Abstract

Might extrinsic rewards be used to encourage the development of intrinsic motivation and thus the cultivation of character (Watts, Fullard & Peterson, 2021, 2022; Berkowitz, 2022)? Research to date suggests that rewards undermine intrinsic motivation (Ryan & Deci, 2017; Warneken & Tomasello, 2008). But these studies have focussed only on those who are *already* intrinsically motivated. What about those who are unmotivated? Might rewards be used to *kickstart* intrinsic motivation? The proposed research aims to address this question within the context of altruism. Preliminary screening will identify a sample of children who are unmotivated to engage in helping. These children will proceed to an experiment that is much the same as Warneken and Tomasello's (2008). A mixed-design ANOVA will be used to assess the effects of condition (Reward/Neutral) and time (1/2) on intrinsic motivation to engage in helping, measured via the number of free-choice trials in which a participant helps. Findings and implications will be disseminated to psychologists, character educationists, and teachers.

Keywords: extrinsic rewards, intrinsic motivation, character, altruism, free-choice

Extrinsic Rewards and the Inception of Intrinsic Motivation

Intrinsic motivation refers to the doing of an activity for its inherent satisfactions (e.g., feelings of interest or enjoyment), while *extrinsic* motivation refers to the doing of an activity for some separable consequence (e.g., social approval or an external reward) (Ryan & Deci, 2000; 2017, p. 14). Intrinsic motivation is essential for Aristotelian character education (e.g., Jubilee Centre, 2022).

First, character consists in *virtues*, and it is essential to virtue that people engage in virtuous activity *because it is virtuous* (Curren, 2014). Suppose that you were in need, and I helped you. But I did so not for the inherent satisfaction of helping others, but rather because I was being paid. My action would not count as virtuous. Truly virtuous behaviour must be intrinsically motivated. It may also transpire that the only truly intrinsically motivating behaviour *is* virtuous behaviour. According to Aristotle, the only intrinsically desirable thing in itself is *flourishing*, which is constituted by the ongoing realisation of virtues (Kristjánsson, 2017, p. 25). Thus, intrinsic motivation is at least necessary, but perhaps also sufficient for virtue.

But intrinsic motivation is not without its problems. First, according to operant psychology, all voluntary behaviour is *extrinsically* motivated; "intrinsic motivation" simply refers to baseline behaviours for which extrinsic motivators have not yet been identified (see e.g., Carton, 1996). But if all behaviours were extrinsically motivated, this would imply that "intrinsic" and extrinsic motivation are *additive*: after adding and then removing extrinsic motivators (e.g., rewards), behaviour would return to baseline (i.e., intrinsic motivation would be unaffected). This is not the case. Deci (1971) found that behaviour in fact went *below* baseline. This undermining effect suggests that intrinsic and extrinsic motivation are distinct phenomena.

Intrinsic motivation raises a difficult problem for character education in particular. If virtue must emanate from *within* the self, then how can it be cultivated from *outside*? Considering just the intellectual meta-virtue of *phronesis*, this question manifests as Peters' (1981, p. 52) famous "paradox of moral education": How can the "palace of Reason" be entered uncritically through the "courtyard of Habit and Tradition"? More generally, and in terms of motivation, what role can extrinsic motivators play in the internalisation of virtue? This is a question of recent debate. Watts, Fullard and Peterson (henceforth "WFP") (2021) claim that rewards can be useful in encouraging the internalisation of virtue. Berkowitz (2022) disagrees, claiming that they are at best ineffective, but often undermine the internalisation process. In response, WFP (2022) acknowledge that more empirical research is needed. But where exactly is the gap in the research?

Experiments examining the effect of rewards on intrinsic motivation typically take the following form. Participants engage in an interesting (i.e., intrinsically motivating) activity and are assigned to one of two groups: a reward group or a control group. The reward group receives rewards for engaging in the activity; the control group receives nothing. This is followed by a free-choice period in which participants have the opportunity to engage with the activity. Engagement is then compared between the two groups. In 1999, Deci et al. conducted a meta-analysis of 128 such studies spanning a 25-year period. The analysis shows that rewards have a negative effect on intrinsic motivation (-0.24). This effect is accounted for by *tangible* (as opposed to *verbal*) rewards (-0.34), which in turn is accounted for by *expected* (as opposed to *unexpected*) rewards (-0.36), especially *engagement*-, *completion*- and *performance-contingent* rewards (-0.40, -0.44 and -0.28, respectively) (Deci et al. 1999; Ryan & Deci, 2017).

Regarding character development in particular, several studies have shown that rewards undermine intrinsic motivation to engage in virtuous behaviour (e.g., Fabes et al.,

1989; Warneken & Tomasello, 2008). In Warneken and Tomasello's (henceforth "W&T") (2008) study, 20-month-old children were rewarded for engaging in helping behaviour. During the subsequent free-choice period, these children were found to engage in helping *less* than those who had not received a reward. W&T (2008) concluded that rewards undermine children's intrinsic motivation to engage in helping behaviours.

These results pose a serious challenge to character education, as rewards are ubiquitous in most educational settings. Does this mean that character may struggle to develop in such settings? If so, how should character educationists respond? Redouble their commitment to intrinsic motivation and risk alienating many schools (e.g., Berkowitz, 2022)? Or try to play down intrinsic motivation and risk reducing virtue to mere expediency (e.g., WFP, 2021)? There might yet be way out of the dilemma.

Studies typically focus on activities that are intrinsically motivated, and proceed to show that this motivation is undermined by external rewards. But intrinsic motivation does not inhere in activities in isolation. Indeed, activities require *actors*, and different individuals will find different activities intrinsically motivating (Ryan & Deci, 2017, p. 117). Thus, part of the challenge of character education is to foster intrinsic motivation in *all* students, despite their individual differences. For those who are already intrinsically motivated, research shows that rewards can have an undermining effect (Ryan & Deci, 2017). But what about those who are *not* intrinsically motivated?

To illustrate, W&T's (2008) final sample consisted of 36 children. But the original sample consisted of 54 children. 18 children (33%) were excluded because during the treatment phase they failed to help in at least five trials. W&T may have reasoned as follows: if a child fails to meet this criterion, then they are not motivated to help; so, there would be no point investigating the effect of rewards on their intrinsic motivation, because they have not got any. The proposed research will challenge this assumption by investigating whether

rewards might facilitate the *inception* of intrinsic motivation in those who have none. The gap in the research thus consists in the participants themselves. Whereas W&T (2008) excluded from their sample those who were *not* motivated to help, these children will comprise the sample in the proposed study.

Indeed, it is these children who should be the focus of character education, as they have the most to gain. In the Jubilee Centre's (2022) neo-Aristotelian model of moral development, these children would be on the lower trajectory – "Plan B" – which is largely guided by extrinsic motivators. Yet the model suggests with an upward arrow that Plan B children might ascend to the upper trajectory – "Plan A". The proposed study will consider whether this ascension might be initiated via external rewards.

A final point of clarification. Much research has already considered the effect of external factors (e.g., rewards) on *extrinsic* motivation, specifically, the degree to which it is integrated and internalised (see Ryan & Deci, 2017). That is not the focus of the proposed study. No matter how internalised and integrated an extrinsic motivator may be, it remains an *extrinsic* motivator. The focus here is whether rewards might facilitate the inception of a wholly different kind of motivation – *intrinsic* motivation.

Research Questions

The overarching and most general question that the proposed research aims to address is:

[1] Can extrinsic motivators support character development?

The specific extrinsic motivator in question is that of external rewards (viz. completion-dependent tangible rewards), and the development of character is taken to consist in the internalisation of virtues. Thus, a more specific question:

[2] Can external rewards support the internalisation of virtues?

A virtue is internalised as the motivation for its associated behaviours shifts from extrinsic to intrinsic. For example, the virtue of altruism is internalised as the motivation to help others shifts from extrinsic to intrinsic. Altruism will be the virtue of focus. Thus:

[3] Can external rewards support intrinsic motivation to help others?

The proposed research concerns specifically the *inception* of intrinsic motivation in children who as yet have none. Thus, finally:

[4] Can external rewards *bring about* intrinsic motivation to help others?

This is the specific question that the proposed research seeks to address; the hypothesis being that external rewards can indeed have such an effect. Note that this hypothesis implies affirmative answers to all prior questions. Thus, if external rewards can bring about intrinsic motivation to help others [4], then extrinsic motivators could indeed support the development of character [1]. (Note that if the proposed research finds no evidence in favour of the hypothesis, this would not imply negative answers to the above questions.)

What reason is there to suppose that the hypothesis might be true? There is no empirical evidence; hence the gap in the research. But there is a theoretical reason. In their seminal article, Lepper, Green and Nisbett (1973) note that the appeal of some activities only becomes apparent after prolonged engagement with it. This suggests that some people might not be intrinsically motivated to engage in an activity simply because they have not spent enough time doing it. A child, for example, might not be intrinsically motivated to help others simply because she has not spent enough time helping others. But if an extrinsic motivator – a reward, say – is used to encourage engagement, this might kickstart the development of intrinsic motivation. This would be a reversal of the undermining effect. Whereas in most studies interest in an activity is undermined by an external reward, here the external reward would be undermined by the appeal of the activity. In attributional terms, if an external reward can prompt an external shift in *perceived locus of causality* (PLOC) in those who are

intrinsically motivated (Ryan & Deci, 2017, p. 127), then there is no reason to deny that the inherent satisfaction of an activity might prompt an internal shift in PLOC in those who are extrinsically motivated (e.g., by a reward). It is this latter proposition that the proposed research will put to the test.

Methodology

The proposed study is similar to W&T's (2008) study, with two major differences. First and foremost, the final sample will consist of children who are *not* motivated to help. This will necessitate a preliminary screening phase to identify such children. Second, helping trials will consist of one 30 second phase in which the experimenter focuses solely on the dropped object (never on the child). These differences are explained in more detail below.

Participants

G*Power was used to calculate a minimum sample size of 36 children (F test; ANOVA: repeated measures, between factors; $\alpha = .05$; power = .95; effect size $f = .5$). This number was increased to 50 to account for attrition. But this is the size of the sample of *unmotivated* children. First, these children must be identified via preliminary screening. In W&T's (2008) study, one in three children were unmotivated (18 out of 54). This suggests that the preliminary sample in the proposed study should consist of at least 150 children. But more children will be recruited and tested until 50 unmotivated children have been identified. 150 children (75 boys and 75 girls) who are around 20 months of age will be recruited from the birth register of a city in the United Kingdom (e.g., Birmingham), come from mixed socioeconomic backgrounds, and have English speaking parents. None of the children will have previously participated in a study on helping.

Research Design

Children will be tested in an experimental room, accompanied by a parent who remains passive. Testing will be conducted by two experimenters (E1 and E2) who are

unaware of the hypothesis. E1 will administer the helping tasks; E2 will operate remote-control cameras from an adjacent room and rearrange the setup between tasks (W&T, 2008).

E1 will be sitting at a desk performing activities such as writing a letter with a pen. During the activity, he will drop an object (e.g., the pen) onto the floor and unsuccessfully reach for it with an outstretched arm by bending over the desk and making sounds of effort (W&T, 2008). The child could help by picking up the object and handing it back to E1. Each trial will last up to 30 seconds.

In W&T's (2008) study, trials were split into two 15-second phases. In the first phase, E1 focussed solely on the object. If the child did not help during this phase, then, during the second phase, E1 named the object and alternated gaze from object to child. This latter phase, however, seems to violate the free-choice paradigm of intrinsic motivation (see Measures), according to which participants should have no evaluative reason to engage with the target activity. To restore the free-choice paradigm, the second phase will be omitted from the proposed study. Instead, E1 will focus solely on the object for the full 30 seconds.

There will be six helping tasks involving out-of-reach objects, which will differ only in the kind of object that is dropped. This is to mitigate boredom with the target activity. The tasks will be presented in three blocks of three tasks, one block during each phase: preliminary screening (objects: marker, paper balls, clips), treatment phase (objects: pen, plates, clothes pegs), and test phase (objects: eraser, coin, keys). The order of blocks and the order of tasks will be counterbalanced across participants.

Measures

Motivation (intrinsic or extrinsic) is operationalised via observation of the amount of time that a participant spends working with a target activity – that is, *behavioural persistence*. Intrinsic motivation in particular is operationalised via behavioural persistence during a *free choice-period*, when participants are left alone, are free to choose what to do, and have no

external incentive of evaluative reason to persist. This is Deci's (1971) *free-choice paradigm* of intrinsic motivation (see also Ryan & Deci, 2017, p. 126).

Thus, following W&T (2008), a participant's intrinsic motivation to help will be measured via the number of trials in which they freely choose to help, with a minimum threshold of five trials. The threshold is imposed to ensure that the behaviour is persistent. If a child fails to meet this criterion, they will be considered *unmotivated* to help. Given the interpersonal nature of helping, it is impossible for participants to be alone during free-choice periods (as per the paradigm). But accompanying parents will remain passive, and E1 will only ever focus on the object (see Research Design).

Free-choice behavioural persistence is a *face* valid measure of intrinsic motivation, which is by definition autonomous (Ryan & Deci, 2017, p. 14). *Criterion* validity – the extent to which a measure agrees with an external criterion – does not apply, because there is no external criterion. Intrinsic motivation is *intrinsic*, after all; it is a theoretical construct, which is not directly measurable. Whether free-choice behavioural persistence is *construct* valid as a measure of intrinsic motivation is debatable. Operant psychologists might argue that intrinsic motivation is a spurious construct; self-determination theorists would disagree (see above).

Procedure

Preliminary Screening

The goal of the preliminary phase is to identify of subsample of children who are unmotivated to help. Like W&T's (2008) treatment phase, the preliminary phase will consist of at most four tasks with 3 trials each (12 trials total). But, unlike W&T (2008), these tasks will take place during a free-choice period (see Measures). If a child fails to help in five trials, they will be considered unmotivated to help. Children will be tested until at least 50 unmotivated children have been identified. These children will proceed to the experiment

(i.e., the treatment and test phases). To mitigate against participants becoming bored with helping, the experiment will take place one week later.

Treatment Phase

During an initial warm-up, E2 will introduce the children to an apparatus that has proven to be an effective reward for children (W&T, 2007). When a cube is thrown through an opening in the apparatus, it will slide down a transparent tube into a box and create a jingling sound. The cubes needed to operate the apparatus will be used as rewards for the children.

For the treatment phase, children will be randomly assigned to one of two conditions: Reward or Neutral. In the Reward condition, E1 will reward children's helping with a toy cube. He will offer the cube while reaching for the dropped object and hand it to the child if she helps, stressing the exchange by saying: "For this, you will get a cube." Thus, the cube is a completion-dependent tangible reward. In the Neutral condition, E1 will not address the child in any way. This condition will serve as a baseline for children's intrinsic motivation to help. As in preliminary screening, the treatment phase will consist of four tasks with 3 trials each (12 trials total).

Test Phase

The test phase will be the same as W&T's (2008) test phase. All children will be presented with three helping tasks of 3 trails each (9 trials in total). The helping tasks will be the same as those in the treatment phase, except that different objects will be used. This time, E1 will offer no response to any of the children. In other words, these tasks will take place during a free-choice period. As in W&T's (2008) study, children will be given the opportunity to play with distractor toys. The distractor toys will be different instruments, installed on a box. When children press one of the buttons, the instruments will play different melodies and sounds and light up. There will be three such distractor toys (violin, trumpet, guitar), one for

each of the helping tasks administered during the test phase (in counterbalanced order). Before each task, E2 will bring in one of the distractor toys and install it on the floor opposite E1's desk. Thus, children will have to stop playing the instruments and leave the distractor in order to help. This is to mitigate against a near ceiling effect in helping (W&T, 2008).

Observation and Coding

All sessions will be videotaped and coded by myself. The test phase will be coded first. A random sample of 25% of children will be independently coded by a research assistant to assess interrater reliability. Following W&T (2008), we will code:

(1) whether the child performed the helping behaviour, defined as picking up the object and handing it to E1;

(2) latencies of helping, occurring from the moment that E1 is reaching for the object until the moment the child puts it in E1's hand; and

(3) whether the child plays with the distractor toy during the trial.

Analysis Plan

Preliminary analyses will be conducted to determine whether there was an effect of gender, task, or task order on any of these measures (1–3, above). If not, further analyses will be collapsed across these factors. The average time taken to help will be calculated and compared across conditions.

The dependent variable in this study is intrinsic motivation to engage in helping, measured via the number of free-choice trials in which a participant helps. There are two independent variables: *condition* (Reward and Neutral) and *time* (time 1 and time 2). Condition is a between-participants variable, while time is a within-participants (repeated measures) variable. Thus, a mixed-design analysis of variance (ANOVA) will be used to assess the effects of condition and time on intrinsic motivation to engage in helping. ANOVA assumptions of normality, homogeneity, and sphericity will be checked.

The hypothesis would predict, and thus be evidenced by, the following results:

- a main effect of condition on intrinsic motivation to help;
- no overall effect of time;
- an interaction effect between condition and time, such that when observing the effect in the Reward condition, there are significant increases in intrinsic motivation from time 1 to time 2;
- no significant change in the Neutral condition from time 1 to time 2.

The Reliable Change Index (Zahra et al., 2016) will be calculated to assess whether any observed change is accounted for by test-retest unreliability.

Ethics

Given the strong connection between intrinsic motivation and wellbeing (Ryan, Curren & Deci, 2013), it reasonable to question the ethics of studies that result in participants *losing* intrinsic motivation, especially for a virtuous or prosocial behaviour such as helping (e.g., W&T, 2008). In the proposed study, however, after preliminary screening, the sample will consist of children who are *not* intrinsically motivated to help. These children could only stand to make gains in intrinsic motivation. To ensure equality of opportunity, those allocated to the Neutral condition will be given access to the treatment after the study.

Informed parental consent will be acquired for each participant, and the study will be approved by an ethics committee before data collection. Participants will be treated according to the British Psychological Society Code of human Research Ethics and will be allowed to withdraw, or be withdrawn by their parents, at any point.

Implications and Dissemination Plan

The proposed research would be of interest to developmental and educational psychologists, especially self-determination theorists. Research to date shows that intrinsic motivation is undermined by external rewards, and that extrinsic motivation can be more or

less integrated and internalised (Ryan & Deci, 2017). But none has sought to determine the effect of rewards on the intrinsic motivation of those who are unmotivated. The proposed research would fill this gap. It would be submitted in the first instance for publication in the *Journal of Personality and Social Psychology*. This journal was chosen because among the journals whose scope includes the proposed research, this one has the highest *h*-index (Scimago Journal & Country Rank, 2023).

The research would help settle an enduring debate in character education. Most recently, WFP (2021, 2022) suggest that rewards might have a positive effect on intrinsic motivation; Berkowitz (2022) disagrees. Research to date supports Berkowitz's position: external rewards undermine intrinsic motivation (Ryan & Deci, 2017). But if the proposed research found evidence that rewards can facilitate the inception of intrinsic motivation, this would settle the debate in favour of WFP (2021, 2022). If no such evidence was found, however, this would further strengthen Berkowitz's (2022) position. These implications could be disseminated to character educationists via presentation at the Jubilee Centre's annual conference.

Findings could inform reward policies in schools. If the hypothesis is supported, this would suggest that teachers might use external rewards to kickstart intrinsic motivation in unmotivated students. Note, however, that in many cases, where students already have some intrinsic motivation, the use of rewards may remain counterproductive. If the hypothesis is unsupported, this would suggest that there is no role for external rewards in the development of intrinsic motivation, and that teachers should look to other strategies (see e.g., Berkowitz, 2022). These implications and caveats could be communicated to teachers via presentation at an education conference (e.g., Warwick Education Studies Conference).

The proposed research may also give rise to further interesting questions. Suppose that external rewards were found to facilitate the inception of intrinsic motivation. The

proposed explanation is that the inherent satisfaction of the activity has prompted an internal shift in PLOC. This suggests that for maximum effect: (1) rewards should be *small*, just big enough to motivate sufficient engagement in the target activity; and (2) once sufficient engagement is achieved, the reward contingency should be withdrawn. These corollaries (1–2) lend themselves readily to empirical testing. The theory might be further corroborated by rerunning the experiment but with an inherently *frustrating* activity. In this case, the theory would predict that no matter how much engagement is motivated via external rewards, there would be no increase in intrinsic motivation to engage with the activity.

References

Berkowitz, M.W. (2022). Introducing the complexity of character education: A review of *Understanding character education: Approaches, applications and issues*. *Journal of Moral Education*, 51(4), pp.589–594.

Carton, J.S. (1996). The differential effects of tangible rewards and praise on intrinsic motivation: A comparison of cognitive evaluation theory and operant theory. *Behavior Analyst*, 19(2), pp.237–255.

Curren, R. (2014). Motivational aspects of moral learning and progress. *Journal of Moral Education*, 43(4), pp.484–499.

Deci, E.L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18(1), pp.105–115.

Deci, E.L., Koestner, R. and Ryan, R.M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), pp.627–668.

Fabes, R.A., Fultz, J., Eisenberg, N., May-Plumlee, T. and Christopher, F.S. (1989). Effects of rewards on children's prosocial motivation: A socialization study. *Developmental Psychology*, 25(4), pp.509–515.

Jubilee Centre for Character and Virtues (2022). *Framework for character education in schools*. [online] Birmingham: University of Birmingham. Available at: https://www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/character-education/Framework%20for%20Character%20Education.pdf [Accessed 24 Oct. 2022].

Kristjánsson, K. (2017). *Aristotelian character education*. Abingdon: Routledge.

Lepper, M.R., Greene, D. and Nisbett, R.E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the 'overjustification' hypothesis. *Journal of Personality and Social Psychology,* 28(1), pp.129–137.

Peters, R.S. (1981). Reason and habit: The paradox of moral education. In: *Moral development and moral education*. London: Routledge.

Ryan, R.M., Curren, R.R. and Deci, E.L. (2013). What humans need: Flourishing in Aristotelian philosophy and self-determination theory. In: A.S. Waterman, ed., *The best within us: Positive psychology perspectives on eudaimonia*. American Psychological Association, pp.57–75.

Ryan, R.M. and Deci, E.L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), pp.54–67.

Ryan, R.M. and Deci, E.L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford Press.

Scimago Journal & Country Rank (2023). *Journal Rankings on Psychology*. [online] www.scimagojr.com. Available at: https://www.scimagojr.com/journalrank.php?area=3200 [Accessed 13 May 2023].

Warneken, F. and Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy,* [online] 11(3), pp.271–294.

Warneken, F. and Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental Psychology*, 44(6), pp.1785–1788.

Watts, P., Fullard, M. and Peterson, A. (2021). *Understanding character education: Approaches, applications, and issues*. London, England: Open University Press, McGraw-Hill Education.

Watts, P., Fullard, M. and Peterson, A. (2022). Response to Berkowitz' extended book review: 'Introducing the complexity of character education: A review of

Understanding character education: Approaches, applications and issues'. *Journal of Moral Education*, 51(4), pp.595–599.

Zahra, D., Hedge, C., Pesola, F. and Burr, S. (2016). Accounting for test reliability in student progression: the reliable change index. *Medical Education*, 50(7), pp.738–745.