Journal of the
Experimental Analysis of Behavior

The power of one reinforcer: The effect of a single reinforcer in the context of shaping

MARY HUNTER AND JESÚS ROSALES-RUIZ

UNIVERSITY OF NORTH TEXAS

During shaping, if the organism is engaged in behaviors other than the current approximation, the amount of time between reinforcers increases. In these situations, the shaper may resort to what is referred to as a "desperation-driven click." That is, after a period of no reinforcement, the shaper delivers one reinforcer for a nontarget approximation. Reports from professional animal trainers suggest that the animal may continue performing this new behavior, even if it is reinforced only once. This study attempted to model this phenomenon with college students. Results from the study demonstrated that a desperation-driven click situation can be reliably produced in a controlled setting. When participants received one reinforcer for interacting with a new object following a period of no reinforcement, they interacted with the new object for a longer or equal amount of time as compared to an object that had a longer history of reinforcement. The results of this study have implications for the understanding of how reinforcement controls behavior.

Key words: acquisition, reinforcement, stimulus control, extinction, humans

In 1952, William Schoenfeld observed, "We don't even know the effect of a single reinforcer presentation on an individual response" (Mechner, 1994, p. 55), and Mechner argued that Shoenfeld's statement continued to be valid in 1994. Even today, there is still truth in this statement. While some studies have looked at the effect of a single reinforcer, questions remain regarding the effect of a single reinforcer during the acquisition, maintenance, and reinstatement of behavior.

Several studies have examined the effect of a single reinforcer by reinforcing only one response during the entire experiment. For example, Skinner (1933) discovered that a substantial extinction curve could be produced when just a single response was reinforced. A rat was habituated to the operant chamber and learned to approach the magazine when it was sounded. Then, the lever was introduced. After about 20 min, the rat pressed the lever for the first time, and this was reinforced with food. After this, the magazine was disconnected. During the 40 min that followed, the rat pressed the lever more than 60 times. This began with a sustained high rate of responding, which was followed by pauses and short bursts of responding. Eventually, pauses became longer until the overall rate of responding was reduced to near zero. This demonstrated that even one reinforcer could have a large effect on the occurrence of a new behavior.

In another study with rats, Iversen and Mogensen (1988) investigated spatial and temporal patterns of responding after the delivery of a single reinforcer in a vertical holeboard apparatus. Each rat received a single 10-min session in an apparatus that contained a matrix of 45 or 54 holes on one side. After 5 min, a single food pellet was presented in one hole. Nose poking through the holes was measured both before and after the delivery of the food pellet. Iversen and Mogensen observed low levels of nose pokes before the delivery of the pellet. After the delivery of the pellet, they reported an increase in nose pokes with the majority of visits occurring in the hole where the pellet had been delivered or the surrounding holes. This increase occurred during the first 2 min after the presentation of the pellet with most rats returning to near baseline levels for the final 3 min.

In a similar study, Mal, McCall, Newland, and Cummins (1993) used a wooden grid with

Mary Hunter, Department of Behavior Analysis, University of North Texas; Jesús Rosales-Ruiz, Department of Behavior Analysis, University of North Texas.

Correspondence concerning this article should be addressed to: Mary Hunter, Department of Behavior Analysis, 1155 Union Circle, Box 311070, University of North Texas, Denton, TX 76201. E-mail: mary@behaviorexplorer.com; Jesús Rosales-Ruiz, Department of Behavior Analysis, 1155 Union Circle, Box 311070, University of North Texas, Denton, TX 76201. E-mail: Jesus.Rosales-Ruiz@unt.edu

doi: 10.1002/jeab.517

40 compartments to deliver one reinforcer to weanling horses halfway through a 10-min session. Their results were similar to those described by Iversen and Mogensen (1988). After finding and consuming the food, the horses visited the apparatus more frequently and visits were clustered around the compartment where food had been delivered. These three studies all support the same conclusion: A single reinforcer delivery for a response that has not been previously reinforced can cause a momentary, but substantial, increase in the frequency of that response.

Other researchers have studied the effect of a single reinforcer by delivering a free reinforcer after the response has been extinguished. This research differs from the previous three studies because it examines the effect of a single reinforcer on a behavior that had previously been reinforced. For example, Reid (1958) conducted research with pigeons, rats, and students. For each subject, an initial response was trained and then extinguished. When the response decreased to a minimal level, a single free reinforcer was given. This reinforcer was "free" in the sense that it was delivered independent of the organism's behavior. In each case, most subjects emitted the previously reinforced response at least once during the observation period that followed. Reid concluded that the single free reinforcer produced this effect because it also functioned as a stimulus, recreating conditions similar to those present during the acquisition of the initial response. That is, the delivery of the reinforcer served as a discriminative stimulus to reestablish the previously taught behavior. Similarly, in research on operant reinstatement, a reinforcer is presented after a period of extinction. Exposing the organism again to the reinforcer leads to an increase in the previously taught response (see Vurbic & Bouton, 2014).

Recent research has also examined how reinforcement signals upcoming contingencies. This type of research differs from the studies cited above because it attempts to measure the effect of reinforcement during concurrent schedules and after much longer exposure to the contingencies. Cowie and colleagues (Cowie, Davison, & Elliffe, 2011, 2014; Cowie, Elliffe, & Davison, 2013), for example, have shown that a reinforcer for Response A may result in an immediate increase in Response B, if the reinforcer predicts that the next reinforcer will more likely be delivered for Response B. According to Cowie and Davison (2016), it may be better to conceptualize reinforcement as a stimulus control effect, rather than as a strengthening effect (see also Baum, 2012). These studies have extended Reid's (1958) finding that a reinforcer may increase behavior because of its discriminative properties.

Many of studies described above examined the effect of giving a single reinforcer for a new response (Iversen & Mogensen, 1988; Mal et al. 1993; Skinner, 1933), the effect of giving a free reinforcer after extinguishing a response (Reid 1958), or the effect of reinforcement in a choice situation with two alternative responses (e.g., Cowie & Davison, 2016). Little research has examined the effect of a single reinforcer on another response after the extinction of the original response. In his investigations regarding superstitious behavior, Skinner (1948) noted that a new behavior could be learned as the result of just one reinforcer. As well, Sidman (2010) discussed that new learning can happen immediately, in one trial, when the learner has previously mastered all of the necessary prerequisite skills.

In applied settings, animal trainers have observed increases in a response after the delivery of a single reinforcer (Bailey, 2013). When shaping novel or complex behaviors, trainers sometimes have difficulty moving from one approximation to the next. This can happen because, when the trainer changes to a new criterion, the animal repeatedly emits behavior appropriate only for a prior criterion. This can lead to extended periods without reinforcement, sometimes eliciting counterproductive emotional responding in the animal (e.g., a dog may begin barking or whining, or leave the context entirely). Sometimes the trainer, in an effort to increase the rate of reinforcement, reinforces a response that does not meet the new criterion. Animal trainer Bob Bailey (2013) called this situation a "desperation-driven click." Anecdotal reports from Bailey and other trainers suggest that the animal repeatedly performs the behavior that was followed by one reinforcer even in the absence of further reinforcement for that behavior. If this behavior persists, it may interfere with the acquisition of the terminal performance.

This study examined the effect of one reinforcer in the context of shaping by creating a period of no reinforcement followed by a single reinforcer for a new response. In particular, we investigated whether, during a 1-min period of extinction, human participants spent more time performing a new behavior that received a single reinforcer after a brief period of no reinforcement or if they spent more time performing a behavior that had been reinforced multiple times.

Method

Participants, Setting, and Materials

Participants included five undergraduate students, two males and three females, from a large public university. Participants were recruited using flyers placed in university buildings and by announcements made in undergraduate behavior analysis classes. Each participant reviewed an informed consent form at the beginning of the study and was informed that participation was voluntary. The consent form included information about the general nature of the study, compensation ($5 per day), and potential risks. Participants were told that the experiment would be filmed, but that only their hands would be visible on the video footage and that their identity would be kept confidential. All procedures and the informed consent form were reviewed and approved by the university's Institutional Review Board.

Apparatus

The game PORTL (Portable Operant Research and Teaching Lab) served as the response apparatus (Rosales-Ruiz & Hunter, 2016). The experiment was conducted in a small room on a folding card table. The participant sat on one side of the table, and the experimenter sat on the opposite side of the table. Stimuli were placed on the table between the experimenter and the participant.

Materials included a collection of 12 small objects, a StarMark dog training clicker, green wooden blocks, and a small blue ceramic dish. The green wooden blocks measured 1 cm x 0.8 cm x 0.8 cm. The 12 small objects were used as manipulanda. They varied in shape, color, and size and consisted of a purple Lego block, a yellow rubber duck, a dice, a red bottle cap, a pink rubber stamp with a picture of a dog, a gray key, a red pencil sharpener, a red domino-shaped wooden rectangle, a green domino-shaped wooden rectangle, an off-white rectangular eraser, a white foam ring, and a blue plastic toy chair (Fig. 1). The smallest item was the dice, which measured 1.5 cm x 1.5 cm x 1.5 cm, and the largest was the blue chair, which measured 2.8 cm x 2.8 cm x 5 cm.

A Canon digital camera attached to a tripod was positioned to the side of the table between the experimenter and the participant. The camera was aimed so that only the table and the participant's hands and arms were visible on the video. In addition, a stopwatch application on an Android phone was used to monitor the passage of time.

Dependent Variables and Measurement

The main dependent measure was the time spent interacting with each object. Interaction with an object was defined as any time that the participant's hand or finger(s) contacted the object. This could include any part of the participant's hand below the wrist. Adding or switching fingers or moving the object in the hand counted as a continuation of the same interaction, as long as part of the hand was constantly in contact with the object. The interaction ended when the participant's hand was no longer in contact with the object. If the participant touched two or more objects at the same time, each object was recorded as a separate interaction. Finally, if the object was out of view, either because it was off-screen, blocked by the participant's hand, or blocked by other objects, an interaction was not recorded until the observer could see part of the hand touching the object. The exception to this rule was if the observer could see the object move and there was no other possible way for the object to move except by contact with the participant's hand.

Data were collected from the recorded videos. Videos were scored frame by frame using the program QuickTime. Data were recorded from the time the experimenter said "start" until the participant had received 10 reinforcers, except for the one-click condition during which recording ended 1 min after the participant touched the 10th object. Duration of time spent interacting with each object was measured as the cumulative number of seconds

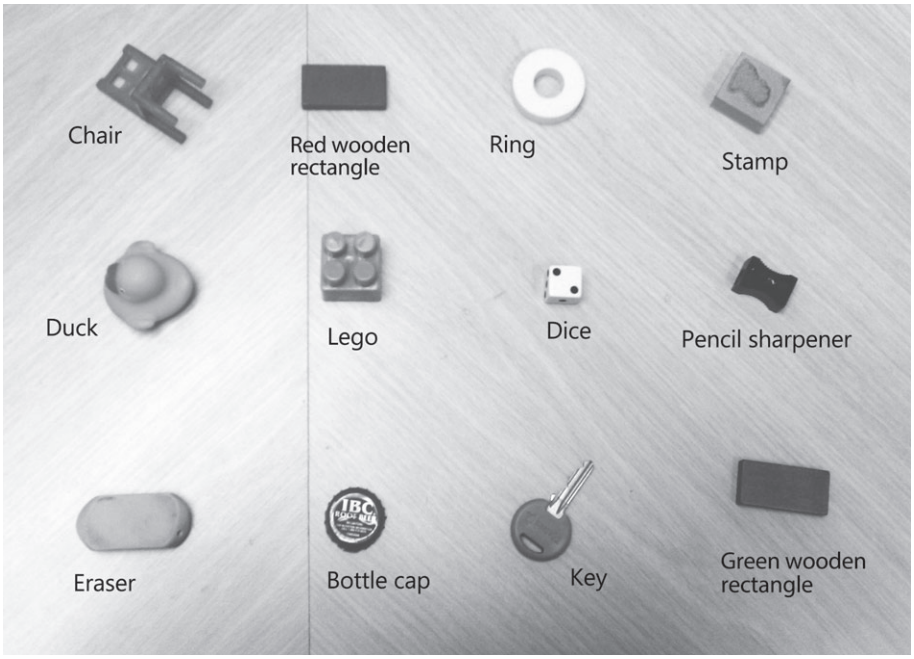

Fig. 1. The 12 objects that were used as experimental stimuli.

during which the participant contacted each object. An interaction was recorded for each time stamp on the video if the participant began, continued, or ended an interaction during that second. Interactions that lasted less than a second were counted as a full second. Because interactions were rounded to a second, if two separate interactions with the same object occurred within the same time stamp, this was counted as a duration of only one second.

Procedures

Preliminary training. Preliminary training included instructions regarding the rules of the game PORTL and an explanation of the clicker and green blocks. First, the participant was told: "Today we are going to play a game with the objects you see sitting on the table. In a minute, I will spread out the objects in the center of the table. Then, when I say 'start,' you can touch or interact with any of the objects in any way you wish. Sometimes, however, when you do certain actions or interact with certain objects, you will hear a click sound and I will hand you a little green block. Your goal today will be to earn as many little green blocks as possible." The participant was also instructed to place the blocks in the blue dish. In addition, the participant was asked to use only one hand when interacting with the objects and was informed that sometimes he or she would have access to all of the objects and other times he or she would have access to only certain objects. Participants were given no further instructions regarding whether they should interact with certain objects or how they should behave with respect to the objects. The experimenter then demonstrated sounding the clicker and handing the participant a block several times.

The clicker and blocks are used in PORTL to create a behavior chain that is analogous to the chain of responses that occur during reinforcement delivery in the operant chamber. The sound of the clicker is functionally similar to the sound of the magazine. It serves as a conditioned reinforcer for correct responses and also as a discriminative stimulus for the participant to hold out a hand to receive a green block. Receiving a block is a discriminative stimulus for the participant to place the block in the dish. This is functionally similar to a consummatory response in an operant chamber (e.g., picking up a food pellet and placing it in the mouth).

In PORTL, the participant can use only one hand to interact with the objects, and blocks are always delivered to this hand. If the participant still has an object in his or her hand after the click, the experimenter holds out the block, but waits for the participant to put the object down before giving the block. This means that the participant's hand must break contact with the object(s) to collect the block and place it in the dish, just as the rat must leave the lever to collect a food pellet from the magazine. If blocks were not used, the participant might continue to hold an object in his or her hand after the click. The use of the blocks allows each response to have a clear beginning and end, which creates a repeatable unit of behavior.

Experiment structure. Each participant took part in the experiment on two separate days. The participant experienced the same set of conditions in the same order on each day. The experimental conditions were divided into smaller units that were called rounds. At the beginning of each round, the experimenter spread out the objects in front of the participant and then said "start." Each round consisted of the delivery of 10 clicks and blocks. The exception to this was the one-click condition, during which only two clicks and blocks were delivered. At the end of each round, the experimenter informed the participant that it was time for a break and removed the objects. Each break lasted approximately 1-2 min. This allowed the experimenter to add or remove certain objects.

Experimental conditions. There were five conditions, the single object training condition, the multiple object training condition, the target object condition, the one-click condition, and the reinforce-all condition. On each day, each condition was implemented for one or more rounds in the order listed above (see Fig. 2). The first two conditions were used to acclimate participants to the apparatus and to teach participants to perform different actions with a single object.

Single object training condition. During the single object training condition, 10 clicks and blocks were delivered during each round. Only one object, the training object, was available. The participant received clicks and blocks for performing a succession of three different actions with this object. Examples of actions included responses such as touching the object, picking up the object, turning the object on its side, pushing the object across the table, turning the object upside down, and twirling the object in the hand. At the beginning of the round, the experimenter delivered clicks and blocks contingent upon the first action the participant performed with the object. If the participant performed other actions, the experimenter waited for the participant to return to this initial action before delivering another click. After three to four clicks and blocks had been delivered, the experimenter stopped clicking for this action and waited for the participant to emit another action. When the participant did so, the experimenter clicked and delivered the next block. Two additional clicks and blocks were then delivered only for this second action. Finally, the experimenter waited for the participant to emit a third action with the object. The final three to four clicks and blocks were delivered when the participant emitted this action. Having the participant perform different actions with the object was meant to simulate shifting between different reinforcement criteria during shaping, thus promoting variations within the response class.

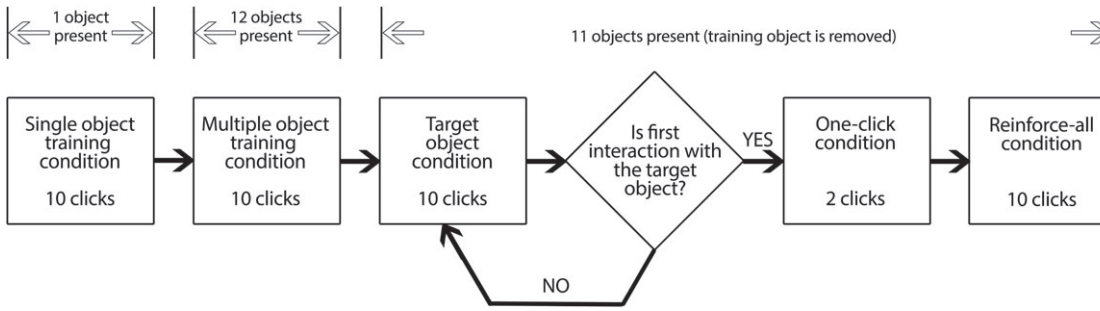

Fig. 2. The order of conditions that was used for each participant.

Multiple object training condition. The multiple object training condition was implemented for the second round on each day. Ten clicks and blocks were delivered during this round. The multiple object training condition was similar to the single object training condition, except all 12 objects were available. Once again, the participant had to emit three different actions with the training object, changing to a new action after the delivery of every three or four clicks and blocks. The participant was allowed to repeat actions that had produced clicks in the previous round. Allowing actions to be repeated deviates from a typical shaping procedure, in which a new criterion would be reinforced at each step. However, this simplification was made because there was a limited number of actions that could be performed with some objects. This also made it easier for the experimenter to keep track of the actions that had been performed and monitor the participant's behavior in real time. During this condition, if the participant interacted with objects other than the training object, the experimenter did nothing. If the participant then returned to the training object and performed an action that no longer earned clicks and blocks, the experimenter did nothing. However, if the participant then left the training object again, but later returned to it, the experimenter clicked and delivered a block. This was done to encourage more interaction with the training object.

Target object condition. The target object condition was meant to simulate reinforcing particular criteria during a shaping session. This condition was also used to establish a history of reinforcement with a particular object, so that responding to this object could be compared later to responding to an object that received only one reinforcer. During this round and all subsequent rounds on a given day, the training object was removed, leaving just 11 objects. Each round in the target object condition consisted of the delivery of 10 clicks and blocks. At the beginning of the first round of the target object condition, the experimenter waited for the participant to touch the 10th object. That is, the experimenter privately tracked which objects the participant had touched and then clicked when the participant touched the penultimate object. This object became the target object for the rest of the day. For the remainder of this round, clicks and blocks were delivered only for interactions with the target object. Similar to the training conditions, the participant had to perform three different actions with the target object, changing to a new action after every three to four clicks.

Response criteria for beginning the one-click condition. Two criteria were used to determine if the target object condition was repeated for more than one round. If the participant did not emit three different actions with the target object, the experimenter implemented the back-up procedure with the target object during the following round. Each round of the back-up procedure was identical to the target object condition, except fewer objects were present. During the first round of the back-up procedure, the participant was given the target object and two other objects. If the participant successfully performed three different actions with the target object, the target object condition was then repeated during the next round with all 11 objects present. If the participant still did not perform three different actions with the target object when only three objects were present, the target object was presented alone in the following round. This was continued for multiple rounds, if necessary, until the participant performed three different actions with the target object when only the target object was present. After this, the experimenter implemented one round of the target object condition with three objects present and then an additional round of the target object condition with all 11 objects present.

Once the participant had performed three different actions with the target object within a single round with all 11 objects present, the experimenter observed if the participant returned immediately to the target object at the beginning of the next round to determine what condition to implement. If the participant began by interacting with an object other than the target object, the target object condition was continued during this round. If the participant began by interacting with the target object at the very beginning of the round, the experimenter implemented the one-click condition during this round.

One-click condition. The one-click condition was implemented for one round on each day. This was the only round that consisted of the delivery of 2 clicks and blocks, rather than 10. When the participant began an interaction

with the target object at the very beginning of the round, the experimenter clicked and delivered the first block. Next, the experimenter stopped delivering clicks and watched as the participant interacted with the other objects. When the participant contacted the 10th object (the second to last object), the experimenter clicked and delivered a second block. This object became the one-click object. Both of the clicks in this condition were delivered contingent on the beginning of any type of interaction with the specified object.

Following the delivery of this second click and block, the experimenter checked the time on the timer and withheld all clicks and blocks for a period of 1 min. The round was finished when this 1-min period had elapsed. This extinction period was limited to 1 min because pilot participants displayed long pauses (without object interaction) and emotional responding (vocal sounds, throwing objects at the experimenter) when longer periods were tried during the initial development of the procedures.

The one-click condition modeled a desperation-driven click situation because a click and block were delivered for a class of responses with a history of reinforcement (interactions with the target object). Then, after a period during which no clicks or blocks were delivered, one click and block were delivered for a different response (interacting with the one-click object). The following 1 min of extinction was then used to assess whether participants spent more time interacting with the one-click object, the target object, or other objects.

Reinforce-all condition. The reinforce-all condition, the final condition, tested whether the participant was more likely to interact with the target object, the one-click object, or some other object after the one-click condition. During this round, the experimenter clicked and delivered a block whenever the participant began any sort of interaction with any object. The participant did not have to perform certain actions with the objects. This round ended after 10 clicks and blocks had been delivered.

Experimental design. The main comparison was between the duration of time spent interacting with the target object after the delivery of the final click and block for the target object at the beginning of the one-click condition and the duration of time spent interacting with the one-click object after the delivery of the final click and block later in the one-click condition. To determine if we could replicate the effect, each participant returned on a second day and all five conditions were repeated in the same order.

Interobserver agreement. Five trained observers collected interobserver agreement (IOA) data on object contacts. Training consisted of explaining the study and data collection procedures. The observer then practiced coding two 20-s video clips under the guidance of the experimenter. Finally, the observer coded a test video of four 20-s video clips without aid from the experimenter. The observer was considered trained if his or her results from the test video showed 90% or greater agreement with data the experimenter collected from the same video clip. IOA data were collected for the entire experiment. However, IOA data were collected only for the training object, the target object, and the one-click object. IOA was calculated for each object during each session using the formula $A/(A + D)*100$, where A was agreements and D was disagreements. IOA was 99.1% for Participant 1 (range 98.4-100%), 98.7% for Participant 2 (range 96.2-99.5%), 90.7% for Participant 3 (range 85.7-97.8%), 99.1% for Participant 4 (range 97.8-100%), and 95.2% for Participant 5 (range 75.3-100%).

Results

Results were graphed as the cumulative number of seconds during which the participant interacted with each object. On each graph, time, as well as the cumulative line for each object, resets to zero at the beginning of each round. In addition, for the one-click condition, the lines for each object and the time reset to zero when the participant received the final click and block. The graph was reset at this point to make it easier to compare the amount of time spent interacting with each object after the last click for the target object to the amount of time spent interacting with each object after the single click for the one-click object. On each graph, the one-click object is always displayed as a solid black line, the target object is displayed as a dashed black line, and the training object is displayed as a dotted black line. The other nine objects are displayed in solid lines of shades of gray.

Individual results for the five participants are shown in Figures 3–7. The top and bottom graphs depict data from each participant's first and second day, respectively. Each graph shows the five conditions implemented for each participant, the single object training condition, multiple object training condition, target object condition, one-click condition, and reinforce-all condition.

On their first day, each participant interacted repeatedly with the training object during the single object training condition and multiple object training condition. When the other objects were first present at the beginning of the multiple object training condition, most participants interacted some with them. However, when each participant returned to interacting with the training object, this produced clicks and blocks, and responding to the rest of the objects decreased to near zero.

The exception to these patterns was Participant 5 (Fig. 7), who interacted with all of the objects during the single object training condition. During this round, she picked up the training object twice, receiving a click and block both times. Then, however, she began reaching across the table and picking up the other objects that were piled next to the experimenter, moving them one by one to her side of the table. Once the participant had moved all 11 of the other objects to her side of the table, she returned to interacting with the training object, and only rarely interacted with any of the other objects. During the multiple object training condition, Participant 5 interacted with only the training object, something not observed with any of the other participants on the first day.

During the target object condition, each participant interacted primarily with the target object across both days and rounds. The target object condition was continued for a second round for Participant 3 on Day 1 (Fig. 5) because this participant did not immediately return to the target object after the first round of the target object condition. For the same reason, the target object condition was repeated on Day 2 for Participants 1, 4, and 5.

During the one-click condition, each participant received one click and block at the very beginning of the round for interacting with the target object. After receiving this first click and block, each participant continued to interact with the target object for some time before beginning to interact with the rest of the objects. When the participant touched the 10th object, the second click and block were delivered. The time between receiving these two clicks varied between participants. For Day 1, the average was 37.2 s, and the range was 30-49 s (Table 1). During this time interval, all participants spent more time interacting with the target object than with any other object.

After receiving the second click and block, no more clicks and blocks were given for the following 1 min. On the first day, most participants interacted repeatedly with the one-click object at the beginning of this period, switching later to interacting with the target object and, to a lesser extent, the other objects. Participant 1 (Fig. 3) and Participant 3 (Fig. 5) spent more time interacting with the one-click object (27 s and 19 s, respectively) than with any other object. For these two participants, the second largest amount of time was spent interacting with the target object. Participant 2 (Fig. 4) spent 11 s interacting with both the one-click object and the target object and 8 s interacting with a third object. This participant did not spend more than 4 s interacting with any of the other objects. Participant 4 (Fig. 6) started out interacting with the one-click object, then switched back and forth between this object and the target object. This participant spent a total of 16 s interacting with the one-click object and 18 s interacting with the target object, which was more than double the amount of time spent with any of the other objects. Participant 5 (Fig. 7) started out interacting with the one-click object, then switched to interacting mainly with another object. These two objects received the most interaction. This participant was the only participant who interacted only rarely with the target object during this part of the experiment.

The final round for each participant was the reinforce-all condition. A summary of the objects that participants interacted with during this condition is displayed in Table 2. On the first day, Participant 1 (Fig. 3) and Participant 5 (Fig. 7) interacted only with the target object, and Participant 3 (Fig. 5) interacted only with the one-click object. Participant 2 (Fig. 4) interacted with eight different objects, and Participant 4 (Fig. 6) interacted exclusively with a single object, which was neither the target object nor the one-click object.

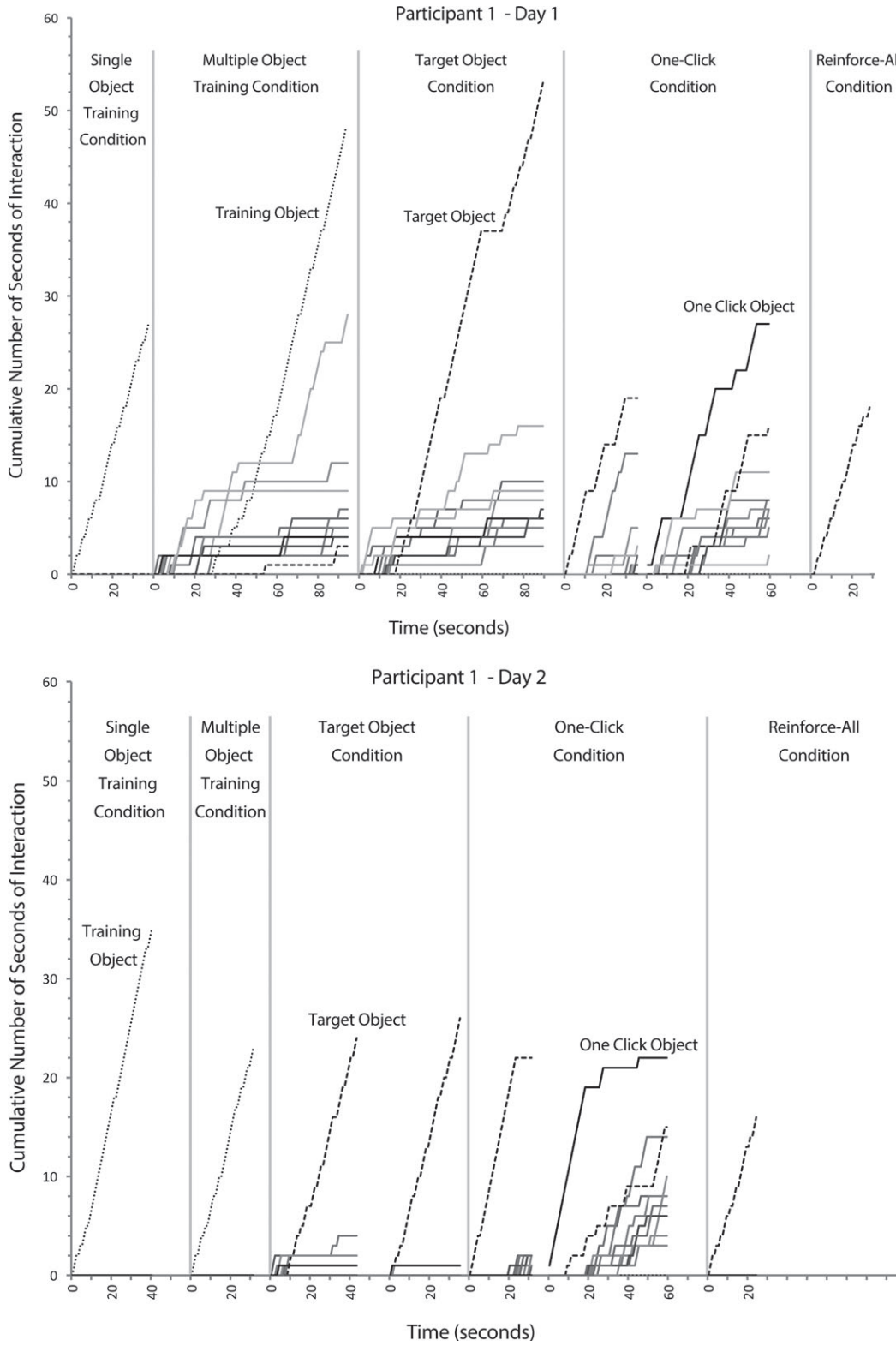

Fig. 3. The cumulative number of seconds that Participant 1 spent interacting with each of the objects during Day 1 (top graph) and Day 2 (bottom graph).

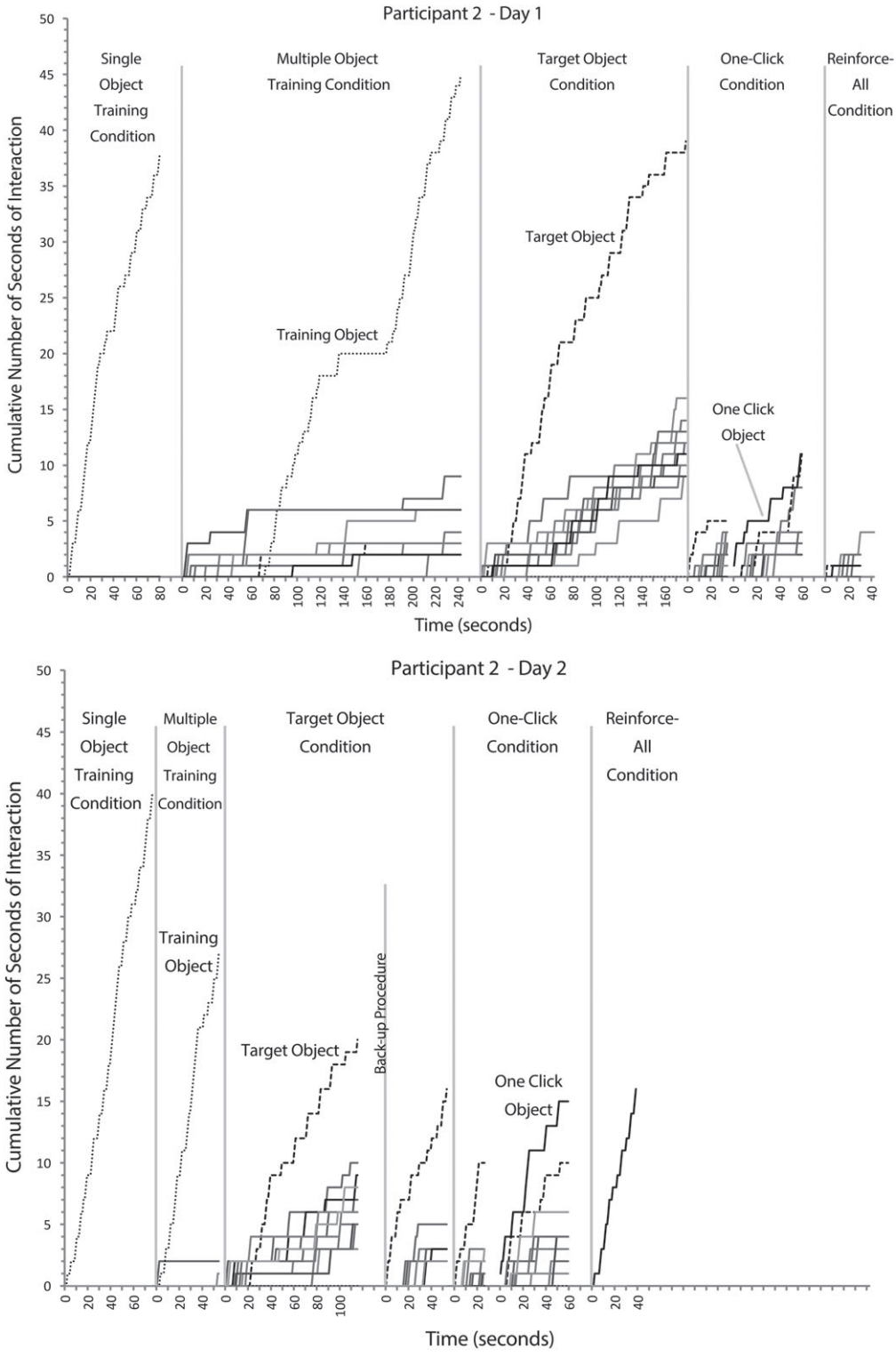

Fig. 4. The cumulative number of seconds that Participant 2 spent interacting with each of the objects during Day 1 (top graph) and Day 2 (bottom graph).

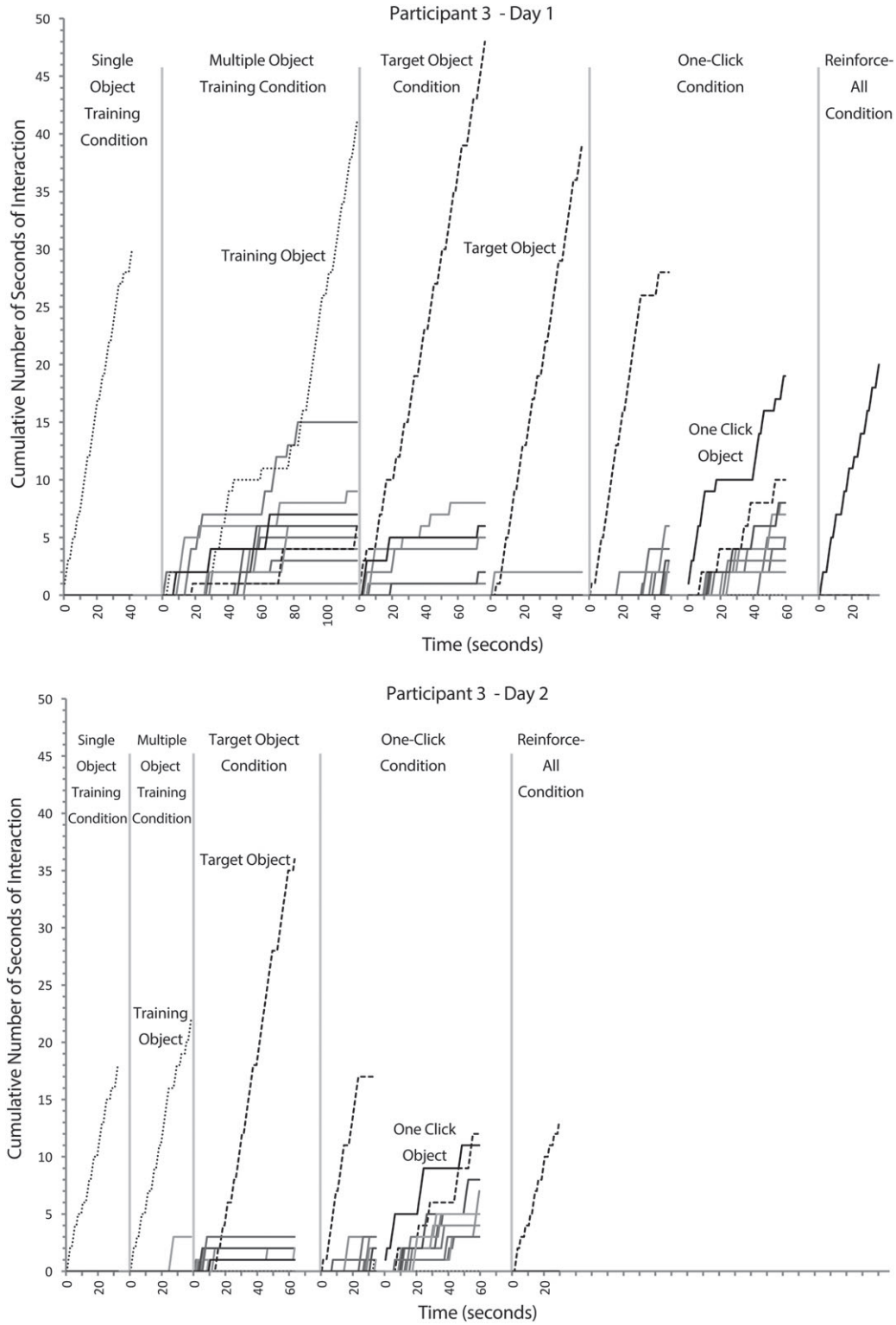

Fig. 5. The cumulative number of seconds that Participant 3 spent interacting with each of the objects during Day 1 (top graph) and Day 2 (bottom graph).

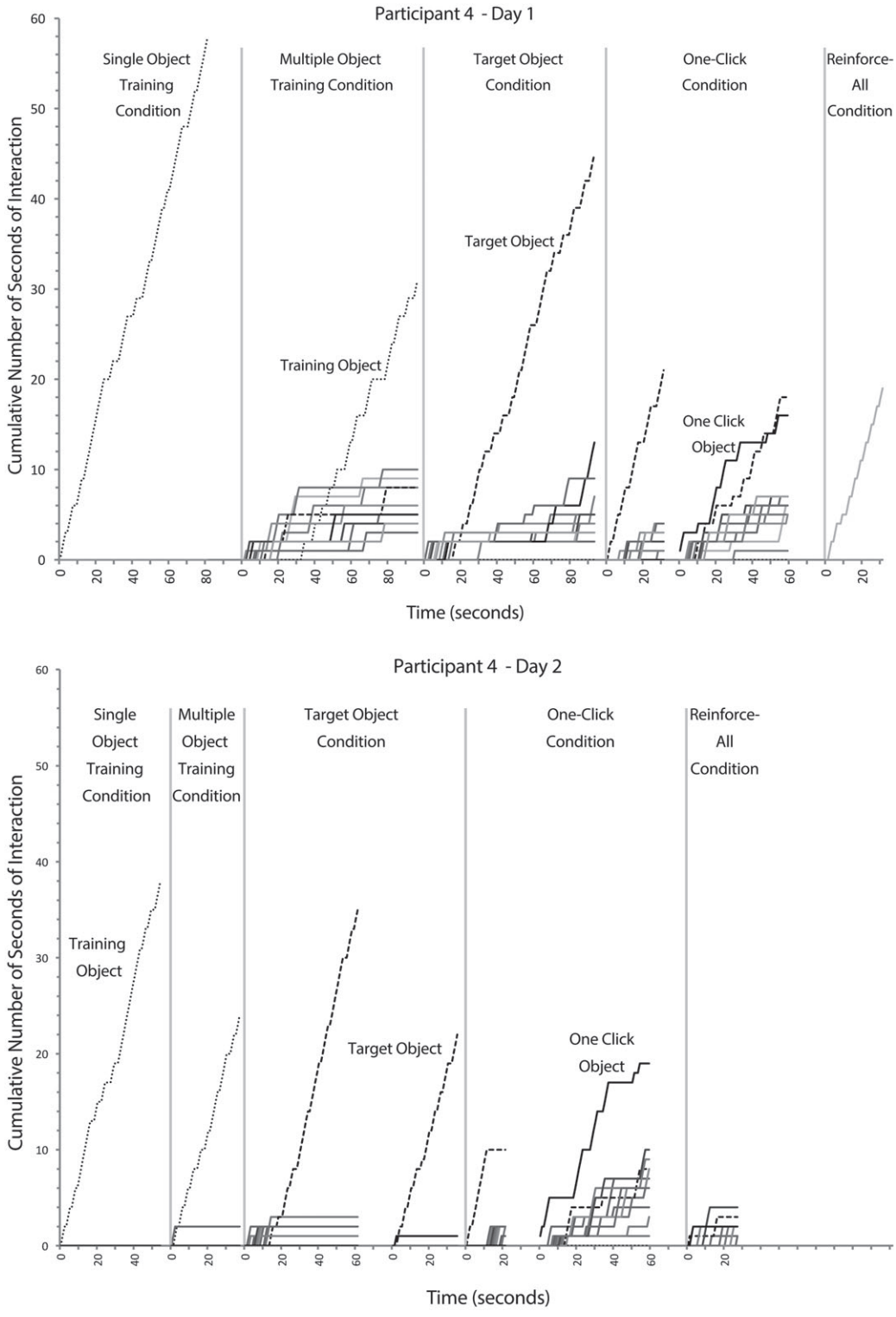

Fig. 6. The cumulative number of seconds that Participant 4 spent interacting with each of the objects during Day 1 (top graph) and Day 2 (bottom graph).

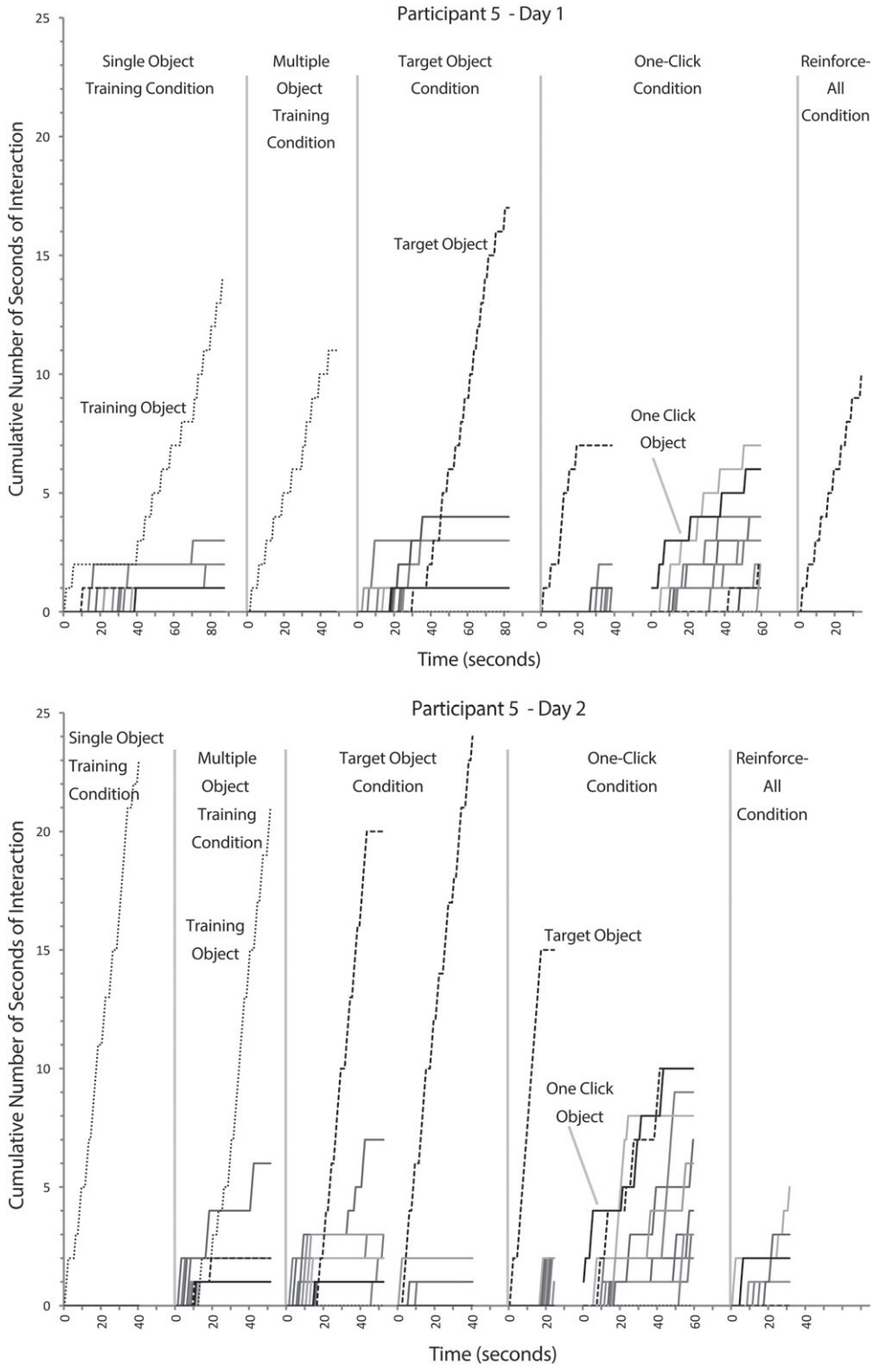

Fig. 7. The cumulative number of seconds that Participant 5 spent interacting with each of the objects during Day 1 (top graph) and Day 2 (bottom graph).

Participants showed similar patterns of responding when the experimental conditions were repeated on the second day. During the multiple object training condition and target object condition, participants interacted primarily with the training object or target object and rarely with the other objects.

During the one-click condition, response patterns were generally similar to what was observed for each participant on his or her first day. This was particularly true for Participant 1, who showed very similar patterns of responding during this condition on both days. Participants 3 and 5 still spent a considerable amount of time interacting with the one-click object. However, the effect was slightly less, as compared to their first day. Participants 2 and 4 seemed to spend slightly more time interacting with the one-click object, as compared to Day 1. For all participants, the time between the two clicks during the one-click condition was shorter on the second day (Table 1). During this time, each participant interacted mainly with the target object (range: 10-22 s), interacting minimally with the other objects.

During the reinforce-all condition on the second day, Participant 1 (Fig. 3) and Participant 3 (Fig. 5) interacted with only the target object, and Participant 2 (Fig. 4) interacted with only the one-click object (Table 1). Participant 4 (Fig. 6) interacted with nine different objects, and Participant 5 (Fig. 7) interacted with six different objects.

Discussion

The results from this study showed that, when participants received a single reinforcer for interacting with a new object following a period of no reinforcement, they spent as much or more time, during extinction, interacting with this new object than another object for which interaction had been reinforced far more often. Also, response allocation substantially favored this new response over other responses that had never been reinforced.

Although a single reinforcer increased the time spent interacting with the one-click object, its effects did not eliminate the participants' history of reinforcement with the target object. As time elapsed without any additional reinforcers, most participants reverted to interacting with the target object more frequently than with the other objects. Thus, like Epstein (1983), we found that a history of reinforcement remains important in determining how behavior is allocated in the face of a change in contingencies.

One main pattern of responding was observed during the 1 min of extinction that began after the one click and block were delivered for interacting with the one-click object. The participant first repeatedly interacted with the one-click object. Then, as responding to the one-click object diminished, the participant began interacting again with all 11 objects. However, the participant interacted more frequently with the target object than with the other nine objects, for which interactions had never been reinforced. We observed two variations to this pattern of responding. In the first variation (seen during Participant 4's second day and during Participant 5's first day), the participant first repeatedly interacted with the one-click object. As responding to the one-click object began to decrease, the participant interacted for approximately the same duration of time with the other objects. However, the participant was still responding to the one-click object at the end of the minute. Possibly, the participant would have interacted more with the target object if the extinction period had lasted for longer than 1 min. In the second variation (seen on Participant 4's first day and Participant 5's second day), the participant interacted nearly equally with the one-click object and the target object

Table 1

Time Interval between the two clicks in the One-Click Condition

Participant	Day 1	Day 2
1	35 s	31 s
2	34 s	26 s
3	49 s	34 s
4	30 s	20 s
5	38 s	23 s

Table 2

Objects touched during the Reinforce-All Condition

Participant	Day 1	Day 2
1	Target object	Target object
2	8 objects	One-click object
3	One-click object	Target object
4	Dice	9 objects
5	Target object	6 objects

throughout the 1 min of extinction. In these two cases, the participant returned much faster to interacting with the target object, yet still continued to interact with the one-click object.

The present study demonstrated that the effect of a single reinforcer could be observed in the context of behavior acquisition. Similar to Skinner (1933), Iversen and Mogensen (1988), and Mal et al. (1993), our data showed that a single reinforcer could momentarily increase the frequency of a new behavior. This momentary increase can be observed when no other responses are reinforced during the experiment, as was demonstrated in these three previous experiments. It can also be seen in the context of acquisition when a previous behavior has been reinforced, as was shown in this experiment. Studying this phenomenon in the context of acquisition allowed us to compare whether participants spent more time performing a previously reinforced behavior or a new behavior that received a single reinforcer.

These results have important implications for therapists, teachers, and animal trainers. An accidental reinforcer may occur if a professional makes an error while applying a procedure, if a professional does not have well-defined criteria for what to reinforce, or if a professional reinforces an unrelated behavior in an effort to increase the rate of reinforcement. Research in applied settings has examined what level of treatment integrity is needed for procedures to be effective (see DiGennaro Reed & Codding, 2014). Practitioners generally accept 80% as an acceptable level of treatment integrity. Recently, some researchers have found that even this level of integrity may interfere with learning (see Bergmann, Kodak, & LeBlanc, 2017). However, this current study suggests that even one accidental reinforcer may result in an increase in an unrelated behavior, which could make it difficult to bring the learner back to approximations to the goal behavior.

The current results differ from what might be expected, based on the Law of Effect (Skinner, 1938; Thorndike, 1911), which would favor the target object. During the first part of the experiment, the participant received multiple reinforcers for interacting with the target object. If these repeated reinforcers strengthened this behavior, it could be predicted that the participant would spend significantly more time interacting with this object during extinction, rather than interacting with an object that received only a single reinforcer. However, this was not what was observed.

Instead, the results could be interpreted as a stimulus control effect (see Baum, 2012; Cowie & Davison, 2016; Reid, 1958). In the initial parts of the present experiment, when the participant received the first reinforcer for interacting with a new object, this was followed by at least 10 more reinforcers for interacting with that object, as well as reinforcement for performing different actions with the object. Additionally, the participant received reinforcement for interacting with only one object during each round. The participant had experienced these conditions with both the training object and the target object. Because of this, when the participant received the first reinforcer for interacting with a new object, this would signal that more reinforcement would likely be forthcoming for interacting with this object. This should lead the participant to interact repeatedly with the new object, as was observed in this study. One interpretation of the present results could be that the delivery of the reinforcer for interacting with the one-click object created conditions that were similar to the conditions that were present when the first reinforcer was given for interacting with the training object and again when the first reinforcer was given for interacting with the target object.

Similar to Reid (1958), in this study a reinforcer was given after a period of extinction. However, Reid reported an increase in the original behavior after the delivery of the free reinforcer (see also Neuringer, 1970), whereas the present study observed an immediate increase in a new behavior. Yet, Reid's experiment does not report the rates of responding for alternative behaviors. It could be that the free reinforcer in Reid's experiment led to a large increase in a new behavior and also to an increase in the original behavior, with only the latter being measured. If that was the case, the stimulus control effect from Reid's experiment and the effect observed in this study would be similar.

Future research should investigate what factors contribute to the single reinforcer effect observed in this experiment. It would be

worthwhile to investigate if the same effect could be obtained if the participant had a very minimal history of reinforcement for interacting with the target object. For example, the two training conditions and the target object condition could be omitted, and the participant could receive just a single reinforcer for interacting with the target object at the beginning of the one-click condition. If the increase in responding to the one-click object is produced as a result of stimulus control, the amount of responding to the one-click object would be expected to be much less. This is because a different stimulus control rule would be established. That is, an interaction with a particular object will be reinforced only once. In summary, this study extended the knowledge about the effect of a single reinforcer in the context of shaping. These results have important implications for practitioners concerned with the effects of individual reinforcers during the acquisition of behavior. They also have implications for basic researchers concerned with how behavior is controlled by reinforcement. Even a single reinforcer may direct the learner away from the target behavior. As Davison and Baum (2003) suggest, every reinforcer counts.

References

Bailey, B. (2013, March). *What did Bailey do, and why: A training potpourri*. Presentation at the 5th Annual Art and Science of Animal Training Conference, Denton, TX.

Baum, W. M. (2012). Rethinking reinforcement: Allocation, induction, and contingency. *Journal of the Experimental Analysis of Behavior, 97*(1), 101-124. https://doi.org/10.1901/jeab.2012.97-101

Bergmann, S. C., Kodak, T. M., & LeBlanc, B. A. (2017). Effects of programmed errors of omission and commission during auditory-visual conditional discrimination training with typically developing children. *The Psychological Record, 67*(1), 109-119. https://doi.org/10.1007/s40732-016-0211-2

Cowie, S., & Davison, M. (2016). Control by reinforcers across time and space: A review of recent choice research. *Journal of the Experimental Analysis of Behavior, 105*(2), 246-269. https://doi.org/10.1002/jeab.200

Cowie, S., Davison, M., & Elliffe, D. (2011). Reinforcement: Food signals the time and location of future food. *Journal of the Experimental Analysis of Behavior, 96* (1), 63-86. https://doi.org/10.1901/jeab.2011.96-63

Cowie, S., Davison, M., & Elliffe, D. (2014). A model for food and stimulus changes that signal time-based contingency changes. *Journal of the Experimental Analysis of Behavior, 102*(3), 289-310. https://doi.org/10.1002/jeab.105

Cowie, S., Elliffe, D., & Davison, M. (2013). Concurrent schedules: Discriminating reinforcer-ratio reversals at a fixed time after the previous reinforcer. *Journal of the Experimental Analysis of Behavior, 100*(2), 117-134. https://doi.org/10.1002/jeab.43

Davison, M., & Baum, W. M. (2003). Every reinforcer counts: Reinforcer magnitude and local preference. *Journal of the Experimental Analysis of Behavior, 80*(1), 95-129. https://doi.org/10.1901/jeab.2003.80-95

DiGennaro Reed, F. D., & Codding, R. (2014). Advancements in procedural fidelity assessment and intervention: Introduction to the special issue. *Journal of Behavioral Education, 23*, 1-18. https://doi.org/10.1007/s10864-013-9191-3

Epstein, R. (1983). Resurgence of previously reinforced behavior during extinction. *Behavior Analysis Letters, 3*, 391-397.

Iversen, I. H., & Mogensen, J. (1988). A multipurpose vertical holeboard with automated recording of spatial and temporal response patterns for rodents. *Journal of Neuroscience Methods, 25*(3), 251-263. https://doi.org/10.1016/0165-0270(88)90140-9

Mal, M. E., McCall, C. A., Newland, C., & Cummins, K. A. (1993). Evaluation of a one-trial learning apparatus to test learning ability in weanling horses. *Applied Animal Behaviour Science, 35*(4), 305-311. https://doi.org/10.1016/0168-1591(93)90082-Z

Mechner, F. (1994). *The revealed operant: A way to study the characteristics of individual occurrences of operant responses* (3rd ed.). Cambridge, MA: Cambridge Center for Behavioral Studies.

Neuringer, A. J. (1970). Superstitious key pecking after three peck-produced reinforcements. *Journal of the Experimental Analysis of Behavior, 13*(2), 127-134. https://doi.org/10.1901/jeab.1970.13-127

Reid, R. L. (1958). The role of the reinforcer as a stimulus. *British Journal of Psychology, 49*(3), 202-209. https://doi.org/10.1111/j.2044-8295.1958.tb00658.x

Rosales-Ruiz, J., & Hunter, M. (2016). PORTL: Your portable Skinner box. *Operants*, Quarter IV, 34-36.

Sidman, M. (2010). Errorless learning and programmed instruction: The myth of the learning curve. *European Journal of Behavior Analysis, 11*, 167-180. https://doi.org/10.1080/15021149.2010.11434341

Skinner, B.F. (1933). "Resistance to extinction" in the process of conditioning. *Journal of General Psychology, 9*, 420-429. https://doi.org/10.1080/00221309.1933.9920945

Skinner, B.F. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: Appleton-Century-Crofts.

Skinner, B. F. (1948). Superstition in the pigeon. *Journal of Experimental Psychology, 38*, 168-172. https://doi.org/10.1037/h0055873

Thorndike, E. L. (2000/1911). *Animal intelligence*. New Brunswick, NJ: Transaction Publishers.

Vurbic, D., & Bouton, M. E. (2014). A contemporary behavioral perspective on extinction. In F. K. McSweeney and E. S. Murphy (Eds.), *The Wiley Blackwell handbook of operant and classical conditioning* (pp. 53-76). Malden, MA: John Wiley & Sons.

Received: December 2, 2017
Final Acceptance: April 11, 2019
Editor in Chief: Amy L. Odum
Associate Editor: Iser G. DeLeon