EVALUATIONS OF THE OVERJUSTIFICATION HYPOTHESIS

By

KERRI PATRICIA PETERS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2013
To my loving family and friends who are like family
ACKNOWLEDGMENTS

I would like to acknowledge Timothy Vollmer for outstanding advising and for his mentoring both inside and outside the laboratory; I cannot provide enough thanks. Additionally, I would like to thank Brian Iwata, Scott Miller, and Maureen Conroy for sitting on my defense committee and for comments on a previous draft of this paper. I would like to acknowledge Sean Peters, Michael Berard, and Allison Hamilton for their constant encouragement and support, and Kyle Peters for changing my life and becoming my inspiration.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS ..........................................................</td>
</tr>
<tr>
<td>LIST OF TABLES ........................................................................</td>
</tr>
<tr>
<td>LIST OF FIGURES .......................................................................</td>
</tr>
<tr>
<td>ABSTRACT ..................................................................................</td>
</tr>
<tr>
<td>CHAPTER ...................................................................................</td>
</tr>
<tr>
<td>1 INTRODUCTION ..........................................................................</td>
</tr>
<tr>
<td>2 GENERAL METHOD, STUDIES 1-3 ..............................................</td>
</tr>
<tr>
<td>Subjects, Setting, and Response Definitions ..............................</td>
</tr>
<tr>
<td>Data Collection and Analysis ..................................................</td>
</tr>
<tr>
<td>General Procedure .....................................................................</td>
</tr>
<tr>
<td>Baseline ....................................................................................</td>
</tr>
<tr>
<td>Reward ......................................................................................</td>
</tr>
<tr>
<td>3 STUDY 1: THE EFFECTS OF CONTINGENT TANGIBLES ON ENGAGEMENT WITH A SINGLE ITEM .............................................</td>
</tr>
<tr>
<td>Method ......................................................................................</td>
</tr>
<tr>
<td>Subjects ...................................................................................</td>
</tr>
<tr>
<td>Design and Procedure ..................................................................</td>
</tr>
<tr>
<td>Interobserver Agreement .........................................................</td>
</tr>
<tr>
<td>Results and Discussion ............................................................</td>
</tr>
<tr>
<td>4 STUDY 2: THE EFFECTS OF CONTINGENT TANGIBLES ON ENGAGEMENT WITH MULTIPLE ITEMS ........................................</td>
</tr>
<tr>
<td>Method ......................................................................................</td>
</tr>
<tr>
<td>Subjects ...................................................................................</td>
</tr>
<tr>
<td>Design and Procedure ..................................................................</td>
</tr>
<tr>
<td>Interobserver Agreement .........................................................</td>
</tr>
<tr>
<td>Results and Discussion ............................................................</td>
</tr>
<tr>
<td>5 STUDY 3: THE EFFECTS OF EXTENDED EXPOSURE ON ACTIVITY ENGAGEMENT .........................................................</td>
</tr>
<tr>
<td>Study 3A: Evaluation of the Effects of an Extended Baseline with Multiple Activities ................................................</td>
</tr>
<tr>
<td>Method ......................................................................................</td>
</tr>
</tbody>
</table>
Subjects .......................................................................................................................... 28
Design and procedure ...................................................................................................... 28
Interobserver agreement ................................................................................................. 28
Results and Discussion .................................................................................................... 28

Study 3B: Evaluation of the Effects of Contingent Tangibles on Activity
Engagement with an Extended Baseline Control ............................................................ 30
Method ............................................................................................................................. 32
Subjects ............................................................................................................................ 32
Design and procedure ...................................................................................................... 32
Interobserver agreement ................................................................................................. 33
Results and Discussion .................................................................................................... 33

6 STUDY 4: REPLICATION OF DECI (1971) .................................................................. 39
Method ............................................................................................................................. 41
Subjects and Setting ......................................................................................................... 41
Design and Procedure ...................................................................................................... 41
Interobserver Agreement ................................................................................................. 42
Results and Discussion .................................................................................................... 42

7 GENERAL DISCUSSION ................................................................................................. 48

LIST OF REFERENCES ..................................................................................................... 53

BIOGRAPHICAL SKETCH ................................................................................................. 56
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-1</td>
<td>Condition averages of percentage of session engaged with target item for all sessions in each condition and the last three sessions in each condition for each subject</td>
</tr>
<tr>
<td>5-1</td>
<td>Condition averages of percentage of session engaged with Set 1 target item and Set 2 comparison item for each subject. Condition lengths for Set 2 comparison items were based on condition lengths for corresponding Set 1 item. The top row for each subject displays the average for all sessions in each condition and the bottom row displays the average of the last 3 two-session blocks for each condition</td>
</tr>
<tr>
<td>6-1</td>
<td>Number of Seconds Working on the Puzzle During the Free-Choice Period</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>Percentage of session engaged with the single item for Angel, Terrance, Georgette, and James.</td>
<td>20</td>
</tr>
<tr>
<td>4-1</td>
<td>Percentage of session engaged for Sam and Angel (top two panels display all available items and bottom panels display only target activity for Sam).</td>
<td>24</td>
</tr>
<tr>
<td>5-1</td>
<td>Percentage of session engaged with all available items for Angel, Terrance, Georgette, and James.</td>
<td>30</td>
</tr>
<tr>
<td>5-2</td>
<td>Percentage of session engaged with item with the initially highest percentage of engagement for Angel, Terrance, Georgette, and James.</td>
<td>31</td>
</tr>
<tr>
<td>5-3</td>
<td>Percentage of session engaged with all items for Jack, Kayla, and Dennis for Set 1 and Set 2 (top and bottom panels, respectively). Sessions are combined in to two-session blocks along the x-axis.</td>
<td>36</td>
</tr>
<tr>
<td>5-4</td>
<td>Percentage of session engaged with target items for Jack, Kayla, and Dennis for Set 1 and Set 2 (top and bottom panels, respectively). Sessions are combined in to two-session blocks along the x-axis.</td>
<td>37</td>
</tr>
<tr>
<td>6-1</td>
<td>Number of seconds engaged with the puzzle during each break for individuals in the experimental group.</td>
<td>46</td>
</tr>
<tr>
<td>6-2</td>
<td>Number of seconds engaged with the puzzle during each break for individuals in the control group.</td>
<td>47</td>
</tr>
</tbody>
</table>
The utility of reinforcement-based procedures has been well established in the behavior analysis literature. However, the overjustification effect is one commonly cited criticism of programs that use rewards. Rewards are frequently delivered in school settings and include grades, stickers, social praise, and star charts. The overjustification hypothesis suggests that the delivery of an extrinsic (socially mediated) reward contingent on engagement with an activity that previously occurs at some level without apparent socially mediated reinforcement will result in a reduction in the amount of engagement in that activity from baseline levels when the reward phase is discontinued. This series of studies evaluated the effects of delivering tangible rewards contingent on engagement with age-appropriate leisure activities on the amount of engagement in the condition after reinforcement was discontinued. The subjects in Studies 1-3 were children in first grade. In these studies, tangible rewards were delivered for engaging with the preferred activities under different conditions, for example when one or several items was or were available at a time. The results did not support the overjustification hypothesis; however they suggest that extended exposure to a preferred item may have effects that look similar to the overjustification effect. Study 4 was a direct replication of
a landmark study evaluating the overjustification hypothesis conducted by Deci (1971).
The subjects in Study 4 were undergraduate students. Contrary to previous findings, both groups showed a decrease in engagement across the study and individual data varied greatly, not providing support for the overjustification effect.
CHAPTER 1
INTRODUCTION

The utility of reinforcement-based procedures has been well established in the behavior analysis literature. However, the overjustification effect is one popular criticism of programs that use rewards (Deci, 1971; Kohn, 1993; Lepper & Greene, 1975; Lepper, Greene & Nisbett, 1973; Pink, 2009). The overjustification hypothesis states that the delivery of some extrinsic reward will decrease an individual’s intrinsic motivation to engage in the behavior that produced the reward (Roane, Fisher, & McDonough, 2003). This is often referred to as the “undermining” effect of rewards and reinforcement (Cameron & Pierce, 2004; Vasta, Andrews, Stirpe, & Comfort, 1978). The terms reward and reinforcer have been used interchangeably throughout the overjustification literature. A reinforcer is defined by its effect on behavior. Specifically, a reinforcer is an event that increases the probability of the behavior it follows. A reward is a stimulus that is assumed to have a positive effect on behavior, but is not defined by its effects on behavior; an increase in the probability of behavior has not necessarily been demonstrated (Cameron & Pierce). The commonly used definition of the overjustification effect does not require demonstration of a reinforcement effect, only delivery of an extrinsic reward (e.g., Deci, 1971; Lepper, Greene, & Nisbett, 1973). This is because the behavior of interest is already occurring at a high rate so increases in rate as a result of reinforcement are not necessary to see the effect (i.e., a decrease in rate across a pre-reward phase to a post-reward phase).

According to the overjustification hypothesis, the delivery of an extrinsic (socially mediated) reward contingent on engagement with an activity maintained by automatic reinforcement will, when discontinued, result in a reduction in the amount of
engagement with that activity below baseline levels. Based on the research showing this effect, many educational entities recommend against the use of rewards. For example, the Des Moines Public School ("Des Moines public," 2011) has a webpage dedicated to motivating student learning and the first line reads, “How rewards can backfire and reduce motivation,” and cited Lepper et al. (1973) and Deci (1971). In addition, one of the eight principles of Montessori education is that extrinsic rewards have a negative impact on long-term motivation and learning (Edwards, 2006). Edwards stated that this principle has the “strongest direct evidence,” and cited Deci (1971). Yet, the vast literature on the overjustification effect is riddled with methodological shortcomings, few efforts have been made to understand mechanisms accounting for the effect or lack thereof, and the primary study from which the effect is inferred has never been replicated despite the fact that several important measures were not reported.

One of the most commonly cited series of studies, with over 2,250 citations, on the overjustification effect was conducted by Deci (1971). Deci evaluated the effects of money and verbal praise on undergraduate students’ performances on a puzzle task. Deci concluded that intrinsic interest decreased when money was used as a reward and increased when verbal praise was used. Intrinsic interest was measured by recording the number of seconds the subjects worked on the puzzle task during a break period between sessions in which the subject was alone in the room and the subject’s behavior was recorded covertly.

Subsequently, there were several studies published in the 1970s addressing the issue of the detrimental effects of rewards (e.g., Lepper & Green, 1975; Lepper et al.,
Cameron, Banko and Pierce (2001) reported that in the span of 30 years over 100 studies were performed to evaluate the effects of extrinsic rewards. Tasks have varied across studies and included dot-to-dot activities (Feingold & Mahoney, 1975), puzzles (Deci, 1971, 1972; Lepper & Green, 1975), coloring with markers (Lepper et al., 1973), listening to songs on a tape recorder (Reiss & Sushinsky, 1975), sorting silverware (Roane et al., 2003), and classroom rule following (Akin-Little & Little, 2004), among many others.

In addition to all of the empirical research on the topic, there have been literature reviews. First, Cameron and Peirce (1994) evaluated the results of 96 studies using group designs and 5 studies using single subject designs. Based on the results of their statistical analyses, they concluded overall that rewards do not decrease intrinsic motivation. In 1999, Deci, Koestner, and Ryan published a meta-analysis claiming that Cameron and Pierce’s review was seriously flawed. Deci et al. claimed that Cameron and Pierce used inappropriate procedures and made numerous errors. Based on the results of 128 studies, Deci et al. concluded that tangible rewards do have a substantial undermining effect on intrinsic interest. In 2001, Cameron, Banko and Pierce readdressed the issue, suggested that both of the previous meta-analyses had flaws, and attended to the issues brought up by Deci et al. The authors came to the conclusion that rewards overall do not have a negative effect, but under some conditions (e.g., tangible rewards delivered contingent on performance on high interest tasks) there might be some decrement in interest.

The contradictory results from the meta-analyses of the research on the overjustification effect may be due to the wide variation in methods used to study the
effect (Cameron et al., 2001; Cameron & Pierce, 1994; Deci, Koestner & Ryan, 1999; 2001). The majority of the research in the area used either a pre- and post-test paradigm without taking repeated measures of behavior (e.g., Anderson, Manoogian & Reznick, 1976; Deci, 1971; Lepper & Greene, 1975) or involved repeated measures but only reported averages (e.g., Mynatt, Oakley, Arkkelin, Piccione, Margolis, & Arkkelin, 1978; Lepper et al., 1973). In addition, aspects of the experimental sessions, including types of task, types of reward or reinforcer, and whether or not there are other items present during the session differ across study. Some allowed access to other activities during the experimental sessions (Deci, 1971; Feingold & Mahoney, 1975), others did not include an alternative activity (Lepper & Greene, 1975; Roane et al., 2003; Bright & Penrod, 2009), and some studies did not specify (Anderson et al., 1976).

A result of the mixed meta-analytic outcomes and mixed methodology is that the overjustification effect is still commonly cited in psychology textbooks (e.g., Franzoi, 2010; Griggs, 2009; Lilienfield, Lynn, Namy, & Woolf, 2010) and in the popular media (e.g., Bronson & Merryman, 2009; Kohn, 1993; Pink, 2009) as a problem with the use of rewards. Also, no study has attempted to directly replicate Deci’s original procedures. Finally, no experiments have attempted to identify the conditions under which overjustification-like effects would and would not occur.

The current dissertation involves a series of studies evaluating the overjustification hypothesis. The purpose of Study 1 was to evaluate the effects of contingent tangible rewards on activity engagement when a single activity was available in order to evaluate whether the overjustification effect is observed when there are no other items available with which to engage. Study 2 examined the effects of contingent
tangible rewards on activity engagement when multiple activities were available in order to evaluate whether the overjustification effect is observed when the subject has continuous opportunity to engage with other activities. Although on the surface Study 2 may seem to be a comparison to Study 1, this was not intended to be the case. Rather, Study 2 is simply designed to test for an overjustification effect when multiple items are available. The purpose of Studies 3A-B was to evaluate the effects of extended exposure to activities when multiple activities were available. Study 3A was a descriptive study that assessed the effects of an extended baseline with multiple activities. Study 3B was an extension of Mynatt et al. (1978) that compared the effects of contingent tangible rewards on activity engagement when multiple activities were available to a within-subject control for exposure (i.e. extended baseline). Given that Deci (1971) is frequently cited in support of the overjustification effect and yet individual data were not presented and there has never been a replication of this study, Study 4 was a direct replication of the first experiment of Deci (1971) with college-age students in order to assess within-subject measures.
CHAPTER 2
GENERAL METHOD, STUDIES 1-3

Subjects, Setting, and Response Definitions

Subjects were recruited from two first-grade classrooms or from the afterschool program at a local elementary school. All subjects were 6 or 7 years-of-age at the beginning of the study. Sessions took place at a large table in a quiet area of the school library or the school cafeteria. For all of the studies experimenters measured engagement with each activity; engagement was defined as touching the item or items corresponding with that activity with a three second onset and offset criterion (i.e., if the duration of contact was less than 3 s, engagement was not scored). Activities were chosen based on age-appropriateness (listed on the packaging) and arbitrarily assigned to sets. Subjects could be engaged with multiple activities at the same time.

Data Collection and Analysis

Data were collected using handheld devices (HP® iPAQ) with a behavioral recording program called Instant Data. The program allowed frequency and duration measures to be collected for multiple responses across time. For Studies 1-3A, a second, independent, observer collected data using a handheld stopwatch and for Study 3B the second observer collected data using a second handheld device programed with Instant Data. Duration measures for activity engagement were converted to percentages by dividing the total duration of engagement by the session time and multiplying by 100.
General Procedure

Baseline

During baseline conditions there was one or more (depending on the study) age-appropriate leisure activities (e.g., puzzles, coloring books, and sticker books) within arm's reach on the table in front of the subject. There were no programmed consequences for engagement with the activities. Sessions lasted 5 min and one or two sessions were conducted per day. At the beginning of each baseline session the subject was told, “You have five minutes to play. You may play with whatever you like on the table. I will tell you when five minutes has ended.”

Reward

Reward conditions were set up the same as the baseline sessions. At the beginning of each reward session the subject was asked to choose from an array of small edible items. One small piece was placed on a plate in front of the subject contingent on 30 s continuous engagement with the target activity; the target activity was determined from the baseline levels of engagement. An activity with somewhat high (>30%) engagement in baseline was selected as the target activity in the reward condition. The 30% value was chosen to ensure that there was some level of responding to the activity, or intrinsic interest, prior to the contingent reward. At the beginning of each reward session the subject was told, “You have five minutes to play. You may play with whatever you like on the table. You will only get (preferred edible) if you play with (the target activity). I will tell you when five minutes has ended.”
The availability of other activities for the child to engage with during the baseline and experimental sessions is either not reported or inconsistent across studies evaluating the overjustification effect. Several studies did not include an alternative activity during any of the sessions (Bright & Penrod, 2009; Lepper & Greene, 1975; Roane et al., 2003) and several studies specified that alternative activities were available throughout the sessions (Deci, 1971; Feingold & Mahoney, 1975). Anderson et al. (1976) told the subjects that if they did not want to engage in the target activity, drawing, they could do something else in the room, but whether there were other activities in the room was not reported. The availability of alternative activities could have direct effects on the level of engagement with a target activity. The purpose of Study 1 was to evaluate the effects of contingent tangible rewards on activity engagement when a single item was available and to assess whether overjustification-like effects are seen when only one item is available.

Method

Subjects

Subjects were four typically developing first-grade children recruited from two classrooms in the same local elementary school. There were two male and two female subjects; Terrance, James, Angel, and Georgette.

Design and Procedure

An ABA design was used to evaluate the effects of contingent tangibles on activity engagement, where A was no tangible reward and B was tangible reward. There was a single age-appropriate leisure activity placed within arm’s reach on the table in
front of the subject. Baseline and reward conditions were as described in the General Method.

**Interobserver Agreement**

Interobserver agreement (IOA) was collected by a second observer simultaneously, but independently, collecting data on 33%, 40%, 32%, and 52% of sessions for Angel, Terrance, Georgette, and James, respectively. Total agreement scores were calculated by dividing the smaller duration of observed engagement by the larger duration of observed engagement in each session and multiplying that value by 100. Average agreement scores were 99.6% (range, 99% to 100%), 99.2% (range, 96.7% to 100%), 99.6% (range, 98.7% to 100%), and 99.2% (range, 98.9% to 99.3%) for the four subjects, respectively.

**Results and Discussion**

Figure 3-1 shows the results of Study 1 for Angel, Terrance, Georgette, and James. None of the four subjects showed a sustained detrimental effect of contingent tangibles on levels of activity engagement. In fact, three of the four subjects were engaged with the activity over 98% of the session for all baseline sessions, contingent reward, and return to baseline sessions. James was engaged with the activity 100% of the session for all baseline and contingent reward sessions and above 70% of the session for all return-to-baseline sessions. Thus, the overjustification effect was not observed when only one activity choice was available.
Figure 3-1. Percentage of session engaged with the single item for Angel, Terrance, Georgette, and James.
CHAPTER 4
STUDY 2: THE EFFECTS OF CONTINGENT TANGIBLES ON ENGAGEMENT WITH MULTIPLE ITEMS

The results of Study 1 appear to debunk the overjustification hypothesis. However, a follow-up question is whether the overjustification effect would be obtained if alternative activities were available. Thus, the purpose of Study 2 was to examine the effects of contingent tangible rewards on activity engagement when multiple activities were available and to assess whether overjustification-like effects were observed under these conditions.

Method

Subjects

Three subjects were originally enrolled for Study 2. One of the subjects was excluded from the study because of a diagnosis of Autism Spectrum Disorder (ASD). His data have been included in a follow-up study with subjects with ASD (not reported here. The remaining subjects were Sam and Angel. Sam was a typically developing male in first-grade. Angel was a typically developing female and was also a subject in Study 1.

Design and Procedure

An ABA design was used to evaluate the effects of contingent rewards on activity engagement. There were more than one age-appropriate leisure activities placed within arm’s reach on the table in front of the subject. Sam had five items available and Angel had two. Baseline and reward conditions were as described in the General Method. In the reward condition a small edible item was delivered contingent on 30 s of continuous engagement with the target activity.
The target activity was determined based on the levels of engagement observed during the baseline sessions. The activity was chosen as the target activity if the average percentage engagement with the activity was above 30% for the last three sessions. The 30% value was chosen to ensure that there was some level of responding to the activity, or intrinsic interest, prior to the contingent reward.

**Interobserver Agreement**

Agreement was collected and calculated in the same manner as Study 1. Agreement data were collected during 56% and 30% of sessions, and average agreement scores were 95.1% (range, 82% to 100%) and 99.2% (range, 96.7% to 100%) for Sam and Angel, respectively.

**Results and Discussion**

Figures 4-1 display the results of Study 2 for Sam and Angel. The top two panels of Figure 4-1 show the patterns of responding for all available items. Stamps were chosen as the target activity for Sam, and benders were chosen as the target activity for Angel. Due to the large number of data paths some of the effects may be obscured, so the bottom panel of Figure 4-1 displays only the percentage of session engaged with the target activity for Sam. For both subjects there was a decrease in activity engagement from the first to the second baseline condition. During the reward condition for both subjects, engagement with the target activity increased to 100% of the session. For Sam, engagement with the stickers was above 80% of the session at the end of the first baseline and dropped to zero after the second session of the second baseline. Angel engaged with the benders during 100% of the session for five consecutive sessions in the first baseline. Immediately, engagement with benders dropped to 0% of the session during the second baseline. Table 4-1 shows the average percentage of
session engaged with the target activity in each condition for each subject. In addition
the average of the last three sessions in each condition is shown for both subjects.
From the first baseline to the second baseline there was a 56.68 and 71.43 decrease in
percentage of session engaged for Sam and Angel, respectively. The decreases in
engagement with the target activities for both subjects could provide support for the
overjustification hypothesis or could be a result of repeated exposure to the activities
(one possibility would be satiation). The purpose of Studies 1 and 2 was not to evaluate
the effects of single versus multiple available activities on the probability of post reward
condition decrements, rather to assess separately whether the overjustification-like
effects would be observed given these two conditions. However, post reward
decrements were only observed when multiple activities were available so future
research could focus on a within-subject comparison of these two conditions.
Figure 4-1. Percentage of session engaged for Sam and Angel (top two panels display all available items and bottom panels display only target activity for Sam)
Table 4-1. Condition averages of percentage of session engaged with target item for all sessions in each condition and the last three sessions in each condition for each subject

<table>
<thead>
<tr>
<th>Subject</th>
<th>BL 1</th>
<th>Reward</th>
<th>BL 2</th>
<th>BL 2 - BL1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sam</td>
<td>76.58</td>
<td>75.33</td>
<td>19.9</td>
<td>-56.68</td>
</tr>
<tr>
<td></td>
<td><strong>Last 3</strong></td>
<td>68.78</td>
<td><strong>100</strong></td>
<td><strong>-68.78</strong></td>
</tr>
<tr>
<td>Angel</td>
<td>71.42</td>
<td>96.00</td>
<td>0</td>
<td>-71.42</td>
</tr>
<tr>
<td></td>
<td><strong>Last 3</strong></td>
<td><strong>100.00</strong></td>
<td><strong>100.00</strong></td>
<td><strong>-100.00</strong></td>
</tr>
</tbody>
</table>
Mynatt et al. (1978) criticized the use of a single trial to evaluate the overjustification effect and included conditions with several baseline and reinforcement trials in their study. Mynatt et al. included 9 days in the baseline condition, 11 days in the reinforcement condition, and 9 in the second baseline. In addition to using repeated measures within-subjects, Mynatt et al. used a group design to compare the experimental group to a group with no reinforcement trials designed to control for amount of exposure to the activities. The results showed a high interest level in the first baseline and reinforcement conditions and a decrease to levels that match the control group in the second baseline. Interest was defined as number of times a subject chose a game divided by the number of games chosen in each session. Based on these results, the authors concluded that the decline in interest in the high base rate condition was due to satiation, with the rate of satiation not affected by the reinforcement condition. Unfortunately, the authors averaged the sessions together for each condition so within-condition changes cannot be seen. Also, the 6- and 7-year-old subjects collected the data and interobserver agreement measures were not reported. Thus, there has been very limited research with children to date that has used repeated measures and acceptable observation methods when multiple activities are available. Controls for amount of exposure have typically been across groups and there have been no studies to date that have used a within-subject control.

The purpose of Study 3 was to evaluate the effects of an extended exposure on activity engagement when multiple activities were available and provide an extension to Mynatt et al. There were two phases to Study 3. The purpose of Study 3A was to
observe the effects of an extended baseline with multiple activities and no programmed consequences. The purpose of Study 3B was to compare the effects of a contingent tangible reinforcer on activity engagement to a within-subject control-for-exposure condition. Study 3B involved an extended baseline as a control.
Study 3A: Evaluation of the Effects of an Extended Baseline with Multiple Activities

Method

Subjects.

Subjects were four first-grade students recruited from two classrooms in a local elementary school. All subjects included in this study were also involved in Study 1.

Design and procedure.

There were five age-appropriate leisure activities placed within arm's reach on the table in front of the subject. Subjects were exposed to between eight and fifteen 5-min sessions with no programmed consequences for activity engagement. One to two sessions were conducted per day, two to five days per week.

Interobserver agreement

Agreement was collected and calculated in the same manner as Studies 1 and 2. Agreement data were collected on 50% of sessions for all subjects, and average agreement scores were 90.3% (range, 74.3% to 100%), 90.5% (64.7% to 100%), 97.9% (96.4% to 99.6%) and 98.1% (94.8% to 99.7%) for Angel, Terrance, Georgette, and James, respectively.

Results and Discussion

Figure 5-1 displays the percentage of session engaged with each of the activities and Figure 5-2 displays the percentage of session engaged with the activity with the highest initial engagement for Angel, Terrance, Georgette, and James. For all four subjects, activity engagement was initially stable; engagement with one of the activities remained at approximately 100% of the session for between three to five sessions. Then, for all subjects, engagement with the activity that was at the highest percentage
of the session dropped to zero and subsequently became variable as the subjects sampled the other activities. This finding supports the notion that the presumed overjustification effect may be a satiation effect as described by Mynatt et al. By reinforcing engagement (in a typical overjustification arrangement), the subject spends more time than they ordinarily would have with the item and the decrease in engagement is therefore not entirely surprising.
Figure 5-1. Percentage of session engaged with all available items for Angel, Terrance, Georgette, and James
Figure 5-2. Percentage of session engaged with item with the initially highest percentage of engagement for Angel, Terrance, Georgette, and James
Study 3B: Evaluation of the Effects of Contingent Tangibles on Activity Engagement with an Extended Baseline Control

The results of Study 3A show that if activity engagement with a single activity occurs at a high rate when other activities are available, over time, engagement with that activity will decrease, become more variable, or both. As Mynatt et al. (1978) suggested, reinforcement in the reward condition may simply maintain temporarily the level of engagement and once the reward condition is discontinued the level of engagement will drop to approximately the levels observed in an extended baseline. Thus, the purpose of Study 3B is to extend on Mynatt et al. by conducting a within-subject comparison by including two sets of activities implementing a reward condition for a target activity in one set while keeping the other set in baseline.

Method

Subjects.

Subjects were three typically developing first-grade students, two males and one female. Jack and Kayla were recruited from a classroom in a local elementary school, and Dennis was recruited from the school’s afterschool program.

Design and procedure.

A combined ABA and multiple baseline design was used to compare the effects of a contingent tangible on activity engagement to an extended baseline condition. Subjects were exposed to two sets of materials; one set of materials in each 5-min session with two sessions per day. There were five age-appropriate leisure activities in each set. The activities in each set were placed within arm’s reach on the table in front of the subject. The presentation of sets was alternated each session.
For one set of toys, an ABA design was used to evaluate the effects of contingent tangibles for activity engagement. This manipulation was the same as Study 2. Baseline and Reward conditions were as described in the General Method. In the Reward condition a small edible item was delivered contingent on 30 s of continuous engagement with the target activity. The target activity was defined in the same manner as in Study 2. The second set of toys remained in an extended baseline as described in Study 3A.

**Interobserver agreement**

A second observer collected data during 29%, 43%, and 30% of sessions for Jack, Kayla, and Dennis, respectively. Percentage agreement scores were computed by dividing the session length into consecutive 10-s intervals. The smaller duration of observed responses was divided by the larger duration of observed responses in each interval and multiplied by 100. These values were averaged across the session. Average agreement scores were 99.0% (range, 94.5% to 100%), 99.2% (range, 96.74% to 100%), and 98.2% (range, 87.4% to 100%) for Jack, Kayla and Dennis, respectively.

**Results and Discussion**

Figures 5-3 and 5-4 display the results for all three subjects. For each subject, Set 1 is displayed on the top in an ABA design and Set 2 is displayed on the bottom in an extended baseline. Sessions are combined in two session blocks along the x-axis for Figures 5-3 and 5-4. In addition, due to the large number of data paths some of the effects may be obscured, so Figure 5-4 displays only the percentage of session engaged with the target activity. The results from all three subjects are consistent with the results reported by Mynatt et al. (1978). Table 5-1 shows the average percentage of session engaged with the Set 1 target item and Set 2 comparison item for each subject.
Condition lengths for comparison items were based on the condition lengths for corresponding Set 1 item. There was an average decrease in percentage of session engaged with Set 1 items from the first baseline condition to the second baseline of 45.65, 55.22, and 43.53 for Jack, Kayla, and Dennis, respectively. The average decrease in percentage of session engaged with the Set 2 comparison item was 50.56 and 54.79 for Jack and Kayla. For Dennis, there was an increase in percentage of session engaged with the Set 2 comparison item of 7.97, but when only the last 3 two-session blocks of each condition were averaged together there was a decrease of 28.22. These decreases are similar to those observed with the Set 1 target activity.

For all three subjects, there was an increase in the level of engagement with the target activity in the set of toys exposed to the contingent tangible reward condition, and once the condition was discontinued, the level of engagement dropped to approximately the levels observed in the other set of toys in the extended baseline as seen in Figure 5-3. This finding supports the notion that the presumed overjustification effect may be a satiation effect or at least an effect of repeated exposure, and that the reward condition may simply prolong the amount of engagement with the target item and, hence, provoke the repeated exposure effect.

In the current study the control condition was an extended baseline, consistent with the procedures used by Mynatt et al. (1978). In future research, a different control condition for this type of analysis would be to restrict access to all other activities. Based on the results of Study 1, this would increase time allocation to the target activity. If a decrease in activity engagement follows a phase of only the target activity available
then there would be further evidence to support that the decrement in responding is not due to the reward.
Figure 5-3. Percentage of session engaged with all items for Jack, Kayla, and Dennis for Set 1 and Set 2 (top and bottom panels, respectively). Sessions are combined into two-session blocks along the x-axis.
Figure 5-4. Percentage of session engaged with target items for Jack, Kayla, and Dennis for Set 1 and Set 2 (top and bottom panels, respectively). Sessions are combined into two-session blocks along the x-axis.
Table 5-1. Condition averages of percentage of session engaged with Set 1 target item and Set 2 comparison item for each subject. Condition lengths for Set 2 comparison items were based on condition lengths for corresponding Set 1 item. The top row for each subject displays the average for all sessions in each condition and the bottom row displays the average of the last 3 two-session blocks for each condition.

<table>
<thead>
<tr>
<th>Set #</th>
<th>BL1 1</th>
<th>BL1 2</th>
<th>Reward 1</th>
<th>Reward 2</th>
<th>BL 2 1</th>
<th>BL 2 2</th>
<th>BL 2 – BL 1 1</th>
<th>BL 2 – BL 1 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jack</td>
<td>46.82</td>
<td>52.81</td>
<td>66.97</td>
<td>33.92</td>
<td>1.17</td>
<td>2.25</td>
<td>-45.65</td>
<td>-50.56</td>
</tr>
<tr>
<td>Last 3</td>
<td>40.23</td>
<td>48.11</td>
<td>91.28</td>
<td>31.11</td>
<td>1.17</td>
<td>2.25</td>
<td>-39.57</td>
<td>-45.86</td>
</tr>
<tr>
<td>Kayla</td>
<td>55.22</td>
<td>87.96</td>
<td>99.29</td>
<td>69.42</td>
<td>0</td>
<td>33.17</td>
<td>-55.22</td>
<td>-54.79</td>
</tr>
<tr>
<td>Last 3</td>
<td>56.96</td>
<td>84.50</td>
<td>99.22</td>
<td>63.34</td>
<td>0</td>
<td>33.17</td>
<td>-56.96</td>
<td>-51.32</td>
</tr>
<tr>
<td>Dennis</td>
<td>47.00</td>
<td>34.84</td>
<td>88.20</td>
<td>89.25</td>
<td>3.47</td>
<td>42.81</td>
<td>-43.53</td>
<td>7.97</td>
</tr>
<tr>
<td>Last 3</td>
<td>66.28</td>
<td>33.22</td>
<td>91.90</td>
<td>99.80</td>
<td>5.78</td>
<td>5.01</td>
<td>-60.50</td>
<td>-28.22</td>
</tr>
</tbody>
</table>
CHAPTER 6
STUDY 4: REPLICATION OF DECI (1971)

The results from Studies 1-3 do not provide evidence to support the overjustification hypothesis. The results do, however, suggest that decrements in activity engagement under conditions in which a reward has been previously delivered for engagement could be due to repeated exposure to that particular activity. The discrepancy between the present results and the results reported by Deci (1971) could be due to the analytic methods used to evaluate the effect.

The subjects in Deci’s study were 24 undergraduate psychology students. Deci used a group design with 12 subjects in the experimental group and 12 subjects in the control group. The experimental group was exposed to an ABA design. The experimental phase involved the delivery of one dollar contingent on successful puzzle completion. The control group was exposed to three baseline conditions. For both groups, each condition was a single hour-long session. The subjects were asked to complete puzzle configurations that matched a drawing presented by the experimenter. Four puzzle configurations were presented in each phase. The dependent measure was the number of seconds that the subject was working on the puzzles during an 8-min break occurring in the middle of each phase (i.e., after the second configuration of that phase had been completed). During the break, the experimenter left the room and the behavior of the subject was being monitored through a one-way window.

Deci subtracted the number of seconds that the subject was engaged with the puzzle during the third break from the seconds engaged in the first break. These values were compared across groups. Based on a high alpha value of .10, Deci concluded that there was a significant difference between groups and that these results support the
overjustification hypothesis. The difference was not significant at the .05 level, although that level is commonly used in statistical tests.

Besides the high alpha value, there were several other limitations with the study. First, there were no individual or repeated measures of behavior. The individual differences could not be seen based on the averages and the amount of engagement during the sessions in which the experimenter was present was not reported. Second, the amount of engagement with other available activities was not reported. Third, the number of times each subject contacted the extrinsic reward was not reported. At the extreme, it is possible that a subject never completed a puzzle and, therefore, never contacted the extrinsic reward. The overjustification hypothesis, by definition, requires that a subject contact an extrinsic reward, so it would be inappropriate to include the data from a subject who did not contact reward in the experimental group average. In 1979, MaWhinney reanalyzed the data obtained from Deci’s (1971) Experiment I. Based on MaWhinney’s analysis, only 4 of the 24 total subjects showed a decrease in engagement with the puzzle from break 1 to break 3. In addition, individual patterns did not necessarily indicate a high level of intrinsic interest in that seven subjects in each group engaged with the puzzle for less than 240 s during break 1. Still, how often the subjects completed the puzzle was not reported and how often the subjects in the experimental group earned the money was not reported.

The purpose of Study 4 was to replicate the methods used by Deci and address the limitations by including individual measures of puzzle engagement during sessions and breaks and by reporting how often the subjects completed the puzzle and earned the reward.
Method

Subjects and Setting

Subjects were 24 undergraduate students in psychology recruited via the General Psychology (PSY2012) Research Participation Pool. There were 12 subjects in the experimental group and 12 in the control group. All Study 4 subjects were enrolled in, and compensated with, PSY2012 class credit in accordance with the class protocol (i.e., one credit per every half-hour spent in a study). Sessions for Study 4 took place at a table in an experimental room in the Psychology building on the University of Florida campus. The experimenter measured engagement with available activities. Engagement was defined as touching the item or items corresponding with that activity. Engagement also included orienting toward and looking directly at the materials. Subjects could be engaged with multiple activities at the same time.

Design and Procedure

A combined ABA and group design was used to evaluate the effects of money contingent on puzzle completion. One group was exposed to all three phases as described below; the second group was exposed to three baseline sessions.

Baseline. The subjects were asked to reproduce a shape printed on a piece of paper using block puzzle pieces. There were four puzzle configurations in this condition. The subjects had 13 min to complete the puzzle before the next configuration was presented. If they completed the puzzle before the 13 min was over they were presented with the next configuration. After the second configuration in this phase there was a break. Procedures for the break are described below.
**Experimental Phase.** Procedures for this phase were the same as baseline with one exception; if the subject completed the configuration within 13 min they were given one dollar.

**Break.** There were three breaks during the entire study, one in the middle of each phase of the study. The first occurred after the second puzzle configuration, the second occurred after the sixth puzzle configuration, and the third occurred after the tenth puzzle configuration. Before each break the experimenter said, “Excuse me, I have to leave the room for a few minutes to determine the next configuration. You may do whatever you like while I am gone.” On the table next to the puzzle were a newspaper and a few up-to-date magazines. The subjects were observed through a one-way mirror and data were collected on puzzle engagement, as well as engagement with other items that were available.

**Interobserver Agreement.**

Interobserver agreement (IOA) was collected by a second observer simultaneously, but independently, collecting data on 23% of all breaks. Total agreement scores were calculated by dividing the smaller duration of observed engagement during each break by the larger duration of engagement and multiplying by 100. Average agreement scores for all subjects were 99.5% (range, 97.5% to 100%), 98.9% (range, 95.5% to 100%), and 99.2% (range, 95.48% to 100) for breaks 1, 2, and 3, respectively.

**Results and Discussion**

Table 6-1 shows the number of seconds engaged with the puzzle during each break for all subjects and the means for both groups. For both groups, the number in parenthesis indicates the number of puzzles completed during the corresponding phase.
of the study for that group. If there is no number, then there were no puzzles completed during that phase for that subject. For example, subject 5 in the experimental group completed zero puzzles during the first phase, one puzzle during the second phase (experimental phase), thus earning the money for that configuration, and one puzzle during the third phase with no money delivered following completion of that configuration. Interestingly, of the 288 configurations that were presented to both of the groups, only 22 were completed in less than 13 min and, furthermore, only four subjects in the experimental group actually earned money (subjects 5, 6, 7 and 10). This means that eight subjects never contacted the reward, so any changes in puzzle engagement could not be attributed to receiving the reward for these subjects. In addition, subjects 3, 9, and 11 showed a decrease in puzzle engagement from break 1 to break 3, but did not contact the reward during phase 2 of the study. Contrary to Deci, both of the groups in the current study showed an average decrease from break 1 to break 3, with the control group showing a larger decrease.

Figures 6-1 and 6-2 show the individual times engaged with the puzzle for the experimental and control groups, respectively. There were three general patterns of individual results in the experimental group: 1) five subjects showed a decrease in the amount of puzzle engagement from break 1 to break 2 and a further decrease from break 2 to break 3 (subjects 3, 5, 9, 10 and 11), 2) two subjects showed an increase from break 1 to break 3 (subjects 1 and 8), and 3) five subjects did not engage with the puzzles during any of the breaks (subjects 2, 4, 6 and 7). For the experimental group, when the amount of puzzle engagement during each break is averaged across subjects, the results are consistent with those reported by Deci. In the current study, individual
patterns of engagement did not show a high level of intrinsic interest, meaning that there was little engagement with the puzzle during the breaks when the subject was left alone. Several of the subjects across both groups had less than 240 s of engagement during break 1, and eight subjects did not engage with the puzzle at all during the entire study. Based on MaWhinney’s (1979) analysis, more than half of the subjects in the original study engaged with the puzzle for less than 240 s during the first break. Future evaluations should include a task with a higher level of engagement during the first break, or implement an exclusion criterion for subjects with a low level of engagement during the first break.
<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Control Group</th>
<th>Break 1</th>
<th>Break 2</th>
<th>Break 3</th>
<th>Break 3 – Break 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>85.92</td>
<td>85.92</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>479.04</td>
<td>437.76</td>
<td>408.48</td>
<td>-70.56</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>472.99</td>
<td>386.02 (1)</td>
<td>183.98 (1)</td>
<td>-289.01</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0 (1)</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0 (1)</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>70.99</td>
<td>0</td>
<td>81.98</td>
<td>10.99</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>478.99</td>
<td>480.00</td>
<td>28.80</td>
<td>-450.19</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>166.99 (1)</td>
<td>0 (1)</td>
<td>0 (1)</td>
<td>-166.99</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>470.40</td>
<td>427.20</td>
<td>379.20 (1)</td>
<td>-91.20</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>472.80</td>
<td>425.33</td>
<td>354.58</td>
<td>-118.22</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>470.40</td>
<td>480.00</td>
<td>460.80</td>
<td>-9.60</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>474.00</td>
<td>316.99</td>
<td>26.02</td>
<td>-447.98</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0 (1)</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>459.98 (1)</td>
<td>392.02</td>
<td>330.00</td>
<td>-129.98</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0 (1)</td>
<td>0 (1)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>345.98 (1)</td>
<td>0</td>
<td>0</td>
<td>-345.98</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>272.02</td>
<td>0</td>
<td>0</td>
<td>-272.02</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>480.00 (2)</td>
<td>480.00 (2)</td>
<td>480.00 (1)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>201.60</td>
<td>338.02</td>
<td>33.02</td>
<td>-168.58</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>459.02</td>
<td>75.98 (1)</td>
<td>0 (3)</td>
<td>-459.02</td>
<td></td>
</tr>
</tbody>
</table>

Mean 178.28          144.25          97.36          -80.92

Mean 302.98          209.03          140.37          -162.62
Figure 6-1. Number of seconds engaged with the puzzle during each break for individuals in the experimental group.
Figure 6-2. Number of seconds engaged with the puzzle during each break for individuals in the control group.
CHAPTER 7
GENERAL DISCUSSION

The results of Studies 1-3 do not provide evidence to support the overjustification hypothesis. Rather, the results demonstrate that extended exposure to the preferred activities may account for a decrease in activity engagement from the first to the second baseline condition, when a decrease is observed. In Study 1, none of the four subjects showed a sustained detrimental effect of contingent rewards on levels of activity engagement and, with the exception of James, the subjects were engaged with the available activity over 98% of the session for all sessions. James showed a slight decrease in activity engagement from the first to the second baseline, but remained engaged over 70% of the session for all sessions in the second baseline. The overjustification effect was not observed when only one item was available and levels of engagement remained high and stable for all four subjects.

Contrary to the results of Study 1, in Study 2, there was a decrease in activity engagement from the first to the second baseline for both subjects. For Sam, a drastic decrease to 0% engagement with the target item, stickers, occurred after the second session of the second baseline. For Angel, the same decrease occurred immediately in the first session of the second baseline. Both subjects were informed of the contingency at the beginning of each session and thus were told at the beginning of the second baseline that they would no longer receive edibles for playing with the target item. These effects are in alignment with the overjustification hypothesis and could provide support for the overjustification effect. However, due to the number of sessions of exposure to the same activities, effects that look similar to the overjustification effect could, instead, be a result of repeated exposure. This possibility was evaluated in Study
3. It is also important to note that in Studies 1 and 2 the effects of contingent rewards differed depending on whether there are other items available with which the child could engage. In previous research, this aspect of the experimental sessions has varied across studies, and in some cases was not even specified (e.g., Anderson et al., 1976). Based on these findings, in future research it will be important to note whether other activity choices are available to the children during the experimental sessions. In addition, within-subject comparisons of one versus many items available should be conducted.

The results of Studies 3A and 3B provide evidence to suggest that the putative overjustification effect may in fact be a result of extended repeated exposure to the activities. Study 3A demonstrates that extended exposure to one set of activities will result in decreases in engagement with a single activity, if engagement with that activity was initially high. The results from all four subjects showed this effect. The results of Study 3B demonstrate that if a contingent reward condition is included and subsequently removed for a target activity in one set of activities, over time, engagement with the target activity will match the levels observed in a set of activities that remained in baseline. These findings are consistent with those reported by Mynatt et al. (1978). Future research should evaluate the effects of other control conditions that would increase time allocation to the target activity without the use of a reward condition, for example, a condition with only the target item available.

Although these results provide support for the notion that the overjustification effect is simply a result of extended exposure to the activities, it is still not conclusive. If the effects of repeated exposure to the activities is minimized or eliminated and if the
overjustification effect is in fact a result of repeated exposure, then activity engagement should remain at high levels throughout the assessment. If the overjustification effect is not an effect of repeated exposure, then when these effects are minimized, activity engagement should remain high for the activities that remain in baseline, but drop below baseline levels for those that have been exposed to a contingent reward.

One possible explanation for the effect of repeated exposure is satiation. There has been ample research showing that varied stimulus presentation will reduce the effects of satiation (e.g., Dunlap & Koegel, 1980; Egel, 1980,1981; Bowman, Piazza, Fisher, Hagopian, & Kogan, 1997; Koehler, Iwata, Roscoe, Rolider, & O’Steen, 2005). Dunlap and Koegel (1980) evaluated the differential effectiveness of two methods of task presentation. In one condition a single task was presented throughout the training session; in the other condition, the target task was interspersed among a variety of other tasks during a single training session. The results showed higher and more stable levels of correct responding during the varied task condition for both subjects.

Egel (1980) found that the presentation of varied reinforcers produced higher rates of bar pressing for ten individuals with autism. Subsequently, Egel (1981) replicated those findings with a population of children diagnosed with autism spectrum disorder or other developmental disabilities in a classroom setting. There were decreasing trends in correct responding and on-task behavior for all three subjects during the constant reinforcer condition; however, correct responding returned to high levels and maintained for 125 trials during the varied reinforcement condition. Varying the presentation of the reinforcers minimized the effects of satiation. Thus, possible future direction would be to minimize the effects of extended exposure to the activities,
perhaps by varying the presentation of sets of activities across sessions, and evaluate the effects of contingent rewards on a target activity in each set.

The terms reward and reinforcer have been used interchangeably throughout the literature (Cameron & Pierce, 2004). The term reward was used throughout the current study because the effectiveness of the edibles as reinforcers was not directly assessed. However, with the exception of Study 1, there was an increase in the level of activity engagement in all reward conditions for all subjects, so it is probably safe to say the edibles were reinforcers given that the baseline lengths were staggered, overall providing a context similar to a nonconcurrent multiple baseline design. In Study 1, all subjects were engaged with the items for 100% of the session in the first baseline therefore, due to a ceiling effect, increases could not be observed. Bright and Penrod (2009) directly evaluated the effects of verbal praise, non-reinforcing stimuli, and reinforcers as identified by a reinforcer assessment and found no evidence to support the overjustification hypothesis in any of the conditions.

Continued research on possible negative effects of reward is warranted. The results of Study 4, a replication of one of the most commonly cited studies, do not provide evidence to support the overjustification hypothesis. Both groups in the current replication showed an average decrease in puzzle engagement from break 1 to break 3 and analysis of the individual data showed a variety of patterns of responding in both groups. Future research should continue to identify the conditions under which overjustification-like effects will and will not occur. Some variations may include evaluations with different types of rewards, (e.g., praise instead of edibles) and different contingencies, (e.g., reward contingent on task completion instead of engagement).
Finally, it is possible that different populations of children and adults would respond differently to the conditions of this study. The children in this study were all typically developing. In some preliminary follow-up work, we are finding that some children with autism are less sensitive to repeated exposure following a reward condition. Thus, we are in the process of evaluating variations of overjustification research in our current studies.


BIOGRAPHICAL SKETCH

Kerri P. Peters graduated from the University of Florida with a Bachelor of Science in psychology and a minor in education in 2004. She began working in the field of behavior analysis in late 2002 as an in-home therapist for a 3 year-old girl diagnosed with autism. From there, her interests in the field grew. She enrolled in all of the classes in the area that were available. In addition to classes, she gained research experience by volunteering in multiple labs. These experiences prepared her for the next step in her academic career: a Master of Science from the Department of Behavior Analysis at the University of North Texas. She continued her graduate studies at the University of Florida in 2007. Since beginning graduate school, Kerri has had the opportunity to work a variety of settings with multiple populations, including individuals with intellectual disabilities, parents at-risk for maltreatment, foster parents and children, and children with behavior disorders. Following graduation, Kerri intends to pursue a career in applied behavior analysis, with her goal being to teach and continue to conduct research in clinical settings.