, in these experiments subjects were presented with a number of trials on each of which they could perform or not perform a particular response (e.g., a key-press) and subsequently observe whether or not a certain event (e.g., a light) appeared as an outcome. The critical variation concerns the contingency between the response and the outcome, i.e., the probability $p(\mathrm{O/R})$ of an outcome (O) given the performance of the response (R), and the probability $p(\mathrm{O}/\neg\mathrm{R})$ of an outcome given that the response is not performed. The task of the subjects is to judge the effectiveness of the action in controlling the outcome. Although judgements have been shown to be biased by several factors, they are generally found to be highly correlated with the difference between both probabilities $\Delta p = p(\mathrm{O/R}) - p(\mathrm{O}/\neg\mathrm{R})$, i.e. with one of the normative measures of the contingency between two binary variables.

Although this finding clearly indicates that humans are sensitive to response–outcome contingencies, it hardly says much about the contribution of R–O learning to instrumental behaviour, as the experimental settings differ too much from the conditions of ordinary goal-oriented behaviour. First, subjects continuously watch the contingency between only one response and one outcome, whereas humans usually aim for steadily changing goals facing multiple R–O contingencies. Second and more important, subjects are explicitly instructed to judge the strength of the experienced R–O contingencies, and it is by no means surprising that they adopt appropriate strategies in order to do so. In contrast, in ordinary behaviour settings the task is not to judge R–O contingencies but to learn what to do in order to attain various goals. Thus, it is more than questionable whether subjects incidentally learn in the same way about multiple R–O contingencies as they explicitly learn if required to judge a single R–O contingency.

The present experiments were designed in order to examine the status of R–O learning in humans under experimental settings which at least resemble the circumstances of ordinary voluntary behaviour. On individual trials subjects were presented with one of several initial conditions and they were asked to perform one of several actions, resulting in the occurrence of one of several outcomes. Moreover, in each individual trial one of the possible outcomes was defined as the goal currently to be reached (i.e., the other outcomes were "devalued"). We were interested to examine what knowledge subjects abstract from their behavioural experiences in order to learn how to behave appropriately. Moreover, we were especially interested to disentangle the contribution of S–R and R–O associations on the resulting representations that control performance.

GENERAL METHOD

Subjects had to accomplish a computer-controlled learning task that required them to learn which action achieves which goal under which initial condition. There were four initial conditions, illustrated by displaying one of four start symbols in the lower left-hand corner of the screen (see Figure 1). There were also four different goals, illustrated by displaying one of four goal symbols in the lower right-hand corner of the screen. Finally, four different actions were available—the pressing of one of four keys on the computer keyboard. Subjects could choose freely among the four keys in each trial and they were told that they should try to select respectively that key that would attain the given goal in the presence of the given start symbol.

The experimental variations first concerned the assignment of reinforcements to the four actions in each of the 16 combinations of start and goal symbols. In Experiments 1 and 2, two reinforcement schedules were applied: In a so-called S–R condition, each of the four keys was consistently reinforced if selected in the presence of a certain *start symbol*. In a so-called R–O condition, each of the four keys was consistently reinforced if selected in the presence of a certain *goal symbol*. Both conditions require learning of a simple mapping of the keys to either the start or the goal symbols in order to always act correctly. In Experiment 3, both schedules were applied again but with the overall R–O contingencies changed. Finally, in Experiment 4 a ''mixed condition'' was applied, comprising both consistent reinforcements of S–R mappings and only frequent reinforcements of R–O mappings.

The second variation concerned the way in which reinforcements were fed back. In Experiment 1, subjects received feedback only on whether the selected key was the correct or an incorrect one. In contrast, in Experiments 2–4, key-presses resulted in the presentation of another symbol on the screen, which was to be compared with the current goal symbol. A key-press was reinforced if it resulted into the presentation of the goal symbol again, i.e., if the outcome

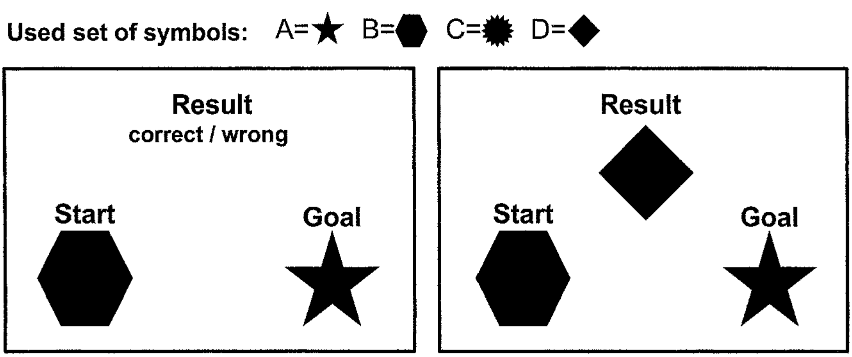

Figure 1. The left-hand side shows a typical situation from Experiment 1 after selecting a response key. Subjects were informed only about whether the selected key was correct or not. On the right-hand side an illustration of the same situation from Experiment 2 is given. Here, the subjects experienced the concrete effects of their actions.

symbol was identical to the goal symbol. In all experiments the mean percentage of correct keys over the learning course and the mean number of trials needed to reach a criterion of 16 consecutive correct key-presses were measured as the dependent variables.

EXPERIMENT 1

In each trial, subjects were presented with a certain start symbol and a certain goal symbol. They were required to press one of four keys in order to attain the current goal. Feedback was provided about whether the current key-press was correct or not. There were two reinforcement schedules: In the S–R condition, each key was consistently reinforced in the presence of a certain start symbol; in the R–O condition, each key was consistently reinforced in the presence of a certain goal symbol. Thus, the ''solution'' was in both cases a mapping of the four keys to either the four start or the four goal symbols.

Method

Subjects. Thirty students from an introductory psychology course served as subjects. The S–R and R–O conditions were randomly assigned to two subgroups of 15 subjects each.

Stimuli and procedure. Four different symbols (A, B, C, and D) were used as start as well as goal symbols (see Figure 1). In consecutive runs of 16 trials each, the 16 possible start–goal combinations were presented in random order. The start symbol was presented first, and after a delay of 1000 ms the goal symbol was presented. Subjects were required to press one of four keys on a computer keyboard (keys ''d'', ''f'', ''j'', and ''k'') in order to attain the given goal. Choosing the correct key was reinforced by the word ''correct'' displayed in the upper middle of the screen. Additionally, an ascending melody was played. If a wrong key was pressed, the word ''wrong'' was displayed together with a descending melody. The next trial was to be initiated by pressing the space key. In the S–R condition, the keys ''d'', ''f'', ''j'', and ''k'' were correct if the symbols A, B, C, and D, respectively, were presented as the *start symbol*. In the R–O condition the keys ''d'', ''f'', ''j'', and ''k'' were correct if the symbols A, B, C, and D, respectively, were presented as the *goal symbol*. Subjects received no hint at any kind of systematicity, they were merely instructed to select respectively those keys that would allow them to attain the current goal, and that the experiment would end after 256 trials (i.e., 16 runs of 16 individual trials) or if they accomplished 16 correct key-presses in a row. Following the learning phase, subjects were asked about any strategy they might have used and whether they could indicate any rule for the selection of the right key.

Results

All subjects, except for one in each group, accomplished the learning task within the limit of 256 trials. The mean number of trials to criterion amounted to 86.8 in the S–R group with a standard error (*SE*) of 18.3 and to 101.6 in the R–O group (*SE* 19.50).[1] The difference is not significant [$t(28) = .56$, $p < .583$]. Figure 2 shows the percentage of correct trials plotted against blocks of four trials each, whereby subjects who had fulfilled the criterion were assumed to be always correct after that. The data show a slightly steeper learning curve for the S–R group, but the difference is not significant [$F(1, 28) < 1$]. Only one subject in the S–R group could not indicate the correct rule after the learning session, and it was due to this subject that the S–R group did not reach 100% correct trials.

Discussion

Subjects quickly learned that each of the four key-presses was successful only if a certain start symbol (S–R condition) or a certain goal symbol (R–O condition) was present. This finding is consistent with the key-presses being determined by S–R associations as well as by conditionalised R–O associations. From the perspective of S–R learning, the key-presses became associated with those symbols that were present in correct trials so that either the start or the goal

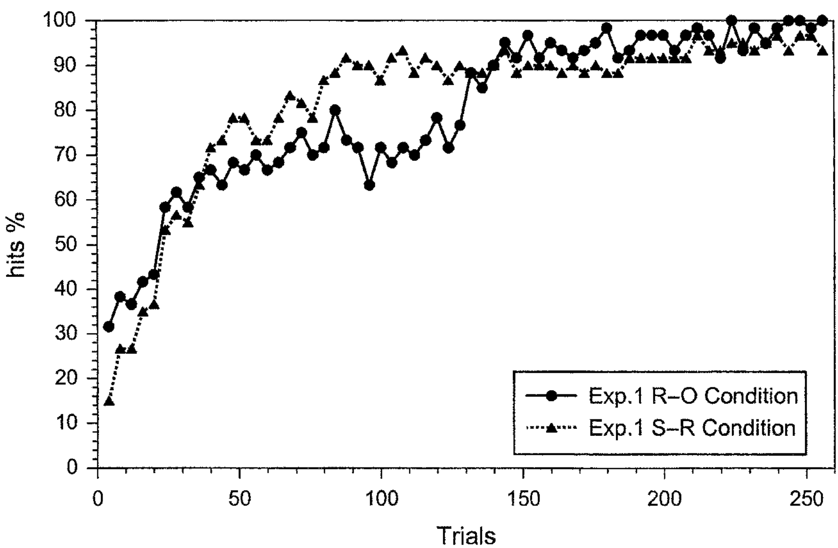

Figure 2. Learning curves for the S–R condition (triangles, dotted line) and for the R–O condition (circles, solid line) plotted against blocks of four trials each, Experiment 1.

[1] Subjects who did not reach the criterion are assigned the total number of trials (256).

symbols increasingly triggered those keys that had always been reinforced in their presence. In terms of R–O learning, however, conditionalised ''key–success'' relations may have been learned: Subjects experienced that all keys sometimes led to the desired outcome, i.e., to positive feedback, and sometimes not, so that they tried to work out under which ''conditions'' each key produced the desired outcome. The data reveal that they can do so equally successfully regardless of whether the critical conditions refer to the start or the goal symbols.

Remember that subjects were instructed to ''attain'' the current goal, which explicitly directed attention towards the goal symbols. Thus, it was to be expected that subjects would learn somewhat faster in the R–O condition than in the S–R condition. However, in the R–O condition subjects learned even more slowly than in the S–R condition. The attentional bias for the goal symbols, set by the instruction, was presumably equalised by presenting the start symbols first and by a tendency to encode the symbols from left to right. Subjects might also have replaced the instructed intention to attain the current goal, which mostly changed from trial to trial, by the constant and only feasible intention for receiving positive feedback, so that both symbols were likewise considered as being initial conditions, despite their different designation as start and goal symbols. In any case, the data of Experiment 1 indicate no difference between learning under the S–R and the R–O condition, so that they provide no hint for a different impact of S–R and R–O associations on performance.

EXPERIMENT 2

The settings of Experiment 1 were purposely chosen so that subjects had only the kind of information available that is, according to the law of effect, exclusively used in instrumental learning, namely whether an action is or is not reinforced. In daily behaviour, however, actions always lead to concrete outcomes, additional to reinforcement information. More precisely, it is the correspondence of actual to desired outcome that makes an action successful. Experiment 2 was designed in order to approach these more ''true to life'' conditions. Each key-press now resulted in the presentation of one of the four symbols on the screen just above and in the middle of the start and the goal symbol (see Figure 1). Subjects were again required to aim in each trial for the present goal. They were also told that a key-press would successfully attain the goal if its outcome symbol was identical to the given goal symbol. Thus, like in daily behaviour, it is the correspondence between the actual and the desired outcome of the performed action that makes it successful.

In order to meet the reinforcement schedule of the S–R condition of Experiment 1, the four symbols were assigned to the outcomes of the four key-presses as shown in Table 1: For each start symbol there was a certain key that always resulted into the presentation of the present goal symbol again, regardless of which goal symbol it was. Moreover, the same key never resulted into the presentation of the required goal symbol if another start symbol was present. For example, if the start symbol A was present, key ''d'' led to symbol A if A was the goal symbol; it led to symbol B if B was the goal symbol, etc. However, if another start symbol was present, pressing key ''d'' was never correct. For example, if the start symbol B was present and the goal symbol A was required, key ''d'' led to symbol D; if B was required symbol C resulted, etc. (cf., Table 1). Thus, as in the S–R condition of Experiment 1, each key was consistently reinforced in the presence of a certain start symbol.

Table 2 shows the assignment of the four symbols to the outcomes of the four keys in order to meet the R–O condition of Experiment 1: for each goal symbol there was now a certain key that always resulted into the presentation of the same symbol, regardless which start symbol was currently present. For example, the key ''d'' and only this key was always successful if symbol A was the goal; the key ''f'' and only this key was always successful if symbol B was the goal, etc. (cf., Table 2). Again, as in the R–O condition of Experiment 1, each key was consistently reinforced in the presence of a certain goal symbol.

A comparison of Tables 1 and 2 reveals that despite the fact that reinforcements (the grey cells) are likewise consistently assigned onto either S–R or R–O

TABLE 1
Structure of the reinforcement schedule for the S–R condition of Experiment 2

		Result/symbol + ascending or descending melody			
Start	*Goal*	*Key ''d''*	*Key ''f''*	*Key ''j''*	*Key ''k''*
A	A	A	D	D	D
	B	B	C	C	C
	C	C	A	A	A
	D	D	B	B	B
B	A	D	A	D	D
	B	C	B	C	C
	C	A	C	A	A
	D	B	D	B	B
C	A	D	D	A	D
	B	C	C	B	C
	C	A	A	C	A
	D	B	B	D	B
D	A	D	D	D	A
	B	C	C	C	B
	C	A	A	A	C
	D	B	B	B	D

TABLE 2
Structure of the reinforcement schedule for the R–O condition of Experiment 2

		Result/symbol + ascending or descending melody			
Start	*Goal*	*Key "d"*	*Key "f"*	*Key "j"*	*Key "k"*
A	A	A	B	C	D
B		A	B	C	D
C		A	B	C	D
D		A	B	C	D
A	B	A	B	C	D
B		A	B	C	D
C		A	B	C	D
D		A	B	C	D
A	C	A	B	C	D
B		A	B	C	D
C		A	B	C	D
D		A	B	C	D
A	D	A	B	C	D
B		A	B	C	D
C		A	B	C	D
D		A	B	C	D

mappings, the two schedules differ with regard to the action–outcome relations. In the S–R condition (Table 1) each key leads to the presentation of every possible symbol. In contrast, in the R–O condition (Table 2), each key leads always to the same symbol. We were interested to see whether or not this difference in the contingencies of the experienced R–O relations influences learning in comparison to Experiment 1. If learning is based exclusively on reinforcement information, the same results are to be expected because reinforcements are scheduled in the same way as in Experiment 1. If, however, learning takes R–O relations into account, it is to be expected that contingent R–O relations will support learning and non-contingent R–O relations will impair it.

Method

Subjects and design. Thirty undergraduate psychology students participated in this experiment; none of them had participated in the first experiment. They were randomly assigned to two groups of 15 subjects each. Subjects from one group were exposed to the S–R condition (see Table 1); subjects from the other group were exposed to the R–O condition (see Table 2).

Procedure. The procedure was almost the same as in Experiment 1, except that after each key-press another symbol was presented additionally to the ascending or descending melody, which again indicated reinforcement or punishment of the current key-press.

Results

All subjects from the R–O group managed the learning task within the limit of 256 trials and indicated the correct rule after the learning session. The mean number of trials to criterion amounted to 62.5 (*SE* 6.15). In the S–R group only 3 of the 15 subjects accomplished the learning task within the limit of 256 trials and these subjects also indicated the correct rule after learning. The remaining 12 subjects neither reached the criterion nor did they report any systematic strategy. Instead, they reported that they despaired more and more and some of them were convinced that there was no solution at all and that they had been fooled by the experimenter. If one assigns the maximal number of trials (256) to subjects who did not reach the criterion (as in Experiment 1), the mean number of trials to criterion amounted to 243.5 (*SE* 8.45). The difference between both groups is significant [$t(28) = -17.3$, $p < .0001$].

Figure 3 shows the percentage of correct trials plotted against blocks of four trials separately for both groups. Subjects who reached the criterion were assumed to always be correct after that. The data confirm that subjects in the

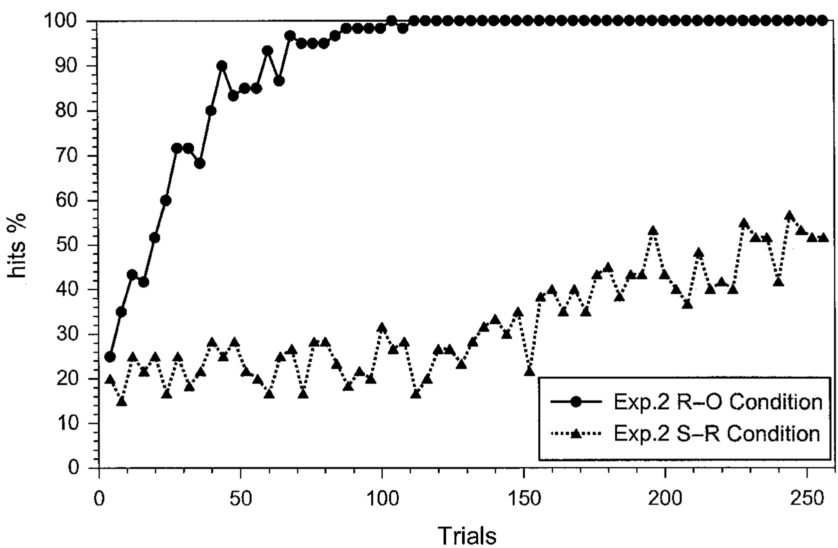

Figure 3. Learning curves for the S–R condition (triangles, dotted line) and for the R–O condition (circles, solid line) plotted against blocks of four trials each, Experiment 2.

R–O group quickly learned which key was to be pressed in order to attain the current goal. In contrast, in the S–R group the random rate of 25% trials correct is exceeded late in the course of learning, due to the three subjects who learned the S–R mappings. The two learning curves differ significantly from each other [$F(1,28) = 321.27$, $p < .0001$].

In order to compare the data of Experiment 1 and 2 an ANOVA with the variables Experiment (1 and 2) and condition (S–R and R–O) as between-subject factors was calculated. In addition to significant main effects of both factors [Experiment: $F(1,56) = 21.023$, $p < .001$; Condition: $F(1,56) = 46.886$, $p < .0001$], there is a significant interaction [$F(1,56) = 54.681$, $p < .0001$]. Whereas learning under the R–O condition was accelerated in Experiment 2 [$t(16.76) = -1.91$, $p < .073$], learning under the S–R condition was dramatically impaired [$t(28) = 7.76$, $p < .0001$].

Discussion

The results of Experiment 2 clearly show that experiencing concrete outcomes of the key-presses differentially influenced learning under the reinforcement schedules of the S–R and the R–O condition. In comparison to Experiment 1, where only successes and failures were fed back, learning in the R–O condition was somewhat facilitated, whereas learning in the S–R condition was substantially impaired. Most subjects in the S–R group remained completely blind to the consistent reinforcement of the S–R mappings. Moreover, subjects were very surprised when we uncovered the simplicity of the rule after the experiment, and they could not believe that they had not detected it, although many of them claimed that they had searched for a dependency between the success of the keys and the start symbol.

The amazing differences between the reinforcement schedules indicate that experiencing concrete outcomes has a strong impact on learning. As subjects now are required to compare the respective goal symbol with the outcome symbol in order to check the success of the key, they necessarily attend the outcomes of their key-presses. Thus, it is not astonishing that the R–O rule is easily learned: Subjects experience that each of the four keys always results into the presentation of a certain one of the four symbols, which quickly leads to the formation of four contingent R–O relations. If a certain goal symbol is then presented, that key is pressed that has been experienced to result into the presentation of the desired symbol.

It is, however, striking that the equally consistent reinforcement of S–R mappings, which were easily learned in Experiment 1, are now almost impossible to learn. Only three of the fifteen subjects of the S–R group realised that each key was consistently reinforced under the presence of a certain start symbol. All other subjects, even after 256 individual trials, remained blind to the consistent reinforcement of the four S–R mappings. If subjects had monitored the reinforcements irrespective of the concrete goal–outcome relations they were based on, the S–R rule would have been as easily learned as in Experiment 1. However, as our data suggest, it seems to be exactly this abstraction from the goal–outcome relations that is almost impossible to accomplish.

Figure 4 illustrates a tentative learning schema that accounts for the efficiency of reinforcements in the R–O condition as well as for the inefficiency of reinforcements in the S–R condition (cf., Hoffmann, 1993; Hoffmann & Sebald, 2000; Hoffmann & Stock, 2000). We assume that if the reinforcement of an action depends on the correspondence of its outcome to what the actor strives for, i.e., to the desired outcome, correspondences primarily lead to a strengthening of associations between the performed action and the desired outcome that has been reached. In other words, associations between actions and *reinforced* outcomes are formed, so that the actions become represented as being instrumental to attain certain desired outcomes. As reinforcements trigger the formation of R–O associations, simultaneous formation of the likewise reinforced S–R relations is blocked. Consequently, in the present experiment, subjects easily form the consistently reinforced R–O associations but they have no chance to form the likewise consistently reinforced S–R associations.

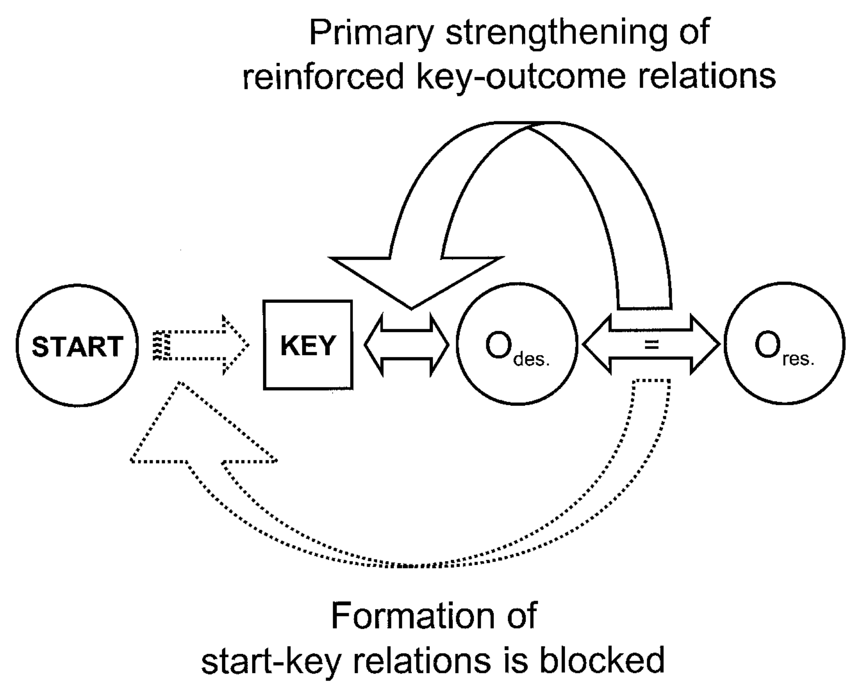

Figure 4. Tentative learning schema to account for the results of Experiment 2. O_{des} = desired outcome = goal symbol; O_{res} = resulting outcome of the current key-press.

Furthermore, we assume that the situational context is taken into account only secondarily if R–O relations are not consistently reinforced. So, if a certain action sometimes allows attainment of a certain goal and sometimes not, the actor starts to search for critical features in the situational context that might be responsible for the success or the failure of the corresponding action. If such features are found, the R–O relations are conditionalised, i.e., S–(R–O) triplets are formed (cf., Rescorla, 1990, 1991 for a similar argument in animal learning). This is what presumably happens if subjects are exposed to the S–R condition. They experience for each key that it sometimes results in the presentation of a certain goal symbol and sometimes not, so that they try to find out, inasfar it depends on the present start symbol, whether or not a certain key successfully attains a certain goal. If subjects had to manage the S–R condition by such conditionalised R–O relations, at least 16 S–(R–O) triplets would have to be formed. This is a task far too complex to be solved in such a limited number of trials here.

It can be suspected that the difficulty of the S–R rule might be more parsimoniously accounted for by arguing that subjects simply did not take notice of the start symbols because they were required to compare the outcomes of their key-presses with the present goal symbols. Although this consideration is in contrast to what subjects reported, we nevertheless looked for a possibility to compensate for this *attentional* bias in order to rule out a pure attentional account. In Experiment 3 the S–R condition was modified in such a way that key-presses frequently led to the presentation of the current *start* symbol. The frequent identity of the start symbols to the outcome symbols should attract attention so that it should be more likely for subjects to take notice of the start symbols. Consequently, if only attention matters, learning in the S–R condition should be likewise easy as learning in the R–O condition.

Another suspicion regards differences in the information subjects can gain from errors. As under the R–O condition of Experiment 2 each key always results in the presentation of the same symbol (cf., Table 2), subjects could learn about the critical R–O relations not only by hits but also by failures. For example, if key ‘‘f’’ has produced symbol B when goal symbol C was required, despite the error subjects gain the correct information that key ‘‘f’’ is appropriate to produce B. In contrast, under the S–R conditions, subjects could not gain any useful information about the critical S–R relations from errors. For example, if key ‘‘f’’ failed in the presence of the start symbol C, no information at all is provided about another start symbol in the presence of which the key ‘‘f’’ might be appropriate. Thus, it can be suspected that the R–O relations are easy to learn because they are consistently confirmed in hits *and* errors, whereas the S–R relations are hard to learn because they are only confirmed in hits. In order to control for this difference, in Experiment 3 the R–O condition is modified so that the critical R–O relations are no longer confirmed in errors.

EXPERIMENT 3

In Experiment 3 subjects again were exposed to either the S–R or the R–O condition. In contrast to Experiment 2, however, the S–R condition was realised in such a way that incorrect key-presses, whenever possible, resulted in the presentation of the current *start* symbol (cf., Table 3). As a consequence, in 40 out of the total of 64 start-goal-key combinations the start symbol was presented as the outcome of the performed key-press. As identical symbols should attract attention, the start symbols should be more likely to be considered as being critical for the success of the key-presses.

The R–O condition was also modified in comparison to Experiment 2: In those cases in which the key-presses were incorrect, they no longer resulted in the presentation of always the same symbol but rather in the presentation of the remaining three symbols with equal frequency. Consequently, the overall R–O relations were as inconsistent as in the S–R rule, i.e., each key resulted in the presentation of every possible symbol (cf., Table 4). Another consequence of this manipulation is that subjects now could no longer learn about the critical R–O relations from errors. For example, if key ''f'' led to symbol D when the goal symbol C was required, this error now did by no means indicate that key ''f'' is appropriate to produce D, as in order to produce D, key ''k'' was to be

TABLE 3
Structure of the reinforcement schedule for the S–R condition of Experiment 3

		Result/symbol + ascending or descending melody			
Start	*Goal*	*Key ''d''*	*Key ''f''*	*Key ''j''*	*Key ''k''*
A	A	A	B	C	D
	B	B	A	A	A
	C	C	A	A	A
	D	D	A	A	A
B	A	B	A	B	B
	B	A	B	C	D
	C	B	C	B	B
	D	B	D	B	B
C	A	C	C	A	C
	B	C	C	B	C
	C	A	B	C	D
	D	C	C	D	C
D	A	D	D	D	A
	B	D	D	D	B
	C	D	D	D	C
	D	A	B	C	D

TABLE 4
Structure of the reinforcement schedule for the R–O condition of Experiment 3

		Result/symbol + ascending or descending melody			
Start	*Goal*	*Key "d"*	*Key "f"*	*Key "j"*	*Key "k"*
A	A	A	C	D	B
B		A	C	D	B
C		A	C	D	B
D		A	C	D	B
A	B	D	B	A	C
B		D	B	A	C
C		D	B	A	C
D		D	B	A	C
A	C	B	D	C	A
B		B	D	C	A
C		B	D	C	A
D		B	D	C	A
A	D	C	A	B	D
B		C	A	B	D
C		C	A	B	D
D		C	A	B	D

pressed. Thus, errors do not only provide no information (as in the S–R rule), they rather provide misleading information about the R–O relations to be learned. Consequently, learning under the modified R–O condition should be substantially impaired if defectively produced outcomes are also taken into account in the formation of R–O associations.

Note, that according to the tentative learning schema of Figure 4, learning is primarily driven by reinforced key-outcome relations, i.e., the keys are above all associated to desired outcomes (goal symbols) that have been attained. With regard to these reinforced key-outcome relations, however, the modified R–O condition of Experiment 3 is identical to the R–O condition of Experiment 2: It is again the case that if subjects aim for a certain goal symbol, there is a certain key that consistently allows them to do so, i.e., the corresponding R–O relation is consistently reinforced. Thus, if learning indeed primarily depends on the reinforcements of these goal-related, or so to speak ''intentionally nested'', R–O contingencies, the abolition of the overall R–O contingencies in Experiment 3 should not matter much.

Altogether, Experiment 3 compares learning under the S–R and the R–O conditions under more equalised conditions than Experiment 2. First, the overall R–O relations in the R–O condition are equally inconsistent, as in the S–R condition. Second, in both conditions errors provide either no or misleading information about the S–R or R–O relations to be learned. Finally, in the S–R condition the frequent presentation of outcomes that are identical to the start symbol should compensate for the attentional bias in favour of the goal symbols.

Method

Subjects. Thirty students from an introductory psychology course served as subjects. None of them had taken part in any of the earlier studies. The S–R and the R–O conditions were randomly assigned to two sub-groups of 15 subjects each.

Procedure. The experimental procedure was identical to that of Experiment 2.

Results

In the S–R group, six of the fifteen subjects accomplished the learning task within the limit of 256 trials and these subjects also indicated the correct rule. The remaining nine subjects found the task as difficult as the majority of subjects who dealt with the S–R condition in Experiment 2 had done. If one assigns the maximal number of trials (256) to subjects who did not reach the criterion (as in the previous experiments), on average 216.7 (*SE* 15.22) trials to criterion were taken, which does not significantly differ from the mean number of trials to criterion in Experiment 2 (243.5, *SE* 8.45), [$t(21.89) = 1.53$, $p = .139$]. Likewise, the learning curves do not differ significantly in both Experiments, [$F(1, 28) = 1.805$, $p = .190$].

In the R–O group all subjects except one reached the criterion and indicated the correct rule. The mean number of trials to criterion increased from 62.5 in Experiment 2 to 120.2 (*SE* 18.75) in Experiment 3. This difference is significant, [$t(16.98) = -2.92$, $p < .010$], and also the learning curves differ significantly between both experiments, [$F(1, 28) = 7.345$, $p < .011$].

Although in comparison to Experiment 2, the S–R condition was somewhat easier and the R–O condition was somewhat more difficult to accomplish, there is still a substantial and significant advantage of learning under the R–O condition, [$t(19.47) = 5.99$, $p < .0001$, for the mean number of trials to criterion, and $F(1, 28) = 47.838$, $p < .0001$ for the learning curves, cf., Figure 5].

Discussion

The experimental manipulations of Experiment 3 were both effective: The fact that under the S–R condition the outcomes of the key-presses were almost always identical to the start symbol did indeed attract attention. Most subjects spontaneously reported that they had noticed this, so that the start symbols were

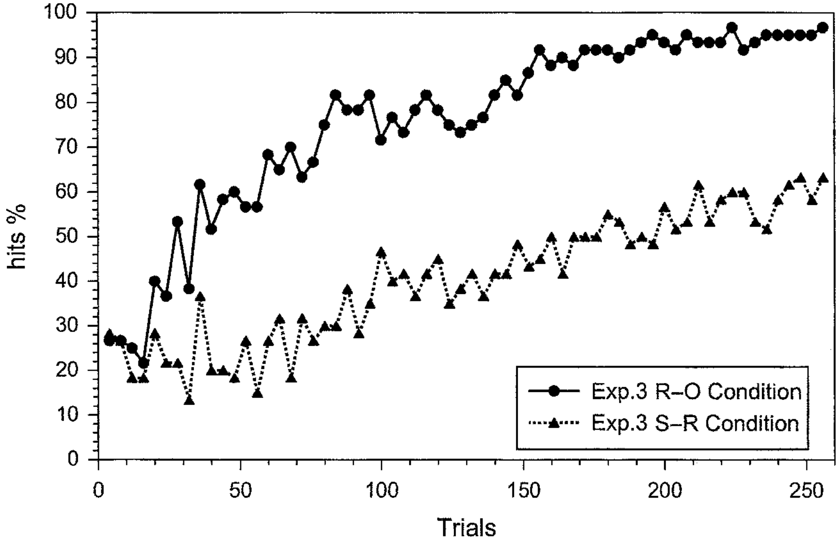

Figure 5. Learning curves for the S–R condition (triangles, dotted line) and for the R–O condition (circles, solid line) plotted against blocks of four trials each, Experiment 3.

presumably attended to more strongly than in Experiment 2. Correspondingly, learning was somewhat improved. However, despite the fact that the start symbols attracted attention, the majority of subjects still remained blind to the consistent reinforcement of the critical S–R mappings. Thus, we conclude that the difficulties in accomplishing the S–R condition are not exclusively due to an attentional bias. No doubt, if we were to explicitly instruct the subjects to search for the dependency of reinforcements from the start symbol, subjects would learn the critical S–R mappings as easily as they learn the critical R–O mappings (cf., Experiment 1). Thus, the point we want to make is not that S–R relations *per se* are difficult to learn. Rather, we argue that if subjects pursue changing goals and if the success of their actions depends on the correspondence between outcomes and goals, learning is focused on the formation of associations between actions and reinforced outcomes. Consequently, the formation of simple S–R associations is blocked, even if subjects attend to the critical situational context.[2]

With regard to the modified R–O condition, the data show that the abolition of the overall R–O contingencies of Experiment 2 does significantly delay

[2] Following the suggestion of one of the reviewers, we performed another experiment in order to examine whether learning under the R–O conditions can be also blocked if subjects are explicitly instructed to find out how far the success of the keys depends on the current start symbol. The results clearly show that the formation of R–O associations is not blocked. The mean number of trials to criterion amounts to 48.5.

learning. Obviously, the formation of consistent R–O relations is supported if the keys always result in the same outcome irrespective of whether this outcome is intended or not. However, the data also show that overall R–O contingency is not necessary in order to accomplish the R–O condition: All subjects (except one) still learned that for each goal there is a certain key that allows its attainment. This finding is consistent with the assumption that the formation of R–O relations relies primarily on successful key presses, i.e., on reinforcement trials (cf., Figure 4). In other words, if a certain key is consistently reinforced every time a certain goal is aimed for, this is sufficient to represent the corresponding key as being instrumental for attaining the corresponding goal, even if the key results into various other outcomes if other goals are aimed for.

EXPERIMENT 4

The priority of the formation of R–O over S–R associations has been so far demonstrated in comparing learning under conditions in which either S–R mappings or R–O mappings were consistently reinforced. In Experiment 4, subjects were exposed to a mixed reinforcement schedule which comprises both consistent reinforcements of S–R relations and only partly consistent reinforcements of R–O relations. There was, so to speak, an optimal S–R solution and a sub-optimal R–O solution of the learning task. If, as we assume, reinforcements are primarily referred to R–O instead of to S–R relations, subjects should increasingly behave according to reinforced R–O relations, despite them not being consistently reinforced.

Method

Subjects. Fifteen undergraduate psychology students participated in this experiment. None of them took part in any of the preceding experiments.

Procedure. The procedure was the same as in the previous experiment, except for the assignment of the four symbols to the outcomes of the four keys (cf., Table 5). There was again a consistent reinforcement of S–R relations, as the keys ‘‘d’’, ‘‘f’’, ‘‘j’’, and ‘‘k’’ were always correct if the respectively start symbols A, B, C, and D were present. Moreover, the R–O relations were also systematic to some extent, as each key produced a certain symbol more frequently, i.e., keys ‘‘d’’, ‘‘f’’, ‘‘j’’ and ‘‘k’’, respectively, produced the symbols A, B, C, and D in nine of sixteen cases. If one considers the goal-dependent R–O contingencies, which have been shown to be decisive for learning in Experiment 3, each goal symbol could be consistently attained by a certain key except in one of four cases. For example if symbol A is required, key ‘‘d’’ is always successful, except if the start symbol D is present, i.e., the goal-dependent reinforcement rate of the key amounts to 75%. In summary, the present learning task can be accomplished either by mapping the keys to the four

TABLE 5
Structure of the mixed reinforcement schedule of Experiment 4

		Result/symbol + ascending or descending melody			
Start	*Goal*	*Key "d"*	*Key "f"*	*Key "j"*	*Key "k"*
A	A	A	B	C	A
	B	B	B	C	A
	C	C	B	C	A
	D	D	B	C	A
B	*A*	A	A	B	D
	B	A	B	B	D
	C	A	C	B	D
	D	A	D	B	D
C	*A*	A	C	A	D
	B	A	C	B	D
	C	A	C	C	D
	D	A	C	D	D
D	A	D	B	C	A
	B	D	B	C	B
	C	D	B	C	C
	D	D	B	C	D

start symbols, or by mapping the keys to the four goal symbols and by additionally learning the correct keys for the respective exceptions.

Results

Four of the fifteen subjects accomplished the learning task within the limit of 256 trials. Three of them indicated that they had adapted to an S–R rule in the course of learning. These three subjects needed 166 (*SE* 28.16) trials to criterion on average. The remaining 12 subjects took 243.5 (*SE* 12.5) trials to criterion on average and they reported no concrete strategy. However, most of these subjects reported that they had noticed that the keys predominantly resulted in the presentation of a certain symbol and that there was not enough time to figure out the respective exceptions. None of them noticed any relation between the start symbols and the respective correct keys. Thus, the verbal reports suggest that most subjects relied preferably on the experienced relations between keys and the correctly produced symbols, i.e., on R–O relations.

Further analysis was performed in order to validate the verbal reports by behavioural data. In order to do so, a type of data analysis was needed that allows one to assess whether or not individual key-presses are based on reinforced S–R or R–O relations. For such an analysis, only those constellations of start and goal symbols that can be responded to either with a key according to the consistently reinforced S–R relations, or with another key according to the frequently reinforced R–O relations, are suitable. Consider, for example, the case in which the start symbol B and the goal symbol A are present. Subjects are correct if they press either key ‘‘d’’ or key ‘‘f’’. As pressing key ‘‘d’’ corresponds to the fact that key ‘‘d’’ is frequently reinforced if goal A is to be attained, it indicates the use of an R–O rule. On the other hand, pressing key ‘‘f’’ corresponds to the fact that key ‘‘f’’ is consistently reinforced in the presence of the start symbol B, indicating the use of an S–R rule. There were eight such cases in which the choice of one of two correct keys indicates that the key is selected either according to S–R or R–O associations. These eight cases are printed in italics and underlined in Table 5.

The total number of key-presses that indicated a choice according to R–O associations was divided by the total number of key-presses indicating a choice according to S–R associations. This results in what we call the goal dependency index (GDI). If the index exceeds the value of 1, it indicates that subjects relied more on R–O associations, and if the index is below the value of 1, it indicates that subjects relied more on S–R associations. Figure 6 shows the mean GDI calculated over blocks of 64 trials (i.e., 32 critical cases) for those 12 subjects who reported no concrete strategy and for the 3 subjects who reported to have adapted to an S–R rule, separately. An ANOVA reveals that the GDI for the group of 12 subjects significantly increases in the course of learning, reaching a

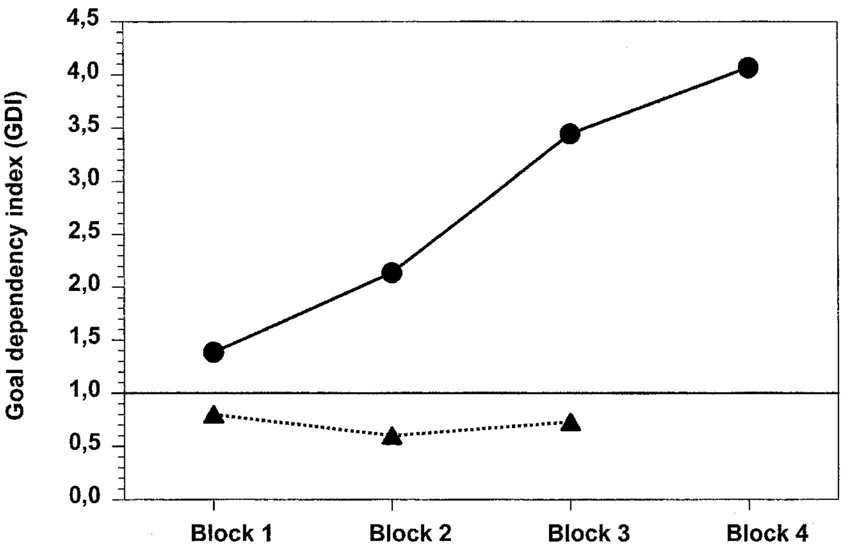

Figure 6. Goal dependency index (GDI) plotted against blocks of 64 trials each. The GDI is separated to those 12 subjects who reported no concrete strategy (circles, solid line) and to the three subjects who recognised the S–R condition (triangles, dotted line).

mean value of 4.06 in the last block of 32 critical cases, [$F(3, 21) = 9.547$, $p < .001$]. This confirms that these subjects indeed increasingly behaved in accordance with reinforced R–O relations. The GDI of the other three subjects lies below 1, but it did not change in the course of learning, [$F(2, 14) < 1$]. This confirms that these subjects mostly (but not always) behaved in accordance with reinforced S–R relations.

Discussion

The results of Experiment 4 confirm that most subjects relied more on relations between their key-presses and successfully produced symbols than on relations between start symbols and successful key-presses. Twelve of fifteen subjects increasingly preferred that one of two similarly successful keys that was frequently reinforced in the presence of the current goal symbol instead of that key that was always reinforced in the presence of the current start symbol. As a consequence, these subjects decreasingly experienced cases in which the keys they had selected were consistently reinforced in the presence of a certain start symbol so that they remained blind to the S–R relations that would have allowed them to handle the task in a very easy way.

GENERAL DISCUSSION

The present experiments departed from the ideomotor hypothesis, according to which the acquisition of behavioural competence is based on the formation of bi-directional action–outcome associations (R–O), so that once an R–O association has been established the desire for the outcome addresses that action which has been learned as resulting in the desired outcome. Recently, several studies have confirmed that the (re)activation of a formerly experienced outcome does indeed facilitate the initiation of the action that consistently resulted in this outcome before (cf., Hommel, 1996, 1998; Kunde, 2000; Kunde, Hoffmann, & Zellmann, 1999, in press; Prinz, 1990, 1997, 1998). Although these findings have revived the discussion of the ideomotor hypothesis, the *formation* of R–O relations is a topic of interest only in animal learning so far (Dickinson, 1994; Holland, 1992; Rescorla, 1990, 1991). In human associative learning, concept formation, covariation learning (including response–outcome covariations), and serial learning dominate the scientific dispute, i.e., paradigms that primarily deal with learning of structural dependencies between stimuli (see Shanks, 1995 for a recent overview). The present study was inspired by the conviction that a systematic investigation of the learning mechanisms that underlie the acquisition of knowledge for the control of goal-oriented behaviour is also needed (cf., Hoffmann, 1993).

In four experiments, subjects had to accomplish a computer-controlled learning task. In each trial an initial condition was illustrated by presenting one of four start symbols on the screen. Moreover, a certain goal to be reached was given by displaying one of four goal symbols. Finally, four different actions were available, consisting of pressing one of four keys. Pressing a key resulted in the presentation of another symbol on the screen (except in Experiment 1). Subjects were required to select the key that would result in the presentation of the current goal symbol, i.e., key-presses were reinforced if their outcome was identical to the goal and they were punished if the two symbols did not correspond. Thus, the experiments resemble the ubiquitous requirement to learn what to do in order to meet changing goals under changing conditions. As we were especially interested to disentangle the contribution of S–R and R–O associations on what subjects learn, the correct keys were either mapped onto the start symbols (S–R condition) or onto the goal symbols (R–O condition). In the S–R condition each key was consistently reinforced when a certain start symbol was present, and in the R–O condition each key was consistently reinforced when a certain goal symbol was present.

Experiment 2 shows that a consistent reinforcement of R–O relations results in fast learning. In contrast, the equally consistent reinforcement of S–R relations are not learned at all by the majority of subjects. Most subjects remained completely blind to the reinforced S–R relations and despaired during the course of the task. Experiment 3 shows that consistent reinforcements of S–R relations remain ineffective for most subjects even if attention is directed to the critical start symbol. Experiment 3 also reveals that the consistent reinforcement of R–O relations remains effective even if the R–O contingencies are abolished in unsuccessful trials. Finally, Experiment 4 shows that if subjects experience both a consistent reinforcement of S–R relations and only a 75% reinforcement of R–O relations, they nevertheless adapt behaviour to the unreliably reinforced R–O relations.

The findings are consistent with the ideomotor hypothesis, as they confirm that, if subjects aim for certain goals, learning mechanisms are addressed that primarily aim at the formation of associations between acts and their reinforced outcomes. Moreover, three specifications of the underlying mechanism are suggested (cf., Figure 4). First, subjects seem to learn above all from successful trials: If a current goal is reached, the successfully produced outcome is associated with the performed action (or an already existing association is strengthened). The outcomes of unsuccessful trials seem not to contribute much to the formation of R–O associations (Experiment 3). Consequently, actions are increasingly connected to the *goals* they consistently allow people to attain, so that overall not R–O but rather goal-related R–O contingencies are learned.[3] Second, as in successful trials the actions are preferredly associated to their outcomes, a simultaneous association to the present situation is suppressed, so

[3] This also makes it plausible that one and the same action can easily be learned as being appropriate for attaining several goals without interference between the various R–O relations to be formed.

that most subjects remain blind to the consistently reinforced S–R relations in Experiments 2–4. Finally, we assume that situational features are taken into account only if R–O relations are not to be established, i.e., if certain acts do not consistently allow attainment of a certain goal. In this case, subjects seem to start to search for features in the current situation which might be responsible for the success or the failure of a certain action in achieving a certain goal. If such features are found, we assume that the R–O relations are conditionalised to them, i.e., S–(R–O) associations are formed.

From an ecological perspective, the presumed learning mechanism can be regarded as being very efficient as it restricts learning only to those R–O relations that are critical to the attainment of the goals the organism really strives for. However, the benefit also has its cost, as dependencies of behavioural success on situational features become hard to detect. The most striking result of the present studies is that most subjects remained blind to the possibility of assigning each key to one start symbol, in order to accomplish the S–R conditions.

If one takes the presumed learning mechanism for granted, the S–R conditions in Experiments 2–4 are indeed challenging: Each of four goals could sometimes be attained by each of four actions and sometimes not. Whether an action was appropriate to attain a certain goal depended on a particular one of four different start symbols. Thus, at least 16 conditionalised S–(R–O) relations were to be formed in order to accomplish the task. This is too difficult a task to be solved in the limited number of 256 trials, especially if one considers that the goals randomly changed from trial to trial so that subjects could not systematically check the appropriateness of the actions for attaining certain goals. Subjects indeed reported that in the search for a solution they also took the start symbol into account as a possible critical condition but mostly without success. That despite these efforts the consistent reinforcement of four simple S–R contingencies remained undetectable for most subjects is consistent with the assumption that subjects remain fixated on the formation of conditionalised S–(R–O) relations.

The ordinary circumstances in pursuing daily goals in general are much easier than in the S–R conditions of the present experiments. First, objects mostly allow only few manipulations and each action usually results in a consistent outcome. Thus, what humans ordinarily experience are consistent three-term relations S–R–O, i.e., in a particular situation only particular goals are attained, always by carrying out a particular action. Second, actions mostly change a given situation so that situations and goals/outcomes are clearly different from each other instead of being exchangeable, as in the present experiments in which the start and the goal/outcome symbols were drawn from the same set. Third, critical features, which the success of an action depends on, are mostly not only statistically but also functionally related to the action–outcome relations they influence. Finally, humans usually strive for a goal until it is reached, i.e., until the appropriate action has been found, instead of being forced to change to another randomly selected goal on each trial. All these circumstances surely influence learning and they might also modify the extent to which situational features are associated with actions in the course of learning. Experiment 1 can be considered as being an illustration for a fast association of actions to situational features under certain conditions.

In Experiment 1, subjects only received feedback about whether or not the actual action was successful. Thus, the only goal subjects could really strive for was to receive positive feedback. In other words, there was only one goal which subjects could continuously try to reach. The results show that under this condition subjects easily learn that it depends on the current situation (i.e., the current start or the current goal symbol) which action is to be performed in order to attain positive feedback. As already discussed, this finding is also consistent with the notion that S–R as well as S–(R–O) associations are formed. In the light of the results of Experiments 2–4 we favour the S–(R–O) account, i.e., we assume that during learning the start/goal symbols become associated not with the keys alone but with key-success expectations.

However, if such S–(R–O) relations have been formed once and the same S–R–O triplets are repeated over and over, the situations may increasingly take over the function of triggering the actions. That situations can indeed acquire the ability to evoke a readiness for the behaviour that repeatedly has been performed in their presence is well documented since, at the latest, the seminal work of Ach (1932) and Lewin (1926). Ach termed it ‘‘voluntionale Objektion’’, Lewin spoke of the ‘‘Aufforderungscharakter’’ of a situation, and Gibson (1979) introduced the concept ‘‘affordances’’. Thus, we do by no means claim that S–R associations are not formed, but we argue that the formation of S–R associations results from training of established S–(R–O) units rather than being a substantial part of their acquisition during initial learning.

We already acknowledged that the intended comparison between consistent reinforcements of S–R and R–O relations led us to use experimental conditions which differ in many respects from ordinary learning environments. It surely is necessary to examine the influence of all of the aspects we have mentioned in order to receive a more complete picture of the mechanisms underlying the acquisition of behavioural competence. We consider the paradigm we have used as being in principle a useful tool for this purpose. The number of different situations, actions, and outcomes can be manipulated as easily as the contingencies between S, R, and O. Likewise, the paradigm allows one to manipulate sequences of situations and goals and thus offers the possibility of controlling the sequence of the events subjects experience. Furthermore, latent learning during free exploration can be compared with intentional learning, etc. Such further research seems to us to be worthwhile, as it will lead to a better understanding of the mechanisms that on the one hand make humans so efficient in the acquisition of the knowledge they need for the control of voluntary behaviour, and on the other hand make us blind to other useful structures in the environment.

Manuscript received May 1999
Revised manuscript received June 2000

REFERENCES

Ach, N. (1932). Finale Qualität (Gefügigkeitsqualität) und Objektion. In *Archiv für die Gesamte Psychologie* (Zweiter Ergänzungsband, pp. 267–366). Leipzig, Germany: Akademische Verlagsgesellschaft.

Allan, L.G. (1980). A note on measurements of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*, 147–149.

Allan, L.G., & Jenkins, H.M. (1980). The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Psychology*, *34*, 1–11.

Alloy, L.B., & Abramson, L.Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, *108*, 441–485.

Colwill, R.M., & Rescorla, R.A. (1985). Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 520–536.

Dickinson, A. (1994) Instrumental conditioning. In N.J. Mackintosh (Ed.), *Animal learning and cognition* (pp. 45–79). San Diego, CA: Academic Press.

Dickinson, A., Shanks, D.R., & Evenden, J. (1984). Judgment of act–outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, *36A*, 29–50.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton.

Greenwald, A.G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review*, *77*, 73–99.

Harleß, E. (1861). Der Apparat des Willens. *Zeitschrift für Philosophie und philosophische Kritik*, *38*, 50–73.

Herbart, J.F. (1825). *Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik und Mathematik. Zweiter analytischer Teil*. Königsberg, Germany: August Wilhelm Unzer.

Hoffmann, J. (1993). *Vorhersage und Erkenntnis: Die Funktion von Antizipationen in der menschlichen Verhaltenssteuerung und Wahrnehmung*. Göttingen, Germany: Hogrefe.

Hoffmann, J., & Sebald, A. (2000). Lernmechanismen zum Erwerb verhaltenssteuernden Wissens. *Psychologische Rundschau*, *51*, 1–9.

Hoffmann, J., & Stock, A. (2000). Intention als psychischer Prozess—eine Suche nach Spuren in der allgemeinpsychologischen Forschung. *Zeitschrift für Psychologie*, *208*, 304–321.

Holland, P.C. (1992). Occasion setting in Pavlovian conditioning. In D.L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 28, pp. 69–125). San Diego, CA: Academic Press.

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research*, *59*, 176–186.

Hommel, B. (1998). Perceiving one's own action—and what it leads to. In J.S. Jordan (Ed.), *Systems theory and apriori aspects of perception* (pp. 143–179). Amsterdam: North Holland.

James, W. (1981). *The principles of psychology* (Vol. 2). Cambridge, MA: Harvard University Press. (Original work published 1890).

Kunde, W. (2001). Response–effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 387–394.

Kunde, W., Hoffmann, J., & Zellmann, P. (1999). Die Wirkung antizipierter Verhaltenseffekte auf die Verhaltensinitiierung. In I. Wachsmuth & B. Jung (Eds.), *KogWis99, Proceedings der 4. Fachtagung der Gesellschaft für Kognitionswissenschaft* (pp. 232–237). Sankt Augustin, Germany: infix.

Kunde, W., Hoffmann, J., & Zellmann, P. (in press). The impact of anticipated action effects on action planning. *Acta Psychologia.*

Lewin, K. (1926). *Versatz, Wille und Bedürfnis mit Vorbemerkungen über die psychischen Kräfte und Energien und die Strukturen der Seele.* Berlin, Germany: Springer.

Lotze, H. (1852). *Medicinische Psychologie oder Physiologie der Seele.* Leipzig, Germany: Weidmann'sche Buchhandlung.

Münsterberg, H. (1889). *Beiträge zur Experimentalpsychologie. Heft 1.* Freiburg, Germany: J.C.B. Mohr.

Pearce, J.M. (1997). *Animal learning and cognition* (2nd ed.). Hove, UK: Psychology Press.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167–201). Heidelberg: Springer.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9,* 129–154.

Prinz, W. (1998). Die Reaktion als Willenschandlung. *Psychologische Rundschau, 49,* 10–20.

Rescorla, R.A. (1990). Evidence for an association between the discriminative stimulus and the response–outcome association in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes, 16,* 326–334.

Rescorla, R.A. (1991). Associative relations in instrumental learning: The eighteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology, 43B,* 1–23.

Rescorla, R.A. (1995). Full preservation of a response–outcome association through training with a second outcome. *Quarterly Journal of Experimental Psychology, 48B,* 252–261.

Shanks, D.R. (1985). Continuous monitoring of human contingency judgment across trials. *Memory and Cognition, 13,* 158–167.

Shanks, D.R. (1995). *The psychology of associative learning.* Cambridge, UK: Cambridge University Press.

Shanks, D.R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 229–261). San Diego, CA: Academic Press.

Thorndike, E.L. (1913). *Educational psychology: The psychology of learning* (Vol. 2). New York: Teacher's College.

Tolman, E.C. (1932). *Purposive behavior in animals and men.* New York: Appleton.

Tolman, E.C. (1949). There is more than one kind of learning. *Psychological Review, 56,* 144–155.

Wasserman, E.A., Chatlosh, D.L., & Neunaber, D.J. (1983). Perception of causal relations in humans: Factors affecting judgments of inter-event contingencies under free-operant procedures. *Learning and Motivation, 14,* 406–432.

Wasserman, E.A., Elek, S.M., Chatlosh, D.L., & Baker, A.G. (1993). Rating causal relations: Role of probability in judgments of response–outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 174–188.

Watson, J.B. (1930). *Behaviorism* (Rev. ed.). Chicago: University of Chicago Press.

Wundt, W. (1893). Grundzüge der Physiologischen Psychologie (Vol. II, 4th ed.). Leipzig, Germany: Engelmann.