

A Structural Model of Intrinsic Motivation:

On the Psychology of Means-Ends Fusion

Arie W. Kruglanski

University of Maryland

Ayelet Fishbach

University of Chicago

Kaitlin Woolley

Cornell University

Jocelyn J. Bélanger

New York University Abu Dhabi

Marina Chernikova

University of Maryland

Erica Molinario & Antonio Pierro

University of Rome “La Sapienza”

In press, *Psychological Review*

Authors' note: Portions of the present data were presented at professional meetings, among them the Social Psychology Spring Conference, March 2016, Cortina d'Ampezzo, Italy, the Society for the Science of Motivation Annual Conference, May 2017, Boston, the Society of Personality and Social Psychology Annual Meeting, January 2017, San Antonio, and the Association for Consumer Research Annual Meeting, October 2017, San Diego. The data from the recent publications are posted on OSF.

Abstract

The term *intrinsic motivation* refers to an activity being seen as its own end. Accordingly, we conceptualize intrinsic motivation (IM) as (perceived) means-ends *fusion* and define an *intrinsicity continuum* reflecting the degree to which such fusion is experienced. Our Means-Ends Fusion (MEF) theory assumes four major antecedents of activity-goal fusion: (1) Repeated pairing of the activity and the goal, (2) Uniqueness of the activity-goal connection, (3) Perceived similarity between the activity and its goal, and (4) temporal immediacy of goal attainment following the activity. MEF theory further identifies two major consequences of the activity-goal fusion (i.e., manifestations of intrinsic motivation): (1) Perceived instrumentality of the activity to goal attainment and consequent activity engagement; (2) goal-related affective experience of the activity. Empirical evidence for MEF theory comes from diverse fields of psychological inquiry, including animal learning, brain research, and social cognition.

Keywords: intrinsic motivation, fusion, goals, means

Disclosures and Acknowledgements

There were no funding sources for this research.

All authors contributed in a significant way to the manuscript, and all authors have read and approved the final manuscript.

No authors have any conflicts of interest.

We are indebted to Robert Eisenberger, Tory Higgins, Ian de Houwer, Kasia Jasko, Mathew Roesch, Carol Sansone, Yaacov Trope, and Robert Vallerand for helpful comments on earlier versions of this paper.

A Structural Model of Intrinsic Motivation: The Psychology of Means-Ends Fusion

“Not Ideas About the Thing but the Thing Itself” (Wallace Stevens)

The terms intrinsic and extrinsic motivation enjoy wide currency among psychologists of various stripes, including researchers with interest in animal behavior (Berlyne, 1960; White, 1959), human creativity (Amabile, Hennessey, & Grossman, 1986; Sansone & Harackiewicz, 1998), educational psychology (Day, Berlyne, & Hunt, 1971; Elliot & Harackiewicz, 1994; Sansone & Harackiewicz, 2000), sports psychology (Vallerand, 2007; Vallerand & Losier, 1999), consumer psychology (Babin, Darden, & Griffin, 1994), and the psychology of work (Heath, 1999; Herzberg, Mausner, & Snyderman, 1959; Lin, 2007; Vroom, 1964; Wrzesniewski et al., 2014). Over the last half century, interest in intrinsic/extrinsic motivation phenomena has grown immensely. Whereas PsycInfo yields 146 entries for these terms in the 1950s, the comparable result for 2017 alone is a staggering 24,071, more than for such major psychological terms as ‘incentive’ (16,961), ‘motive’ (19,112), and ‘expectancy’ (20,152).

Yet despite the ample and important work on intrinsic and extrinsic motivation, there seems to exist a deep-seated confusion about the meaning of these terms. Indeed, this general topic has been approached from two different perspectives, addressing two essentially unrelated questions, namely: (1) the *structure* question: what is it about the relation between a means and a goal that makes motivation (toward *any activity*) intrinsic (vs. extrinsic), and (2) the *content* question: what *specific motives* are to be counted as intrinsic or extrinsic. It is the latter, content, question that inspired the lion’s share of intrinsic motivation research. But because the structural conception of intrinsic motivation captures the essential definition of this construct, its relative neglect in research thus far bears rectification. To address this issue, we presently lay out a structural model of intrinsic motivation and explore its features.

As a preview of what is to come, we first offer a summary description of our theory. By way of a historical background, we then discuss the definitional duality concerning the nature of intrinsic motivation, and trace this confounding to early psychological research on animal learning (Berlyne, 1960; White, 1959). We then describe our structural model of intrinsic motivation as means-ends fusion (MEF) and elaborate its postulates and implications. Finally, we consider a body of relevant evidence for our model and discuss its ramifications for everyday motivational concerns.

The Argument in Overview

An activity is *intrinsically* motivated when it is seen as its own end, and is *extrinsically* motivated when it is seen as serving a separate, ulterior end (cf. Day et al., 1971; Lepper & Greene, 1978). We assume that an activity's identity is a matter of subjective perception¹ (Higgins & Trope, 1990; Trope & Liberman, 2003; Vallacher & Wegner, 1985): in some circumstances an activity may be perceived one way (e.g., as its own end) whereas in other circumstances it may be perceived in a different way (e.g., as distinct from its end).

Identification of an activity as its own end reflects a perceptual *fusion* between the two, whereby the activity and its end are seen as inseparable parts of the same entity, thus forming a unified Gestalt (Campbell, 1958; Wertheimer, 1923). We presently posit *degrees* of fusion, depending on the strength of perceived association between the activity and its end. The greater the fusion, the more the activity is experienced as intrinsically motivated. This implies an *intrinsicality continuum* reflecting the extent of the fusion. We further assume that the quality of the goal and the degree of the fusion determine the specific way in which the intrinsic activity is

¹ We assume that subjective perceptions, rather than objective realities, determine the status of activities as intrinsic versus extrinsic. Indeed, it is the organisms' perceptions of reward contingencies in their environment that drive the preponderance of behavior.

experienced. Not all goals were created equal, and the experience of goal attainment, including the specific affective, cognitive, and behavioral reactions it elicits, varies with the goal's unique nature. We are positing a *transfer* of those various properties from the goal to the activity as a function of their degree of fusion (Fishbach et al., 2004). Whereas all goal attainment (both successful approach to desired end-states and avoidance of undesired end-states) elicits some sort of positive affect that transfers to the intrinsically motivated activity, the specific affective experience accompanying an intrinsically motivated activity (e.g., euphoria versus relief, pleasure versus pride) should vary as a function of the specific goal involved.

These notions are explored in subsequent sections of this paper. But first we elaborate on the duality in meaning that has bedeviled the study of intrinsic motivation and obfuscated a clear understanding of its properties and manifestations.

What is Intrinsic Motivation? A Tale of Two Conceptions

The structural definition. As noted earlier, the term “intrinsic motivation” is generally applied to an activity seen as its own end, whereas extrinsic motivation applies to an activity that is distinct from its end. We call this conception *structural* because it depicts the perceived *relation* (of similarity or difference) between *any* activity and its goal, regardless of the activity's specific type or content. Either one of “Going for a walk,” “having a drink,” or “listening to opera” could be intrinsically or extrinsically motivated depending on the circumstances, as could all other activities one envisages. In fact, any instance of goal attainment (or consummatory behavior), whatever that goal might be, is intrinsically motivated by definition: it represents the realization of a desire and a successful conclusion of a behavioral sequence absent any ulterior, transcendental, purpose. The only question is whether the activity (any activity) itself is

experienced as constituting goal attainment (hence as intrinsically motivated), or whether it is perceived as a separate (and hence, as extrinsic) means leading to subsequent attainment.

Intrinsic and extrinsic motives. In contrast to the present, *content-free*, analysis whereby any activity could become intrinsically motivating, popular discussions of intrinsic motivation depict it in *content-specific* terms. For instance, Wikipedia defines intrinsic motivation as “the self-desire to seek out new things and new challenges, to analyze one’s capacity, to observe and to gain knowledge” (“Motivation,” n.d.). Similarly, Ryan and Deci (2000) portray intrinsic motivation as “the inherent tendency to seek out novelty and challenges, to extend and exercise one’s capacities, to explore, and to learn... (representing) natural inclination toward assimilation, mastery, spontaneous interest, and exploration” (p. 70, parentheses added). Again, the latter definitions are content-specific in that they refer to unique motives—those related to exploration, the meeting of challenges, realization of capacities, expression of relationship closeness, and so on.

Note that according to the structural conceptualization, any of the latter “intrinsic” motives could engender an extrinsic activity aimed at its satisfaction. Take *curiosity*, often hailed as a prototypical intrinsic motive. From the structural perspective, it could motivate an extrinsic activity as in taking a plane to Africa in order to satisfy one’s curiosity about its wildlife. After all, the activity of taking a plane as such (going through security checks, sleeping poorly in narrow coach class seats, etc.) isn’t particularly enjoyable, but it is undertaken nonetheless to gratify one’s curiosity, a goal which is extrinsic to air travel.

The discrepancy between the structural and the content-specific depictions of intrinsic motivation was not lost on motivational theorists. In this vein, Higgins and Trope (1990, p. 233) commented (from the structural perspective) that any goal can instigate activities intrinsically or

extrinsically. As they put it: “Any activity can be transformed from an end in itself to a means to an end by setting up contingency conditions, so that a person must engage in that activity in order to engage in another positive activity.... The natural or inherent properties of an activity do not necessarily determine whether it will function intrinsically as an end in itself or extrinsically as a means to an end.”²

Internal motives versus intrinsic motivation. Arguably, the content-specific depictions of intrinsic motivations, for example, as featured in Deci and Ryan’s (1985, 2000) self-determination theory (SDT), were *not* actually meant to address activities perceived as their own ends. Rather, they pertained to different *motives*, some of which were considered “internal” to the individual or *innate*, and hence labeled as ‘intrinsic’ (specifically, the motives for autonomy, competence and relatedness), whereas others (e.g., power, status, money) were considered to be internalized or introjected from the *exterior* through learning, socialization, or acculturation (e.g., see Deci & Ryan’s 1985 distinction between intrinsic motives as “internal structures” and other behaviors as a “function of external controls” [p. 11]).

Internalization or introjection of external goals, although important in and of itself, has little to do with the concept of intrinsic motivation defined structurally as the relation of

² Whereas Higgins and Trope’s (1990) activity engagement theory agrees with the present notion that the concept of “intrinsic motivation” is content-free and can apply to any activity, its focus is very different from that of our MEF model. The MEF model focuses on the *relation* between the activity and its goal (i.e. the degree of fusion between them). Activity engagement theory does not address this relation. Instead, it focuses on the distinction between an activity’s primary and secondary identifications and their implications for individuals’ reactions to the activity. Indeed, activity engagement theory states: “The primary activity identification of an input determines which properties of the input are relevant when the actor is engaged in the activity.... Secondary activity identifications of an input can compete with the primary identification for the attention of the actor. Affective responses and inferential responses become associated with both the primary-identification-related properties of an input and the secondary-identification-related properties of the input” (Higgins & Trope, 1990, pp. 247-248).

isomorphism or fusion between activities and their goals. It follows that the SDT notion of intrinsic motivation is very different from the present structural treatment of this concept in the MEF model. Whereas the SDT reserves the term ‘intrinsic’ to a handful of basic human goals, we view intrinsic motivation as defined by the *relation* (of fusion) between any goal and the activity meant to serve it. Despite identical nomenclature then, the SDT and the MEF model refer to entirely different explananda.³

Historical origins of the definitional duality. The indiscriminate application of the term “intrinsic motivation” to both (1) the relation of an activity (any activity) to its end, and (2) specific goals, is arguably related to the historical circumstances wherein the concepts of intrinsic (vs. extrinsic) motivation first drew the attention of psychological scientists. Specifically, this occurred in reaction to the discovery, revelatory at the time, that animals will explore and manipulate the environment even if not offered external (food and water) rewards for so doing (Berlyne, 1960; Harlow, Dodsworth & Harlow, 1965; Yerkes & Yerkes, 1929; White, 1959). This was taken to imply that exploration of the environment *can* be intrinsically motivated; that is, it can constitute its own end (in accordance with the structural definition discussed earlier). Consequently, the fact that exploration and curiosity *can be* intrinsically motivated (i.e., *can* be a goal in itself) was taken to mean that intrinsic motivation *is* actually defined by curiosity and exploration, suggesting a *content-specific* definition of this concept.

³ The SDT and MEF model use some of the same measures of intrinsic motivation such as activity enjoyment and interest in the activity’s pursuit. However the interpretation differs. According to SDT enjoyment and interest in an activity reflect the satisfaction of an “intrinsic” human goal (i.e., of autonomy, competence, or relatedness), whereas in MEF enjoyment (positive affect) and interest in an activity reflect its fusion with any active goal, of whatever content.

In contrast to the content-specific definition of intrinsic motivation that engendered ample psychological research over the years, the structural perspective received only scant research attention. Our MEF model, described next, seeks to rectify this omission and offers a structural framework for the systematic study of intrinsic motivation, which provides novel conceptual and practical possibilities.

A Means-Ends Fusion Model of Intrinsic Motivation

Action Identification

Actions do not exist in an objective realm. They are subjectively perceived, identified and attributed. The same action performed by a tennis player may be seen as hitting a ball with a racket, serving, or scoring an ace. The fluid nature of action perceptions was stressed in Vallacher and Wegner's (1987, e.g., pp. 3, 13) *action identification* theory as well as in Trope and Liberman's (2010, e.g., p. 441) *construal level* theory (see also Sansone, Sachau, & Weir, 1989).

These notions imply that an action may be identified in terms of its goal or in terms separate from its goal. We further assume that action identification is not necessarily dichotomous, implying either one identification or the other. Instead, we are positing a continuum, reflecting the degree to which the activity is perceptually fused with the goal so that each (i.e., the activity and the goal) calls the other to mind. At the extrinsic end of the continuum, the action and the goal are experienced as completely separate. At the intrinsic end, the two are completely fused. In those instances, the activity is experienced as representing goal attainment (or *consummatory behavior*). Eating when hungry, drinking when thirsty, receiving an award when craving recognition are simultaneously activities and ends in themselves; they are precisely what one desires at a given moment. Between the intrinsic and extrinsic ends of the intrinsicality

continuum lie activities with intermediate degrees of fusion with the goal, representing moderate degrees to which activities are identified, or defined by their goals.

Unidirectionality. We assume that the activity's fusion with its goal is *unidirectional*: the activity "loses itself" in the goal as it were; it is assimilated to the goal and assumes the goal's identity. "Running" when fused with the goal of "fitness," is subjectively experienced as (the attainment of) fitness. "Having a drink at a bar" when fused with the goal of "unwinding" becomes the experience of unwinding, etc. Because the identity of the fused activity is tantamount to that of the goal, it is the goal properties that transfer to the activity, rather than vice versa. We thus assume an asymmetrical process of influence running from the goal to the activity. And because the activity's possible prior identity is largely lost and replaced by the goal's identity, it cannot exert a reciprocal influence on the goal's identity. In this sense, the means-ends fusion process is a "one way street."

Intrinsic Motivation as a Goal-driven Process

All activities, regardless of their degree of intrinsicity, are motivated; they are animated by goals that individuals desire to attain. Because an activity's intrinsicity is determined by its fusion with the goal, that goal's properties determine how the fused activity is experienced. In what follows we discuss in these terms three major goal aspects, namely: *goal activation*, *goal magnitude*, and *goal related affect*.

Goal Activation. For any activity to occur, its goal needs to be active at the time (cf. Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001). The goal constitutes the source of the affective and experiential responses to the activity, hence its fulfillment or suppression should weaken or void those reactions as well. Essentially, intrinsic motivation toward an activity is "borrowed" or derived from a currently active goal with which the activity is fused.

For example, whereas eating can be intrinsically motivated, it will not take place if one were not hungry, that is, if the food consumption goal was inactive at the time. On occasion, one might be intrinsically motivated to paint or play, but not if the dominant goal at the time was to sleep, or study.

Goal Magnitude. It is generally agreed that motivation is a matter of degree, and that it varies with goal magnitude (Kruglanski, Chernikova, Rosenzweig, & Kopetz, 2014). Degree of intrinsic motivation toward an activity thus depends on the magnitude of the goal with which the activity is fused: the higher such magnitude, the stronger the individual's intrinsic motivation toward the fused activity. For instance, one could be intrinsically engaged in drawing (running, writing, eating) if these activities were strongly fused with their goals; yet while performing these activities, the person could experience high, low, or medium degrees of motivation. Some (intrinsically motivated) people might be passionate about drawing, for example, whereas other (equally intrinsically motivated) people may care about it to a lesser degree. We thus assume that motivation strength is *orthogonal* to the degree of intrinsicity; goal magnitude does not change the status of the activity as more or less intrinsic. At any location on the continuum (i.e., at any degree of intrinsicity), the individual's motivation strength toward the activity could vary over the entire gamut of strengths.

Goal-derived affect. Different goals may evoke different qualities of affect upon success in goal attainment and different qualities of affect upon failure of attainment. Some goals (e.g., success at an exam) induce pride upon attainment or shame upon non-attainment, other goals (e.g., having a sumptuous meal) elicit pleasure on attainment and disappointment on non-attainment. Higgins (e.g., 2000) famously distinguished between *promotion* goals, whose attainment elicits joy and elation and whose non-attainment elicits sadness and dejection, in

contrast to *prevention* goals whose attainment elicits relief and calm and whose non-attainment elicits agitation and anxiety. The different qualities of affect that accompany the attainment or non-attainment of different goals are transferred to activities driven by those goals as a function of the activities' degree of fusion with those goals. We elaborate on this point at a later juncture. As for now, we characterize in greater detail our fusion concept and identify its antecedents and consequences.

Fusion's Antecedents

As noted earlier, fusion between the activity and its end connotes the degree that the two are experienced as isomorphic. The *extent* to which an individual feels she is eating because she wants to eat, resting because she wants to rest, or running because she wants to run represents the degree of such fusion and locates a given activity on the intrinsicality continuum.

In identifying the antecedent conditions of fusion we draw on the perceptual principles of Gestalt psychology (e.g., Wertheimer, 1923) concerning factors “which lead discrete elements to be perceived as (forming) a whole organization”, or as parts of a uniform *entity* (Campbell, 1958, p. 17, parentheses added). Specifically, we assume that humans (and members of other species) react to their *perception* of the situation they are in, including the perceived means-ends relations (or reward contingencies) in their environment. Gestalt principles identify the determinants of such perception, hence they constitute a broad conceptual framework for understanding goal driven cognition, affect, and behavior.

Theorists in the cognitive behavioral tradition (e.g., Holyoak, Kohl & Nisbett, 1989; Rescorla & Wagner, 1972; Tolman, 1933) recognized the crucial role of perception in driving the behavior of organisms. Accordingly, they elaborated on how Gestalt principles govern learning, thus forging individuals' situational perspective that guides action. In this vein, Rescorla (1985,

p. 113) noted that “[Pavlovian] conditioning ... is importantly influenced by organizational [Gestalt] factors” (p. 113). And Holyoak (1990, p. 293) asserted that “cognitive science is beginning to establish a firmer basis for the Gestalt intuition that human perception and thinking have a fundamental unity.” Consistent with that view, we now discuss major Gestalt principles as they bear on means-ends fusion, the central phenomenon of present interest.

(1) **Repeated coupling.** The Gestalt principle of *Common Fate* asserts that “elements are...likely to be perceived as parts of the same organization (if they) move together in the same direction, and otherwise in successive temporal observations share a ‘common fate’” (Campbell, 1958, p. 17). As Hsiao (1928, p. 290) observed in this connection, in order to bring out a Gestalt—conjunction between parts “has to be repeated again and again.” Accordingly, we posit that a major antecedent of fusion between an activity and its goal is *repeated coupling* of the two, representing their “movement together in the same direction.”

(2) **Linkage uniqueness.** Gestalt theorists also affirmed that a perceptual unit (i.e. fusion) will be formed to the extent that the mutual resemblance of the parts “is greater than...that of such members and any other parts of the situation” (Kohler, 1940, p. 135). We take this to mean that unit formation should depend on uniqueness of the linkage⁴ between the parts, such that the greater the uniqueness, the greater the fusion. Specifically, the more unique the linkage between the activity and its end, the more the activity should be fused with its end and, hence, should be experienced as intrinsically motivated.

(3). **Similarity.** The principle of similarity in Gestalt theory asserts that similar elements tend to create an experience of a relationship between them, resulting in their perception as a unified whole. In this vein, Wertheimer wrote of “the tendency of like parts to band together”

⁴ Resemblance representing one type of common bond or linkage between the parts.

(Wertheimer, 1923, p. 625). Helson, too, asserted that “parts which are alike tend to go together to form groups of wholes, *ceteris paribus*, rather than unlike elements. [Moreover] the likeness may be with respect to quality, size, position or what not” (Helson, 1933, p. 28). And echoing these notions, Campbell stated that “Similar elements are more likely to be perceived as parts of the same organization” (Campbell, 1958, p. 17).

Presently, we assume that a similarity between the activity and its goal would contribute to the experience of fusion between the two. Such similarity could derive from the semantic relation between the activity and the goal. For instance, labeling an activity as ‘practice’ may be related to the goal of ‘improvement,’ but less so to the goals of ‘relaxation’ or ‘pleasure.’ Labeling the activity as ‘dieting’ semantically evokes the goal of ‘weight loss’ but less so of (eating) ‘enjoyment’, ‘building muscle,’ etc.

(4) **Immediacy.** According to the Gestalt principle of *Proximity*, elements close together are more likely to be perceived as parts of the same organization (Campbell, 1958, p. 17). Helson (1933) translated such proximity into temporal distance between the parts such that “Parts having smaller intervals [of time] between [them] tend to form wholes rather than parts separated by larger intervals” (Helson, 1933, pp. 27-28).

We apply this rule to temporal contiguity between activity completion and goal attainment. Specifically, the greater such contiguity or the immediacy with which goal attainment follows the activity, the greater should be the activity-goal fusion, and the attendant experience of intrinsic motivation.

We assume that the foregoing antecedents of activity-goal fusion are additive in their effects. That is, each may contribute independently to perceived fusion, and none is absolutely necessary for such a perception.

Fusion's Antecedents in Prior Research

Psychological research over the years yielded ample findings pertinent to fusion's antecedents identified above. These results are briefly reviewed below.

1. Repeated coupling. Earlier we noted the close relation between the Gestalt laws of perception and the principles of classical conditioning (cf. Rescorla, 1985; Holyoak, 1990). Consistent with the Gestalt principle of Common Fate, an extensive body of research on classical conditioning (for a review see, e.g., Hilgard & Bower, 1966) attests that repeated pairing of a neutral and an affectively charged stimulus forges an association⁵ between them. As a result of classical conditioning, the neutral stimulus acquires value, and comes to serve as what is known as a *secondary reinforcer* (e.g., see discussion by Robinson & Berridge, 1993).

According to Hilgard and Bower (1966, pp. 121-122), secondary reinforcement constitutes: "A stimulus that *occasions* or accompanies a [primary] reinforcement [and] acquires thereby reinforcing value of its own." In a seminal demonstration of this phenomenon, Wolfe (1936) taught chimpanzees to insert tokens in the slot of a vending apparatus, which then automatically released a grape. Subsequently, the animals were trained to perform simple responses (lifting a lever, pulling a train by means of a cord) in order to obtain the tokens. In other words, obtaining the tokens now acquired a rewarding value and became "fused," as it were, with the original goal of food obtainment.

We are assuming that such an association is also formed as a consequence of repeated coupling⁶ of the activity and the goal (Custers & Aarts, 2005), creating a fusion between the two:

⁵ Throughout the present discussion, the term 'association' is meant to denote a linkage between mental representations of different events.

⁶ Notably, mere repeated coupling is not always sufficient to create fusion, as attested by studies investigating the blocking effect (Kamin, 1969) in which after a given neutral stimulus was conditioned to a UCS, repeated pairing of a new neutral stimulus with the old one (i.e., now the CS) would not result in

the means-ends fusion discussed herein. In other words, intrinsic motivation is created in the process that turns an activity into a *secondary reinforcer*. For instance, for an individual who repeatedly experienced running as associated with (the goal of) exercising, a *running-exercising* bond will be created such that considering running will bring exercising to mind and vice versa. For the individual in question, exercising *is* running (rather than an external outcome of running). In research by Levey and Martin (1975), an affectively neutral photograph paired with a disliked photograph led the neutral photograph itself to become disliked. In this case, pairing the originally disliked photograph (avoidance goal) with a neutral photograph lent the latter the distasteful quality of the former (something to be avoided).

The notion that an activity can serve as a reinforcer was anticipated by Premack (1962), and elaborated by Timberlake and Allison (1974). Specifically, Premack (1962) ran a study with rats, in which he found that drinking could serve as a reinforcer for running, and vice versa: running could be made to serve as a reinforcer for drinking. Timberlake and Allison (1974) elaborated on these findings in their concept of ‘response deprivation’ whereby for a rat deprived of the ability to drink—*drinking* can serve as a reinforcer, or a goal; similarly, for a rat deprived of the ability to run, *running* can serve as a reinforcer, or a goal. Whereas drinking and running are related to basic organismic needs (for hydration and activity respectively) whose deprivation creates the conditions for goal setting (e.g., to drink or to run), any activity could acquire value, and hence become (fused with) a goal (i.e., be experienced as intrinsically motivated) through the classical conditioning process as discussed earlier.

Decoupling. Just as repeated coupling of an activity and a goal ‘solders’ them together,

conditioning of the new stimulus to the UCS. It appears then that once an association between a given CS and a UCS (in the sense of linkage between mental representations of the CS and the UCS) was established, the organism is not motivated to attend to possible additional association as the old one suffices for predictive purposes.

decoupling the two reduces or eliminates the fusion. In animal learning research, repeated non-reward of a behavior is known to result in *extinction*, causing decline in value of the secondary reinforcement; in present terms, this means de-fusing the connection between the activity and goal attainment, rendering the activity less intrinsically motivating. Extinction may occur more or less gradually, depending on how perceptible the decoupling is. For instance, if the original conditioning phase involved *continuous reinforcement* (with the UCS following the CS on every trial), the contrast between the conditioning and the extinction phase (during which the CS is not followed by the UCS) is salient, which results in relatively rapid extinction. However, under *partial* or *intermittent reinforcement* (with the UCS appearing on some trials only), the contrast between the conditioning and extinction phases is less clear (as some conditioning trials were non-reinforced as well), which typically results in slower extinction (for discussion of the partial reinforcement effect (PRE) in classical conditioning see, e.g., Pavlov, 1927, p. 49; Chan & Harris, 2017, p. 20).

Severance of the association between the activity and goal (reward) attainment may be obscured also in the case of avoidance learning, where the activity is intended to evade a negative event, because non-appearance of the aversive event may be interpreted as successful goal attainment. Indeed, research has shown that avoidance learning is particularly resistant to extinction (Solomon, Kamin, & Wynne, 1953). In such a case, fusion between the activity and the goal may endure considerably even though the aversive event might not have happened in absence of the activity. In this sense, avoidance of aversive effects may be intrinsically motivating. For instance, avoiding walking under a ladder because one is superstitious about it may provide relief.

In research with human subjects, decoupling the activity from its goal may be more or

less clear, depending on goal-type. For instance, tangible, material rewards (e.g., money) may be readily decoupled from an activity by mere experimental fiat. In contrast, if the reward is perceived competence on a task, removal of competence *feedback* might not remove the *sense* of competence while performing the activity. In such a case, the activity may not be decoupled from its goal even though the feedback may have been discontinued. For instance, in acquiring a given skill (e.g., playing tennis), the individual might learn what constitutes the correct performance (e.g., the perfect serve) and derive a sense of competence from exhibiting it even in the absence of explicit feedback from the instructor.

2. Links' uniqueness. The *fewer* the number of links between an activity and a goal, the stronger the bond between them is. This notion recalls Anderson's (1983) *fan principle* whereby the greater the number of links between a cue and a concept, the lesser the ability of the cue to activate the concept. In the present context, the bond between an activity and a goal may be diluted by alternative bonds between the goal and other activities that serve it (in an *equifinality* configuration) or between this same activity and its multiple goals (in a *multifinality* configuration). The classic research on intrinsic motivation used this principle to manipulate intrinsic motivation. For example, children were less intrinsically motivated to color after they were promised a reward for coloring (Lepper et al., 1973). Adding a new goal for coloring diluted the association between coloring and other goals (e.g., self-expression), undermining intrinsic motivation to pursue this activity.

3. Similarity. A study by Sansone, Sachau, and Weir (1989) examined intrinsic interest in a computer game. The investigators had participants read instructions as to how to improve their performance at the game. Performance instructions enhanced participants' interest when the task was described as a "game of skill," but reduced interest when the task was described as a

“fantasy game.” From the present perspective, the performance instructions may have augmented the perceived similarity of the activity to that goal. However, when the activity was described as a fantasy game, performance instructions reduced the perceived similarity between the activity and the “fantasizing” goal, thus decreasing interest in the activity.

Freitas and Higgins (2002) suggested that the fit between an action’s strategic orientation and the actor’s regulatory state, or goal, can affect the degree of enjoyment the action provides. The authors found that individuals who were in a regulatory state focused on hopes and accomplishments (i.e., promotion focus) experienced eagerness-related activities more favorably than vigilance-related activities. On the other hand, participants who were in a regulatory state focused on duties and responsibilities (i.e., prevention focus) experienced vigilance-related actions more favorably than eagerness-related activities.

Presumably, eagerness and excitement belong in the semantic universe of accomplishment and advancement (hence, promotion), to a greater extent than vigilance and care. In that sense, an eager manner of performing an activity is more *similar* to the goal of accomplishment than is a vigilant manner. In contrast, vigilance and care belong in the semantic universe of responsibility and protection (hence of prevention) to a greater extent than eagerness and excitement. Hence a vigilant manner of performing an activity is more similar to the goal of responsibility than an eager manner.

Finally, Higgins, Cesario, Hagiwara, Spiegel, and Pittman (2010) found that the “intrinsic” effect of increasing subsequent interest in redoing the activity occurs when the participants’ goal orientation (toward fun, or importance) is similar to the manner of doing the activity (enjoyable, serious), whereas the “extrinsic” effect of decreasing subsequent interest occurs where the activity and its goal are described in dissimilar terms.

4. Immediacy. Research findings attest that *temporal contiguity* between conditioned and unconditioned stimuli is an important determinant of associative learning (e.g., Balsam, Drew & Gallistel, 2010; Gibbon, Baldock, Locurto, Gold, & Terrace, 1977; Gormezano & Kehoe, 1981; Ost & Lauer, 1965; Reynolds, 1945; Smith, 1968; Stein, Sidman, & Brady, 1958; Vandercar & Schneiderman, 1967; Wickens, Meyer, & Sullivan, 1961). Whether two unrelated events become associated depends on their co-occurrence in time. For example, one study found that pairing two flavors *simultaneously* promoted stronger flavor-flavor associations in rats than pairing them *sequentially* (Rescorla, 1980). Accordingly, we assume that the closer the temporal contiguity between activity completion and goal attainment, the greater the perceptual fusion between the two.

Fusion's Consequences

Fusion between an activity and a goal is indicated by the activity being mentally represented in terms of the goal. That is, irrespective of the specific antecedent of the activity-goal fusion, the greater its degree, the greater should be the activity's perceived similarity to the goal, and the perceived feature-overlap between it and the goal. As Helson (1933, p. 28) put it in Gestalt theoretic terms: "If parts have occurred in a structure together and one of them re-appears bearing its membership character, it tends to re-instate the whole" (Helson, 1933, p. 28). We assume that such "reinstatement" of the goal by the activity characterizes intrinsic motivation. Intrinsic motivation in turn, has the following consequences: (1) the activity's perceived *instrumentality* to its goal attainment is enhanced, and (2) the activity comes to elicit *affective and behavioral reactions* that mimic those evoked by the goal's attainment. Prior research assessed intrinsic motivation in terms of (1) activity engagement in the free choice paradigm and (2) activity's positive experience. As will be seen, the former mode of assessment is extended

here through tapping the activity's perceived instrumentality, and the latter mode through tapping the specific goal-related emotions. Let us explain.

1. **Instrumentality.** Because the highly fused activity is mentally represented as isomorphic with goal attainment, it should be experienced as realizing attainment, and in this sense being "instrumental" to accomplishing the goal *the activity is fused with*. In turn, perceived instrumentality should prompt activity engagement en route to goal pursuit (as in the free choice paradigm often used to study intrinsic motivation phenomena). For instance, if for a given individual 'running' was tightly fused with 'fitness', activating the goal of fitness should increase the likelihood of running. Similarly, de-fusing the association between the activity and goal attainment should reduce the likelihood of performing the activity when the goal in question is active.

2. **Attainment-related reactivity.** The greater the fusion between an activity and its goal's attainment, the more the individual's affective and behavioral reactions to the activity should resemble those that characterize the attainment process. For instance, individuals should be motivated to pursue the fused *activity* (just as they are motivated to achieve the *goal*), and while so doing they should experience affect akin to that which accompanies attainment.

Such affect should be positive on the whole, in that attainment of any goal is satisfying by definition. However, the specific quality of affect toward the fused activity should vary in accordance with the goal's specific nature. For instance, whereas *promotion* goals should imbue fused activities with feelings of happiness and enjoyment, *prevention* goals should imbue them with feelings of relief and calm (Higgins, 1987, 2012).

In Summary

The present arguments may be summarized in a series of formal statements as follows:

Definition 1: *The term “intrinsic motivation” refers to an activity being perceived as an end in itself.*

Postulate 1 (Goal dependence): *Intrinsically motivated activity will only occur if the goal it serves is activated.*

Postulate 2 (Activity perception): *Intrinsic motivation (perceiving an activity as an end in itself) is the outcome of fusion between the activity and its end.*

Corollary to Postulate 2 (Intrinsicality continuum): *The fusion experience varies in degree, defining a continuum of intrinsicality; at the intrinsic end, the activity is inextricably interwoven with its goal, whereas at the extrinsic end, the activity and the goal are experienced as entirely separate.*

Postulate 3 (Fusion’s antecedents): *Four major antecedents of activity-goal fusion are: (1) Repeated pairing of the activity and the goal, (2) Uniqueness of the activity-goal connection, (3) Perceived similarity between the activity and its goal, and (4) Temporal immediacy of goal attainment following the activity.*

Postulate 4 (Fusion’s consequences): *Two major consequences of the activity-goal fusion are: (1) Perceived instrumentality of the activity to goal attainment and consequent activity engagement (2) Elicitation by the activity of goal-related reactions.*

Figure 1 here

These notions are represented graphically in Figure 1, in which intrinsic motivation (IM) is treated as a latent construct: (1) variously induced by the four identified antecedents of activity-goal fusion (i.e., repeated coupling, linkage uniqueness, similarity, and linkage immediacy); (2) producing the consequences ascribed to IM (i.e., perceived instrumentality and goal-related reactions). Figure 1 also represents our assumption that activity engagement is

moderated by goal activation: Regardless of whether the activity is more or less intrinsically motivated or fused with its goal, the activity will not take place, nor will effects of its fusion with the goal be apparent, if the goal was inactive.

Empirical Evidence

Antecedent-Consequence Pairings

According to our model, any antecedent of means-ends fusion should determine any of its consequences. We now examine the available evidence for this assertion, considering in turn each of the three fusion antecedents portrayed in Figure 1.

I. Repeated Association: Strengthening the Activity-Goal Link

1. Motivating properties of the fused activity. Repeated association of an activity with successful goal attainment should foster a fusion between the two, i.e. promote an assimilation of the activity to attainment, such that the activity will come to elicit affective and behavioral reactions originally evoked by attainment.

2. Perceived instrumentality of goal-fused activities. In animal learning studies, researchers relied, *per force*, on subjects' overt behavior for inferring the motivating effects of activity-goal associations. In contrast, social cognition researchers can rely in addition on participants' subjective experiences conveyed via self-report measures. In this vein, Zhang, Fishbach, and Kruglanski (2007, Study 4) used a subliminal priming technique to forge a goal-means connection, and assessed its effects on reported instrumentality of the means to goal attainment. Specifically, they repeatedly (and subliminally) flashed words related to the goal of strengthening muscles (e.g., "muscle") concomitantly with a (supraliminal) presentation of a specific means to that goal (i.e., "jogging"). This procedure significantly increased the perceived instrumentality of the means to the goal while concomitantly decreasing the perceived

instrumentality of the same means to an alternative goal (increasing blood oxygen). These findings were conceptually replicated in research by B elanger, Schori-Eyal, Pica, Kruglanski and Lafreni ere (2015, Study 4).

In summary, research from diverse psychological domains (e.g., animal learning, social cognition studies) attests that repeated association of an activity with goal attainment, imbues the activity with properties of the goal; the activity is subjectively experienced as a “*goal surrogate*” as it were, eliciting the motivation to engage in it, and its perception as instrumental to goal attainment.

II. Links’ Uniqueness: Diluting the Activity-Goal Association

Whereas repeated coupling of the activity with goal attainment strengthens their fusion, introduction of alternative links between the activity and other goals or between the goal and other activities should dilute their fusion, as noted earlier (Anderson, 1983; Rescorla & Wagner, 1972). In turn, dilution should reduce the activity’s perceived instrumentality to goal attainment, and reduce the activity’s elicitation of goal-related reactions.

1. Multifinality configuration: Perceived instrumentality. Zhang et al. (2007) investigated these notions in several studies on the *multifinality* configuration, in which a single means is attached to several goals (see Figure 2).

Figure 2 here

Study 1 in their paper manipulated the number of goals (one vs. two) served by each of different activities. Participants read three short essays discussing respectively *aerobic exercise*, *consumption of tomatoes*, and *withdrawing from caffeine*. One sentence in each essay described how each of those means may satisfy one or two goals. For instance, the first essay discussed how aerobic exercises could contribute to people’s health. In the one-goal condition, the essay

read:

“Regular aerobic exercise helps protect you from heart disease [helps you maintain healthy bones]. For the greatest overall health benefits, experts recommend that you do 20 to 30 minutes of aerobic activity three or more times a week and some type of muscle strengthening activity and stretching at least twice a week. However, if you are unable to do this level of activity, you can gain substantial health benefits by accumulating 30 minutes or more of moderate-intensity physical activity a day, at least five times a week.”

In the two-goals condition the first sentence was replaced with “Regular aerobic exercise not only protects you from heart disease but also helps maintain healthy bones.” Participants then rated the extent to which aerobic activity was effective in promoting the first goal listed (i.e., either protection from heart disease or maintenance of healthy bones). It was found that aerobic exercise was perceived as a more effective means with respect to the first/only goal listed in the single goal condition than in the two goals condition. Similar results were obtained with the remaining activities (i.e., consumption of tomatoes and withdrawal from coffee) addressed in that first study (see Figure 3).

Figure 3 here

Further studies in the Zhang et al. (2007) series (1) conceptually replicated these initial findings in a context where participants self-generated the goal (or goals) to a given “means” (namely, “studying in a library”, “attending college” or “using a laptop”) by listing either one or three things that this means could help accomplish (Study 2); (2) showed that the dilution effect is more pronounced when the goals listed are more (vs. less) distinct from one another (Study 3); (3) demonstrated that association strength mediated the dilution effect on perceived instrumentality (Study 5); and (4) showed that the dilution effect on perceived instrumentality

affects actual choice behavior (Study 6).

2. Equifinality configuration: a. Perceived instrumentality. Findings described above support the notion that uniqueness of the goal-means link determines the means' perceived instrumentality to goal attainment. Note, however, that the Zhang et al. (2007) work examined this phenomenon exclusively within the *multifinality* configuration. Yet our theory requires that the same link-uniqueness effects should hold also with ties that run in the opposite direction, connecting a goal to several means (in an *equifinality* configuration; see Figure 4) (Heider, 1958; Kruglanski et al., 2002). Bélanger et al. (2015) investigated these possibilities in an additional series of studies.

Figure 4 here

Specifically, using the experimental paradigm developed by Zhang et al. (2007), Bélanger et al. (2015) demonstrated that presenting additional means (Study 1) and having participants self-generate multiple means to a goal (Study 2) led, as expected, to a reduction in perceived instrumentality of a given means to the goal. Second, this effect was stronger when the means within the set were seen as distinct from each other rather than as similar (Study 3), highlighting the separateness of the goal-means links. The dilution effect was found to influence means selection and the magnitude of means engagement (Study 5). Lastly, presenting additional means to a goal reduced participants' intrinsic motivation to pursue that means; importantly, this effect was mediated by reduced perceived effectiveness of the original means (Study 6).

b. Experienced affect. According to the present theory, one consequence of the fusion between an activity and a goal is that properties of the goal “seep down” and come to characterize the activity as well. To test this proposition, Fishbach, Shah, and Kruglanski (2004, Study 1) had participants generate one vs. two means to their goal. As predicted, participants

subsequently reported more positive affect toward the activity in the singular vs. two means condition. As we have seen, the two (vs. single) means configuration dilutes the association between the activity and the goal, hence it should afford a less pronounced transfer of positive affect from the goal to the activity, consistent with the Fishbach et al. (2004, Study 1) findings.

3. Compensatory effects. Both Zhang et al. (2007) and Bélanger et al. (2015) demonstrated *compensatory effects* of means-ends association whereby increasing the strength of one link (and hence perceived instrumentality) between an activity and one of its goals (in a multifinality configuration) or between a goal and one of its means (in an equifinality configuration) reduces the strength of alternative links between the activity and its other goals, or between the goal and its other means.

Taken as a body, the repeated association (coupling) and dilution findings are consistent with the notion that the stronger the link between a means and a goal, the greater the fusion between them. As a consequence, the means is more likely to be enacted, to be perceived as instrumental to goal attainment, and to elicit positive affect. In other words, the more tightly the activity appears to be connected to its goal, the more it appears to be intrinsically motivated, or to exhibit features characteristic of the goal.

III. Immediacy of the Activity-Goal Link

According to the MEF model (see Figure 1) the greater the temporal proximity, or immediacy, of goal attainment to performance of the activity, the more pronounced should be the activity-goal fusion. As a consequence, stronger temporal proximity of the activity to goal attainment should result in the activity exhibiting features that are typically associated with attainment.

This notion echoes a phenomenon previously documented in conditioning research (De

Houwer, Thomas, & Baeyens, 2001; Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010; Razran, 1954),⁷ whereby the length of the time interval between the conditioned and the unconditioned stimulus determines the frequency of the conditioned response (McAllister, 1953; Wolfe, 1936). In the present conceptualization, the activity itself serves as a conditioned stimulus. Therefore, the immediacy of the activity association with its goal (i.e., the reward or the “unconditioned stimulus”) should determine the degree to which the positive experience of goal attainment (e.g., satisfaction, enjoyment) should transfer to the activity.

1. Affective and behavioral reactions to the activity. Woolley and Fishbach (2017a) had participants watch a clip from a satirical news program and reflect on two outcomes they received from watching it: becoming more informed and gaining conversation topics. Participants either elaborated on how these outcomes were an immediate benefit they received while watching the clip (immediate condition) or a delayed benefit they would receive in the days or weeks after watching the clip (delayed condition). It was found that participants perceived the show as more *enjoyable*, *interesting* and *fun* when they framed these benefits as immediate (vs. delayed), that is, when watching served immediate versus delayed goals. In this sense, Woolley and Fishbach’s (2017a) findings support the notion that immediacy of the activity-goal link is related to a major consequence of intrinsic motivation, the positive affect it elicits as shown in Figure 1 above (see also Woolley & Fishbach, 2015, 2016).

⁷ Though both evaluative and classical conditioning find evidence for time interval effects, our notion of intrinsic motivation is more closely related to classical conditioning in which the CS signals the arrival of some *wanted* state of goal attainment (approaching or avoiding the UCS) than to evaluative conditioning where the UCS is merely a *liked* (or *unliked*) stimulus that does not necessarily represents a goal the individual *wants* to approach or avoid. For a distinction between liking and wanting see Berridge, Robinson, and Aldridge (2009).

Participants in another study (Woolley & Fishbach, 2017a) read that using a moisturizer would have an immediate or delayed effect (of promoting smooth skin). As predicted, those in the immediate versus delayed condition reported greater activity enjoyment. In yet a different study (Woolley & Fishbach, 2017a) participants reported their experience while pursuing a task that delivered either an immediate or delayed outcome, namely chocolate received either while performing the task or following task completion (though no one was allowed to consume the chocolate during task performance). It was found that receiving the immediate (vs. delayed) chocolate increased experienced enjoyment, interest, and intention to pursue the experimental task in the future (for similar measures of intrinsic motivation see Deci, Koestner & Ryan, 1999; Kruglanski, Alon, & Lewis, 1972; Tang & Hall, 1995).

Whereas the study just described assessed mere intentions, Woolley and Fishbach's (2017a) next study found that an immediate bonus (compared with a delayed bonus or no bonus) affected the actual behavioral choice to pursue a task. Participants in this experiment read a book excerpt and learned that they would receive a bonus, either immediately following the experiment or with some delay. After completing the task, participants received a choice to either continue working on the same task (i.e., continue reading the book) or complete a different task, with no opportunity to earn an additional bonus either way. Greater continued engagement in that task occurred when it was previously associated with an *immediate bonus* compared with *delayed* or *no bonus* conditions (see Figure 5).

Figure 5 here

2. Fractional anticipatory goal responses. Decades ago, researchers observed that in proximity to the goal, animals exhibit reactions that appear to constitute muted and partial versions of behaviors emitted during goal consumption; these were, therefore, labeled *fractional*

anticipatory goal responses (FAGR).⁸ Hull (1931, p. 494) mentions the reactions of the dogs in Pavlov's studies as an example of FAGR. As he put it: "The...salivation of Pavlov's dogs evidently also represents anticipatory goal or terminal reactions. The...dogs executed gross mouth movements such as vigorous licking of the lips as well."

Pavlov himself describes the FAGR (though he does not label it as such): "Firstly the animal exhibits a reflex activity directed toward getting hold of the food and eating it or, in the case of rejectable substances, towards getting rid of them out of the mouth; and secondly, in both cases an immediate secretion of saliva occurs, in the case of the food, to start the physical and chemical processes of digestion, and, in the case of rejectable substances, to wash them out of the mouth" (Pavlov, 1927, p. 17). He continues that the salivation "is frequently copious and...is always accompanied by the motor reaction peculiar to the conditioned stimulus which was used" (Pavlov, 1927, p. 93). Pavlov offers another example of differential FAGRs, which occurred when a dog was trained to expect either acid or food from an experimenter: "In the case of the alimentary reflex, the dog turned toward the experimenter, dividing its gaze between him and the food and smacking its lips; in the case of the reflex to acid, the dog turned away from the experimenter, snorting and whining, shaking its head and making ejective movements with its tongue" (Pavlov, 1927, p. 184).

Amsel (1949, p. 786, parentheses added) also quotes Kendler's (1946) description of the fractional anticipatory goal response occurring "in the food box and in the alley leading to the food box (and comprising) anticipatory eating responses (salivation, masticatory movements, etc.). In a similar manner, anticipatory drinking responses become conditioned to the alley leading to water..."

⁸ Note that the FAGRs do not denote mere *anticipation* of goal consumption. Rather, they are "as-if" responses that resemble those the animal emits to the goal object (food or water) as such.

From the present perspective, FAGRs could reflect a fusion between the *goal* and a goal directed *activity* occurring as a function of temporal proximity or immediacy of the activity to the goal: The more immediate or proximal to the activity is goal attainment, the more the activity is fused with the goal. As a consequence, the activity takes on the features of consummatory (goal attainment) behavior as observed by behavioral learning theorists (e.g., Amsel, 1949; Amsel, 1962; Guthrie, 1940; Hull, 1943; Spence, 1956).

Researchers also observed that behaviors analogous to those emitted to the goal object (akin to the FAGRs) are often directed at the CS associated with the goal. This phenomenon was referred to as *auto-shaping*. In the original set of studies addressing this phenomenon, Brown and Jenkins (1968) found that pigeons approached and pecked a small key whose brief illumination (conditioned stimulus, CS) preceded the presentation of grain (US), even though these behaviors had no effect on grain delivery.

Wasserman, Franklin, and Hearst (1974) replicated these effects with pigeons and additionally demonstrated that a withdrawal response can be obtained as a result. Specifically, they found that “pigeons approached and pecked an illuminated key that was positively correlated with food delivery and *positioned themselves relatively far* from an illuminated key that was negatively correlated with food delivery” (p. 616).

In summary then, studies with humans and animals suggest that immediacy of the activity-goal link appropriately impacts the affective and behavioral reactions that the activity elicits: the greater the immediacy, the more those reactions resemble those expected to characterize an intrinsically motivated activity, that is, an activity that constitutes its own goal.

Magnitude of Consequences

According to our theory, *degree* of Means-Ends Fusion should determine the magnitude of the fusion's consequences. This proposition was explored in several studies in which the mental representation of the activity (indicative of the degree of fusion), and some of fusion's postulated consequences (in particular, subjective experience of affect toward the activity) were assessed. Fishbach et al. (2004, Study 2) assessed the degree of association between each of two goals and its related means through a sequential priming procedure (cf. Anderson & Bower, 1973; Higgins, 1996; Meyer & Schvaneveldt, 1971; Srull, 1981). It was found that affect felt toward the means corresponded more to that felt toward the goal when the two were strongly versus only weakly associated (see Figure 6). The greater degree of affective correspondence between the goal and the activity, as function of their degree of association, thus supports our theory.

Figure 6 here

A recent study by Bélanger (2017) used the overlapping Venn diagrams technique to assess the degree of perceived fusion between the activity of “going to the gym” and the goal of “being healthy” (Aron, Aron, & Smollan, 1992). Specifically, participants were presented with a scale consisting of a set of 5 circle sets whose relation to each other ranged from no-overlap, through partial overlap to complete overlap (in which one circle was contained in the other). One of the circles was said to represent “going to the gym” and the other “being healthy” and participants were instructed to select the circle pair that best represented their perceived relation between these two notions. Subsequently, participants' enjoyment and interest in attending the gym was measured using the Situational Intrinsic Motivation Scale (SIMS; Guay, Vallerand, & Blanchard, 2000). Results indicated that scores on the fusion scale and SIMS were positively and significantly correlated.

In summary, findings reviewed above are consistent with the notion that the degree of fusion between the activity and the goal is related to the degree to which affect toward the activity resembles affect toward the goal.

The Mediating Role of Fusion

We argue that the subjective experience of fusion mediates the effect of means-ends association on the expression of intrinsic motivation. A recent study by Woolley and Fishbach (2017a) provides direct evidence for this implication. Specifically, in that study the authors manipulated the *immediacy* of goal attainment in the moisturizer-smooth skin paradigm mentioned earlier (i.e., presented the moisturizer effect as immediate vs. delayed) and measured the mental representation of the activity-goal fusion via the Venn diagram technique mentioned earlier (cf. Aron et al., 1992). Consistent with our theory, it was found that when moisturizing (*the activity*) was said to provide immediate (vs. delayed) outcomes, participants associated it with smooth skin (*the goal*) to a greater degree. Of special relevance, the Venn fusion measure mediated the effect of immediacy (i.e., means-ends association) on expected enjoyment of and positive experience of the activity (reflecting intrinsic motivation).

Participants in another study (Woolley & Fishbach, 2017a) were informed that reading a book excerpt would provide an immediate or delayed bonus (\$0.25 delivered the same day or in a month). As predicted, those in the immediate versus delayed condition reported greater activity enjoyment. The authors further measured the mental representation of the activity-goal fusion via the Venn diagram technique. Consistent with our theory, it was found that when reading (*the activity*) was said to provide an immediate (vs. delayed) bonus, participants associated it with receiving rewards (*the goal*) to a greater degree. Of special relevance, the Venn fusion measure

mediated the effect of immediacy (i.e., means-ends association) on expected enjoyment of and positive experience from the activity (reflecting intrinsic motivation).

Participants in a third study (Woolley & Fishbach, 2017a) were informed that drinking a smoothie beverage would have an immediate or delayed effect (of providing an energy boost). As predicted, those in the immediate versus delayed condition expected to enjoy drinking the smoothie more (i.e., believed it would be tastier and would provide a more positive experience). The authors further measured the mental representation of the activity-goal fusion via the Venn diagram technique (cf. Aron et al., 1992). When drinking the smoothie (*the activity*) was said to provide immediate (vs. delayed) outcomes, participants associated it with receiving energy (*the goal*) to a greater degree. The perceptual overlap between the activity and the goal mediated the effect of immediacy (i.e., means-ends association) on expected enjoyment of drinking the smoothie (reflecting intrinsic motivation).

In summary, there is substantial evidence that the different antecedents of means-ends fusion (repeated association, links strength, similarity, and immediacy) induce the subjective experiences (perceived instrumentality, goal derived affect) and behavioral reactions (choice of, and persistence at, the activity) postulated in our model. Of particular importance, relations between the antecedents and consequences of means end fusion appear to be mediated by the fusion experience (e.g., as measured via the Venn diagram technique described above).

Goal Activation

According to our theory, the tendency of the fused activity to *exhibit* the properties of intrinsic motivation should vary as function of activation of the goal with which the activity is fused. Simply, for the activity to reflect the properties of its goal, the goal needs to be active to begin with and hence capable of having its properties manifest in the activity.

1. Satiation effects. Evidence for this proposition comes from diverse domains of psychological science including animal learning studies as well as social cognition experiments (cf. Redden, 2015). In the former domain, ample studies have found that the conditioned (i.e., goal-fused) activity is sluggish, much less pronounced (e.g., running speed is lower) or absent altogether when the animal is *satiated*, hence the goal that originally propelled the activity is inactive (Hillman, Hunter, & Kimble, 1953, Kintsch, 1962; Porter, Madison, & Senkowski, 1968; Sidman & Stebbins, 1954; Zaretsky, 1966). Research on satiation in humans finds similar results, with decreased goal-directed behavior occurring after repeated exposure to a stimulus such as a food, song, or social interaction (Redden, 2015).

Whereas the foregoing studies attest to the reduced *degree* of motivation attached to the conditioned (fused) activity when the relevant goal is inactive, other experiments on auto-shaping suggest that the *qualitative* features of the activity also fail to appear under satiation. In this vein, Cleland and Davey (1982) found that food satiation significantly reduced the effects of goal-related contact in rats toward the CS of retractable lever (that is, licking and pawing the lever). Similarly, Hilliard and Domjan (1995) found that male Japanese quails displayed significantly less goal-related approach to a visual CS when sexually satiated than when sexually deprived.

2. Momentary goal shifts. Robinson and Berridge (2013) showed that a repulsive stimulus can be transformed into an attractive one, depending on the goal being currently activated. In their study, rats learned repulsion toward a lever that always predicted an unpleasant saltiness sensation. The animals were retested following the injection of substances that mimic sodium deficiency brain signals. In this state they became immediately attracted by the previously aversive lever CS. The Robinson and Berridge (2013) study thus demonstrates that

the attraction (liking) toward a given activity, such as approaching a given CS, is a function of the goal that happens to be activated: If the activated goal is of the positive, desirable variety, the conditioned activity of approach will override that of repulsion that would appear when the goal is of the negative, undesirable variety.

A similar effect with humans was demonstrated in an experiment by Fishbach et al. (2004, Study 3) referred to earlier. Specifically, experimenters primed one out of two goals, *weight-watching* and *food enjoyment*, and examined the effects this had on the perception of *eating fattening foods*, an activity that may promote food enjoyment, but impede weight-watching. It was found that priming the goal of food enjoyment imbued eating high-calorie food with positive affect, whereas priming the goal of weight-watching imbued the very same activity with negative affect.

Other experiments with human subjects confirmed the notion that priming or activating a given goal prompts greater liking of an activity that is instrumental to that goal's attainment. In the participant's mind, the instrumental activity should be more strongly fused with the goal than a non-instrumental activity; thus, it should be higher in intrinsicity and, therefore, liked better than the non-instrumental activity. Accordingly, Ferguson and Bargh (2004) activated the achievement goal by inducing participants to care (achievement goal) or not care (control condition) about their performance in a word-creation game, which they were asked to play during the experiment. Next, participants evaluated objects instrumental to the achievement goal, either while they were still playing the game (unfinished condition) or after they had completed the game (finished condition). Results suggest that participants who were actively engaged in goal pursuit, compared with those who were not pursuing the goal, automatically evaluated goal-relevant objects as relatively more positive than goal-irrelevant objects. In similar research,

Fitzsimmons and Shah (2008) primed the *achievement* goal and found that participants felt closer to (i.e., liked more), and rated as more important, friends who were instrumental to achievement than friends who were not instrumental to that goal.

General Discussion

We reviewed evidence concerning various implications of our MEF model of intrinsic motivation. Specifically, there is support for the notion that each of the postulated antecedents of intrinsic motivation impacts the intrinsic experience of an activity. Thus, (1) repeated association between the activity and its end; (2) uniqueness of the association; (3) similarity of the activity to its goal and (3) immediacy of goal attainment following the activity—all induce fusion between an activity and its end and result in the experience of intrinsic motivation (e.g., Bélanger, 2017; Woolley & Fishbach, 2017a). Moreover, our MEF theory also identifies novel manifestations of intrinsic motivation (beyond traditionally studied engagement under free choice and positive experience). Specifically, we suggest that the fused activity is seen as instrumental to its goal's attainment, and as sharing in the various properties of the goal itself, including its motivating capability and its elicitation of specific goal-related affect (e.g., Fishbach et al., 2004). There is also evidence that the relation between fusion's antecedents and its consequences is mediated by the subjective experience of fusion. Finally, there is evidence that reacting to the fused activity in ways that resemble goal-reactions requires that the goal in question be activated.

The present view of intrinsic motivation has important implications. An intrinsically motivated activity enjoys several advantages: (1) It is accompanied by positive affect, or in the case of avoidance goals, the removal of negative affect (cf. Higgins, 1987) ; (2) it is experienced as *freely* initiated (c.f., Kruglanski & Cohen, 1973; Kruglanski, 1975) as the actor may feel that (s)he is “doing what (s)he wants”; (3) it is perceived as more instrumental to goal attainment,

hence (4) it inspires greater commitment (Kruglanski, Pierro, & Sheveland, 2011) to the activity and fosters (5) greater persistence at it (Woolley & Fishbach, 2016; 2017b) and resumption of the activity following interruption.

Of interest, data supportive of the foregoing assertions comes from widely dispersed psychological inquiries including animal learning studies, brain research as well as social cognition experiments with human participants. The compatibility of findings across a wide span of phenomena attests to the generality of motivational principles being investigated here, and their robustness.

The present analysis suggests procedures that promise to move any activity along the intrinsicity continuum, rendering its identification as more intrinsic or more extrinsic. This discovery could be of considerable practical utility. Thus, if one wished to intensify individuals' engagement with a given activity (e.g., because it was deemed desirable and socially valued), one could do so by following the fusion-enhancing procedures identified in the present theory (related to repeated association, linkage uniqueness, immediacy and similarity). To the contrary, if one wished to reduce activity engagement, one could implement the opposite, fusion-reducing processes.

In a study relevant to these notions, Milkman, Minson, and Volpp (2013) found that strengthening the association between immediate “fun” rewards (listening to exciting audiobooks) and more unpleasant activities (exercising at the gym) increased the likelihood that participants would engage in the latter activities. Their study had three conditions: one in which participants could listen to audiobooks *only* while exercising; a second condition in which participants could listen to audiobooks anytime, including during exercising, which rendered the link between exercising and fun *less unique* than in the former condition; and a third control

condition where no audio books were mentioned. They found that participants exercised more in the first (vs. second and third) condition, which, from the present perspective, represents the strongest fusion between the goal of “fun” and the activity of exercising.

Research by Higgins, Cesario, Hagivara, Spiegel and Pittman (2010, p. 571) provides another example of rendering an undesirable activity (consumption of junk food) as extrinsically motivated, hence lower on the intrinsicity dimension, by getting individuals to attribute the activity to a goal (getting a boost of energy) *dissimilar* from the activity itself.

This logic could be profitably applied to multiple pertinent domains. For example, in the realm of health, augmenting patients’ adherence to their medication regimen could be improved by repeatedly associating regimen adherence with primed thoughts about positive health outcomes. Similarly, in the realm of education, associating studying with thoughts about interesting careers through repeatedly coupling should increase the intrinsic appeal of studying (cf. Yeager et al., 2014). Such undesirable activities as substance abuse, or bullying could be discouraged by diluting their relations to the goals they are believed to serve through the strengthening of alternative links to those particular goals (e.g., the goals of relaxation, a sense of mattering and significance).

Functional autonomy? The concept of extinction and the ample evidence for it, are inconsistent with Allport’s (1937) theory of the functional autonomy of motives, whereby an activity previously associated with goal attainment may acquire independent motivating properties on its own. Researchers agree, however, that though such properties are present for a time, they ultimately disappear if the activity fails to result in goal attainment (e.g., Bertocci, 1940, Chan & Harris, 2017; McClelland, 1942; Pavlov, 1927; Seward, 1963; Rethlingshafer, 1943). As noted earlier, fused activities derive their motivational and affective properties from

the goals they serve; those properties are “on loan” as it were, and they dissipate once the link between the activity and the goal was demonstrably severed. Of course, a semblance of functional autonomy’ may exist when severance of the link between an activity and its goal was blurry or ambiguous, and when a different goal was attached to the activity and lent it its motivating properties.

Undermining intrinsic interest by extrinsic rewards. One of the most intriguing findings in research on intrinsic motivation phenomena was the discovery that intrinsic interest in an activity can be undermined by attaching to the activity an extrinsic reward of some sort; the over-justification effect. For instance, studies have shown that children’s interest in puzzle solving or drawing is undermined when the children are given a monetary reward for the activity (for reviews of such findings see Deci, 1975; Deci & Ryan, 1985; Lepper, Greene, & Nisbett, 1973; Kruglanski, Friedman, & Zeevi, 1971), and charitable donations are reduced when individuals are offered thank-you gifts for the donations (Ariely, Bracha, & Meier, 2009; Newman & Shen, 2012).

From the present perspective, linking a previously performed activity to a new, tangible, reward creates an association between an activity and an additional goal (i.e., the reward in question). Theoretically, such linkage should have two opposite effects: On the one hand, the added goal enhances the total value attached to the activity; in turn, such value “trickles down” or *transfers* to the activity enhancing its interest value or appeal (Fishbach et al. 2004). On the other hand, adding a goal to an activity weakens the strength of association between the activity and the original goal (cf. Zhang et al., 2007). So after the additional goal was removed (as in the “free choice” phase of intrinsic motivation studies), versus being left intact (Maimaran & Fishbach, 2014) the resultant interest in the activity should be determined by the degree to which

the diluted association between the activity and the original goal is overcompensated by the value transfer to the activity from the added (and now removed) goal. Specifically, if the decrement of interest produced by the *dilution effect* is greater than the increment produced by the *transfer effect* we should observe the classic “undermining of intrinsic interest” result. In contrast, if the transfer effect exceeded in magnitude the dilution effect we should observe, in fact, an augmentation in an activity’s interest.

Relevant to this analysis is the classic study by Harackiewicz, Manderlink, and Sansone (1984, Study 1). In their experiment, one third of the participants were offered a *performance-contingent reward*; one third anticipated performance evaluation, but were not promised nor given a reward (*evaluation-only*); and one third neither expected evaluation nor received a reward (*control*). All participants then received positive feedback upon completing the task; half the participants in each condition were informed that their performance level was above the 50th percentile, and the other half were told that their performance was above the 80th.

Consistent with the findings (cf. Zhang et al., 2007) that addition of a goal (here, obtaining a positive evaluation) dilutes the strength of association between the activity and the original goal (e.g. of fun), it was found that evaluation reduced interest, relative to the control group. Yet, the performance-contingent reward condition actually *enhanced* participants’ interest in the activity when compared to the evaluation-only group (though not significantly so when compared to the control group). Relatedly, Goswami & Urminsky (2017) found that although there is a robust decrease in engagement immediately after the incentive is removed, engagement returns to a post-reward baseline that equal or exceeds the initial baseline.

Why might these effects have occurred? From the present perspective, adding a reward to the evaluation might have imbued the evaluation goal with greater importance than was the case

in the evaluation only condition. In turn, enhanced magnitude of the evaluation goal (in the evaluation + reward condition) may have transferred to the activity, thus increasing its interest value as compared to the evaluation only condition. These effects are graphically portrayed in Figure 7 (a, b). Of course, removing the reward in this case might not have severed clearly enough the association between the activity and the demonstration of competence, allowing the transfer effect to be present. Admittedly, the present interpretation is *post hoc* hence in need of its further validation through additional research.

Figure 7 here

One implication of the present, structural, analysis is that interest undermining (or augmenting) effects need not be restricted to traditional “extrinsic” rewards (e.g., monetary pay, or praise) but could also be produced by alerting actors performing an activity for a given reason to an additional intrinsic reason for performing it, that is, an additional goal that the activity could accomplish (cf. Higgins & Trope, 1990). If that second feature was then subtly removed, individuals should display a lowering or heightening of interest in pursuing the activity depending on the tradeoff between the dilution and the transfer effects as discussed above.

In a study specifically demonstrating the lowering effect, children became less interested in reading when the reading was previously combined with coloring (Higgins et al., 1995), both representing highly fused (and hence “intrinsic”) ends of the activity. This finding supports the present claim that the addition of any reward (or goal) might reduce the motivation to perform the activity for the original end by diluting the association between the activity and that end. However, from the present perspective, a less blatant removal of the additional goal, might have allowed it to continue imbuing the activity with interest proportionately to the interest/value-transfer from that particular goal to the activity.

Over two decades ago, a heated controversy erupted over the question whether rewards enhance interest in an activity (cf. Eisenberger, 1992; Eisenberger & Cameron, 1996; Eisenberger & Shanock, 2003) or depress it (Deci et al., 1999; Lepper, Henderlong & Gingras, 1999). In conclusion to a review volume of empirical research on this question, Sansone and Harackiewicz (2000, pp. 443-444) aptly commented that “It depends [as] rewards can have a variety of effects on intrinsic motivation and performance [and] that these...depend on the nature of the activity, the reward contingency, the feedback obtained, the more general context of reward administration and the people offering and receiving the reward.” Eisenberger and Shanock (2003) concurred in noting that “Three decades of research have failed to produce general agreement concerning the effects of reward” (p. 121). The present MEF theory addresses this issue and offers a general framework within which the question of reward effects on activity-interest may be usefully considered, specifically in terms of the tradeoffs between the dilution and the transfer mechanisms elaborated earlier.

Coda

An intrinsically motivated activity is generally understood as one that constitutes an end in itself. Building on this definition, we presented a structural model of intrinsic motivation that portrays it in terms of perceptual fusion between an activity and its end. This conception envisages an intrinsicity continuum, ranging from fully extrinsic to fully intrinsic relations between an activity and its end. Drawing on Gestalt theories of perception, classical conditioning theory, and contemporary research on human motivation, we identified several antecedent conditions as well as consequences of intrinsic motivation. Our model explains key findings in the realm of intrinsic motivation (e.g., the undermining of intrinsic interest by external rewards), and offers a guiding framework for further research on intrinsic motivation. It also proffers

suggestions for possible interventions designed to alter activity engagement in desirable ways.

References

- Allport, G. W. (1937). The functional autonomy of motives. *The American Journal of Psychology*, 50(1), 141-156.
- Amabile, T. M., Hennessey, B. A., & Grossman, B.S. (1986). Social Influences on Creativity: The Effects of Contracted-for Reward. *Journal of Personality and Social Psychology*, 50(1), 14-23.
- Amsel, A. (1949). Selective association and the anticipatory goal response mechanism as explanatory concepts in learning theory. *Journal of Experimental Psychology*, 39(6), 785-799.
- Amsel, A. (1962). Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychological Review*, 69(4), 306-328.
- Anderson, J. R., & Bower, G.H. (1973). *Human Associative Memory*. Washington, D.C: Winston.
- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 261-295.
- Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *The American Economic Review*, 99(1), 544-555.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596-612.
- Babin, B. J., Darden, W. R., & Griffin, M. (1994). Work and/or fun: measuring hedonic and utilitarian shopping value. *Journal of Consumer Research*, 20(4), 644-656.
- Balsam, P. D., Drew, M. R., & Gallistel, C. R. (2010). Time and associative learning. *Comparative Cognition & Behavior Reviews*, 5, 1-22.
- Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, 81(6), 1014-1027.
- Bélanger, J.J. (2017). Means-Ends fusion and intrinsic motivation. Unpublished data. New York University, Abu Dhabi.
- Bélanger, J. J., Schori-Eyal, N., Pica, G., Kruglanski, A. W., & Lafrenière, M. A. (2015). The “more is less” effect in equifinal structures: Alternative means reduce the intensity and quality of motivation. *Journal of Experimental Social Psychology*, 60, 93-102.
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York, NY: McGraw-Hill Book Company.
- Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: ‘liking’, ‘wanting’, and learning. *Current Opinion in Pharmacology*, 9(1), 65-73.
- Bertocci, P. A. (1940). A critique of GW Allport's theory of motivation. *Psychological Review*, 47(6), 501-532.
- Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, 11(1), 1-8.
- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Systems Research and Behavioral Science*, 3(1), 14-25.
- Chan, C. J., & Harris, J. A. (2017). Extinction of Pavlovian Conditioning: The Influence of Trial Number and Reinforcement History. *Behavioural Processes*.

- Cleland, G. G., & Davey, G. C. (1982). The effects of satiation and reinforcer devaluation on signal-centered behavior in the rat. *Learning and Motivation, 13*(3), 343-360.
- Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: on the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology, 89*(2), 129.
- Day, H. I., Berlyne, D. E., & Hunt, D. E. (1971). *Intrinsic motivation: A new direction in education*. Toronto, Canada: Holt, Rinehart & Winston.
- Deci, E. L. (1975). *Intrinsic motivation*. New York, NY: Plenum.
- Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality, 19*(2), 109-134.
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627-668.
- De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*(6), 853-869.
- Eisenberger, R. (1992). Learned industriousness. *Psychological Review, 99*, 248-267.
- Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: Reality or myth?. *American Psychologist, 51*(11), 1153-1166.
- Eisenberger, R., & Shanock, L. (2003). Rewards, intrinsic motivation, and creativity: A case study of conceptual and methodological isolation. *Creativity Research Journal, 15*, 121-130.
- Elliot, A. J., & Harackiewicz, J. M. (1994). Goal setting, achievement orientation, and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 66*(5), 968-980.
- Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Sciences, 8*(1), 33-39.
- Fishbach, A., Shah, J. Y., & Kruglanski, A. W. (2004). Emotional transfer in goal systems. *Journal of Experimental Social Psychology, 40*(6), 723-738.
- Fitzsimons, G. M., & Shah, J. Y. (2008). How goal instrumentality shapes relationship evaluations. *Journal of Personality and Social Psychology, 95*(2), 319-337.
- Freitas, A. L., & Higgins, E. T. (2002). Enjoying goal-directed action: The role of regulatory fit. *Psychological Science, 13*(1), 1-6.
- Gibbon, J., Baldock, M. D., Locurto, C., Gold, L., & Terrace, H. S. (1977). Trial and intertrial durations in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 3*(3), 264-284.
- Gormezano, I., & Kehoe, E. J. (1981). Classical conditioning and the law of contiguity. *Advances in Analysis of Behavior, 2*, 1-45.
- Goswami, I., & Urminsky, O. (2017). The dynamic effect of incentives on postreward task engagement. *Journal of Experimental Psychology: General, 146*(1), 1-19
- Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion, 24*(3), 175-213.
- Guthrie, E. R. (1940). Association and the law of effect. *Psychological Review, 47*(2), 127-148.
- Harackiewicz, J. M., Manderlink, G., & Sansone, C. (1984). Rewarding pinball wizardry: Effects of evaluation and cue value on intrinsic interest. *Journal of Personality and Social Psychology, 47*(2), 287-300.

- Harlow, H. F., Dodsworth, R. O., & Harlow, M. K. (1965). Total social isolation in monkeys. *Proceedings of the National Academy of Sciences*, 54(1), 90-97.
- Heath, C. (1999). On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organizational Behavior and Human Decision Processes*, 78(1), 25-62.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.
- Helson, H. (1933). The fundamental propositions of Gestalt psychology. *Psychological Review*, 40(1), 13.
- Herzberg, F., Mausner, B., & Snyderman, B. B. (1959). *The motivation to work*. New York, NY: John Wiley & Sons.
- Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychological Review*, 94(3), 319-340.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E.T. Higgins & A.W. Kruglanski (Eds.), *Social Psychology: Handbook of Basic Principles*. New York, NY: Guilford.
- Higgins, E. T. (2000). Making a good decision: value from fit. *American Psychologist*, 55(11), 12-17.
- Higgins, E. T. (2012). *Beyond pleasure and pain: How motivation works*. New York, NY: Oxford University Press.
- Higgins, E. T., Cesario, J., Hagiwara, N., Spiegel, S., & Pittman, T. (2010). Increasing or decreasing interest in activities: The role of regulatory fit. *Journal of Personality and Social Psychology*, 98, 559-572.
- Higgins, E. T., Lee, J., Kwon, J., & Trope, Y. (1995). When combining intrinsic motivations undermines interest: a test of activity engagement theory. *Journal of Personality and Social Psychology*, 68(5), 749-767.
- Higgins, E. T., & Trope, Y. (1990). Activity engagement theory: Implications of multiply identifiable input for intrinsic motivation. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior*, 2, (pp. 229-264). New York, NY: Guilford Press.
- Hilgard, E. R., & Bower, G. H. (1966). *Theories of learning*. New York, NY: Appleton-Century-Crofts.
- Hilliard, S., & Domjan, M. (1995). Effects on sexual conditioning of devaluing the US through satiation. *The Quarterly Journal of Experimental Psychology*, 48(1), 84-92.
- Hillman, B., Hunter, W. S., & Kimble, G. A. (1953). The effect of drive level on the maze performance of the white rat. *Journal of Comparative and Physiological Psychology*, 46(2), 87-89.
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychological Bulletin*, 136(3), 390-421.
- Holyoak, K. J. (1990). Problem solving. *Thinking: An Invitation to Cognitive Science*, 3, 117-146.
- Holyoak, K. J., Kohl, K., & Nisbett, R. E. (1989). A theory of conditioned reasoning: Inductive learning within rule-based hierarchies. *Psychological Review*, 96, 315-340.
- Hsiao, H. H. (1928). A suggestive review of Gestalt psychology. *Psychological Review*, 35(4), 280.
- Hull, C. L. (1931). Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review*, 38(6), 487-506.

- Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. New York, NY: Appleton-Century-Crofts.
- Kamin, L.J. (1969). Predictability, surprise, attention and conditioning. In B.A. Campbell & R.M. Church (Eds.), *Punishment and Aversive Behavior* (pp. 279–96). New York, NY: Appleton-Century-Crofts.
- Kendler, H. H. (1946). The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *Journal of Experimental Psychology*, 36(3), 212-220.
- Kintsch, W. (1962). Runway performance as function of drive strength and magnitude of reinforcement. *Journal of Comparative and Physiological Psychology*, 55(5), 882-887.
- Kohler, W. (1940). *Dynamics in psychology: Vital applications of gestalt psychology*. New York, NY: Liveright.
- Kruglanski, A. W. (1975). The endogenous-exogenous partition in attribution theory. *Psychological Review*, 82(6), 387-406.
- Kruglanski, A. W., Alon, S., & Lewis, T. (1972). Retrospective misattribution and task enjoyment. *Journal of Experimental Social Psychology*, 8(6), 493-501.
- Kruglanski, A. W., Chernikova, M., Rosenzweig, E., & Kopetz, C. (2014). On motivational readiness. *Psychological Review*, 121(3), 367-388.
- Kruglanski, A. W., & Cohen, M. (1973). Attributed freedom and personal causation. *Journal of Personality and Social Psychology*, 26(2), 245-250.
- Kruglanski, A. W., Friedman, I., & Zeevi, G. (1971). The effects of extrinsic incentive on some qualitative aspects of task performance. *Journal of Personality*, 39(4), 606-617.
- Kruglanski, A. W., Pierro, A., & Sheveland, A. (2011). How many roads lead to Rome? Equifinality set - size and commitment to goals and means. *European Journal of Social Psychology*, 41(3), 344-352.
- Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y., & Sleeth-Keppler, D. (2002). A theory of goal systems. *Advances in Experimental Social Psychology*, 34, 331-378.
- Lepper, M. R., Greene, D., & 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), 129-137.
- Lepper, M. R., & Greene, D. (Eds.). (1978). *The hidden costs of reward: New perspectives on the psychology of human motivation*. Hillsdale, NJ: Lawrence Erlbaum.
- Lepper, M. R., Henderlong, J., & Gingras, I. (1999). Understanding the effects of extrinsic rewards on intrinsic motivation—Uses and abuses of meta-analysis. *Psychological Bulletin*, 125(6), 669-676.
- Levey, A. B., & Martin, I. (1975). Classical conditioning of human 'evaluative' responses. *Behaviour research and therapy*, 13(4), 221-226.
- Lin, H. (2007). Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions. *Journal of Information Science*, 32(2), 135-149.
- Maimaran, M., & Fishbach, A. (2014) If it's useful and you know it, do you eat? Preschoolers refrain from instrumental food. *Journal of Consumer Research*, 41, 642-655.
- McAllister, W. R. (1953). Eyelid conditioning as a function of the CS-US interval. *Journal of Experimental Psychology*, 45(6), 417-422.
- McClelland, D. C. (1942). Functional autonomy of motives as an extinction phenomenon. *Psychological Review*, 49, 272-283.

- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227-234.
- Milkman, K. L., Minson, J. A., & Volpp, K. G. (2013). Holding the Hunger Games hostage at the gym: An evaluation of temptation bundling. *Management Science*, *60*(2), 283-299.
- Motivation. (n.d.). In *Wikipedia*. Retrieved November 6, 2017, from https://en.wikipedia.org/wiki/Motivation#Intrinsic_motivation
- Newman, G. E., & Shen, Y. J. (2012). The counterintuitive effects of thank-you gifts on charitable giving. *Journal of Economic Psychology*, *33*(5), 973-983.
- Ost, J. W. P., & Lauer, D. W. (1965). Some investigations of classical salivary conditioning in the dog. In *Classical Conditioning: A Symposium*. New York: Appleton-Century-Crofts, 192-207.
- Pavlov, I. P. (1927). *Conditioned reflexes*. Mineola, NY: Dover Publications, Inc.
- Porter, J. J., Madison, H. L., & Senkowski, P. C. (1968). Runway performance and competing responses as functions of drive level and method of drive measurement. *Journal of Experimental Psychology*, *78*(21), 281-284.
- Premack, D. (1962). Reversibility of the reinforcement relation. *Science*, *136*(3512), 255-257.
- Razran, G. (1954). The conditioned evocation of attitudes (cognitive conditioning?). *Journal of Experimental Psychology*, *48*(4), 278-282.
- Redden, J. P. (2015). Desire over time: the multi-faceted nature of satiation. In W. Hofmann & L. F. Nordgren (Eds.), *The Psychology of Desire* (pp. 82-103). New York, NY: Guilford Press.
- Rescorla, R. A. (1980). Simultaneous and successive associations in sensory preconditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *6*(3), 207-216.
- Rescorla, R. A. (1985). Pavlovian conditioning analogues to Gestalt perceptual principles. *Affect, conditioning, and cognition: Essays on the determinants of behavior*, 113-130.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, *2*, 64-99.
- Rethlingshafer, D. (1943). Experimental evidence for functional autonomy of motives. *Psychological Review*, *50*(4), 397-407.
- Reynolds, B. (1945). The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace. *Journal of Experimental Psychology*, *35*(1), 15-30.
- Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Research Reviews*, *18*(3), 247-291.
- Robinson, M. J., & Berridge, K. C. (2013). Instant transformation of learned repulsion into motivational "wanting". *Current Biology*, *23*(4), 282-289.
- Ryan, R. M., & Deci, E. L. (2000) Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, *55*(1) 68-78.
- Sansone, C., & Harackiewicz, J. M. (1998). "Reality" is complicated. *American Psychologist*, *53*(6), 673-674.
- Sansone, C., & Harackiewicz, J. M. (2000). *Intrinsic and Extrinsic Motivation: The Search of Optimal Motivation and Performance*. San Diego, CA: Academic Press.
- Sansone, C., Sachau, D. A., & Weir, C. (1989). Effects of instruction on intrinsic interest: The importance of context. *Journal of Personality and Social Psychology*, *57*(5), 819-829.

- Seward, J. P. (1963). The structure of functional autonomy. *American Psychologist*, 18(11), 703-710.
- Sidman, M., & Stebbins, W. C. (1954). Satiation effects under fixed-ratio schedules of reinforcement. *Journal of Comparative and Physiological Psychology*, 47(2), 114-116.
- Smith, M. C. (1968). CS-US interval and US intensity in classical conditioning of the rabbit's nictitating membrane response. *Journal of Comparative and Physiological Psychology*, 66, 679-687.
- Solomon, R. L., Kamin, L. J., & Wynne, L. C. (1953). Traumatic avoidance learning: the outcomes of several extinction procedures with dogs. *The Journal of Abnormal and Social Psychology*, 48(2), 291-302.
- Spence, K. (1956). *Behavior theory and conditioning*. New Haven, CT: Yale University Press.
- Srull, T. K. (1981). Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory*, 7(6), 440-463.
- Stein, L., Sidman, M., & Brady, J. V. (1958). Some effects of two temporal variables on conditioned suppression. *Journal of the Experimental Analysis of Behavior*, 1(2), 153-162.
- Tang, S. H., & Hall, V. C. (1995). The overjustification effect: a meta-analysis. *Applied Cognitive Psychology*, 9(5), 365-404.
- Timberlake, W., & Allison, J. (1974). Response deprivation: An empirical approach to instrumental performance. *Psychological Review*, 81(2), 146-164.
- Tolman, E. C. (1933). Sign-Gestalt or conditioned reflex. *Psychological Review*, 40(3), 246-255.
- Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, 110(3), 403-421.
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440-463.
- Vallacher, R. R., & Wegner, D. M. (1985). *A theory of action identification*. Hillsdale, NJ: Lawrence Erlbaum.
- Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, 94(1), 3-15.
- Vallerand, R. J. (2007). Intrinsic and extrinsic motivation in sport and physical activity: A review and a look at the future, In G. Tenenbaum & E. Eklund (Eds.), *Handbook of Sport Psychology* (3rd ed., pp. 59-83). New York, NY: John Wiley.
- Vallerand, R. J. & Losier, G. F. (1999). An integrative analysis of intrinsic and extrinsic motivation in sport. *Journal of Applied Sport Psychology*, 11, 142-169.
- Vandercar, D. H., & Schneiderman, N. (1967). Interstimulus interval functions in different response systems during classical discrimination conditioning of rabbits. *Psychonomic Science*, 9(1), 9-10.
- Vroom, V. H. (1964). *Work and motivation*. Oxford, UK: Wiley.
- Wasserman, E. A., Franklin, S. R., & Hearst, E. (1974). Pavlovian appetitive contingencies and approach versus withdrawal to conditioned stimuli in pigeons. *Journal of Comparative and Physiological Psychology*, 86(4), 616-627.
- Wertheimer, M. (1923). Laws of organization in perceptual forms. In *A source book of Gestalt psychology* (pp. 71-88). London, UK: Routledge & Kegan Paul.
- White, R. W. (1959). Motivation Reconsidered: The Concept of Competence. *Psychological Review*, 66(5), 297-333.

- Wickens, D. D., Meyer, P. M., & Sullivan, S. N. (1961). Classical GSR conditioning, conditioned discrimination, and interstimulus intervals in cats. *Journal of Comparative and Physiological Psychology*, *54*(5), 572-576.
- Wolfe, J. B. (1936). Effectiveness of token rewards for chimpanzees. *Comparative Psychology Monographs*, *12*(60), 1-72.
- Woolley, K., & Fishbach, A. (2015). The Experience Matters More Than You Think: People Value Intrinsic Incentives More Inside Than Outside an Activity. *Journal of Personality and Social Psychology*, *109*, 968-982.
- Woolley, K., & Fishbach, A. (2016). For the Fun of It: Harnessing Immediate Rewards to Increase Persistence on Long-Term Goals. *Journal of Consumer Research*, *42*, 952-966.
- Woolley, K., & Fishbach, A. (2017a). *Its about time: Earlier rewards increase intrinsic motivation*. (Working paper).
- Woolley, K., & Fishbach, A. (2017b). Immediate Rewards Predict Adherence to Long-Term Goals. *Personality and Social Psychology Bulletin*, *43*, 151-162.
- Wrzesniewski, A., Schwartz, B., Cong, X., Kane, M., Omar, A., & Kolditz, T. (2014). Multiple types of motives don't multiply the motivation of West Point cadets. *Proceedings of the National Academy of Sciences*, *111*(30), 10990-10995.
- Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology*, *107*(4), 559-580.
- Yerkes, R. M., & Yerkes, A. W. (1929). *The great apes*. New Haven, CT: Yale University Press.
- Zaretsky, H. H. (1966). Learning and performance in the runway as a function of the shift in drive and incentive. *Journal of Comparative and Physiological Psychology*, *62*(2), 218-221.
- Zhang, Y., Fishbach, A., & Kruglanski, A. W. (2007). The dilution model: how additional goals undermine the perceived instrumentality of a shared path. *Journal of Personality and Social Psychology*, *92*(3), 389-401.
- Zhang, Y., & Gao, L. (2016). Wanting ever more: Acquisition procedure motivates continued reward acquisition. *Journal of Consumer Research*, *43*(2), 230-245.

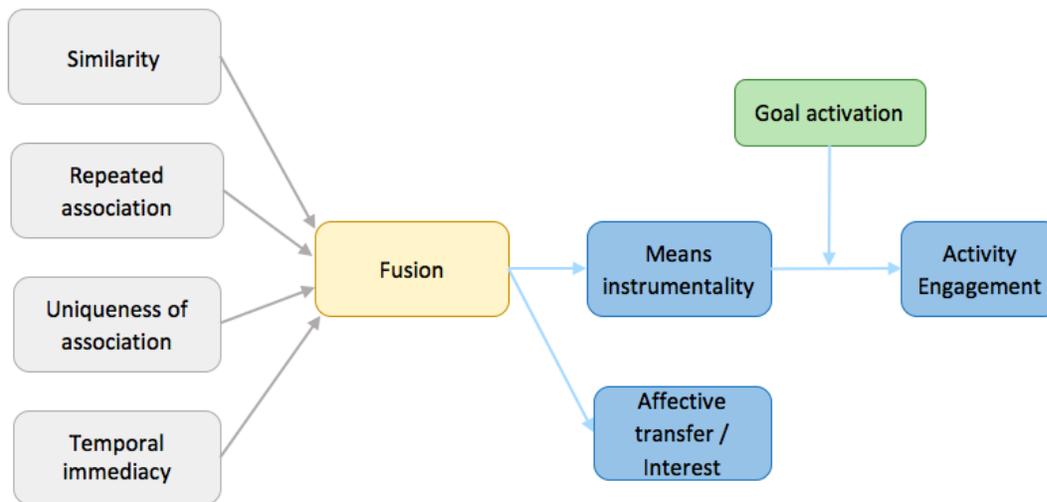


Figure 1. Theoretical model. Gray shapes indicate antecedents of fusion; blue shapes indicate consequences of fusion.

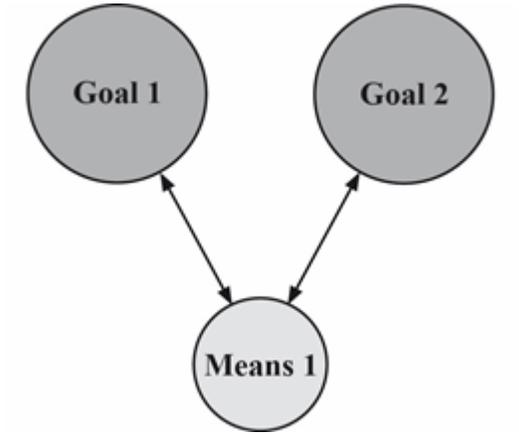


Figure 2. Multifinality configuration.

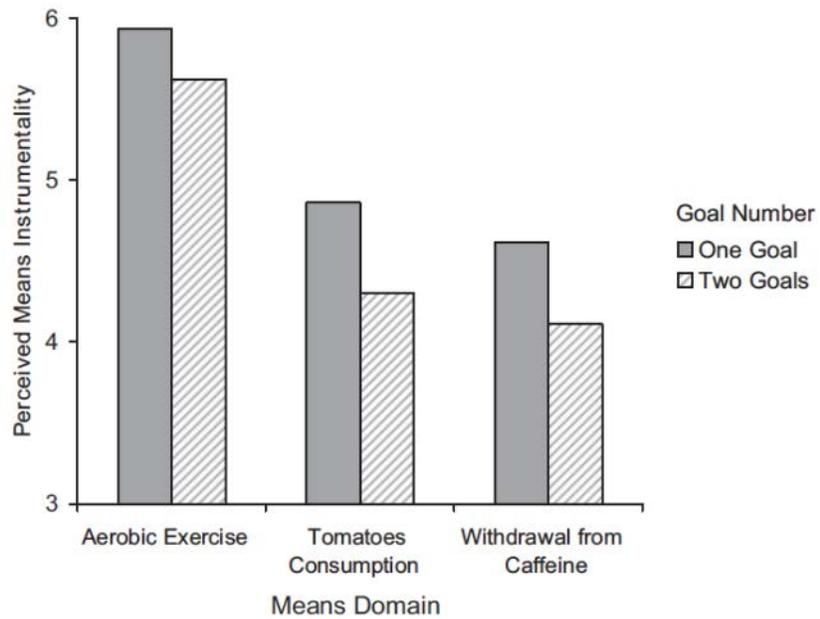


Figure 3. Perceived instrumentality of means for the first listed goal as a function of goal number (Zhang et al., 2007).

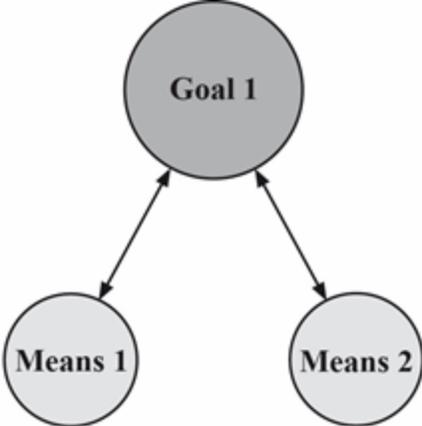


Figure 4. Equifinality configuration.

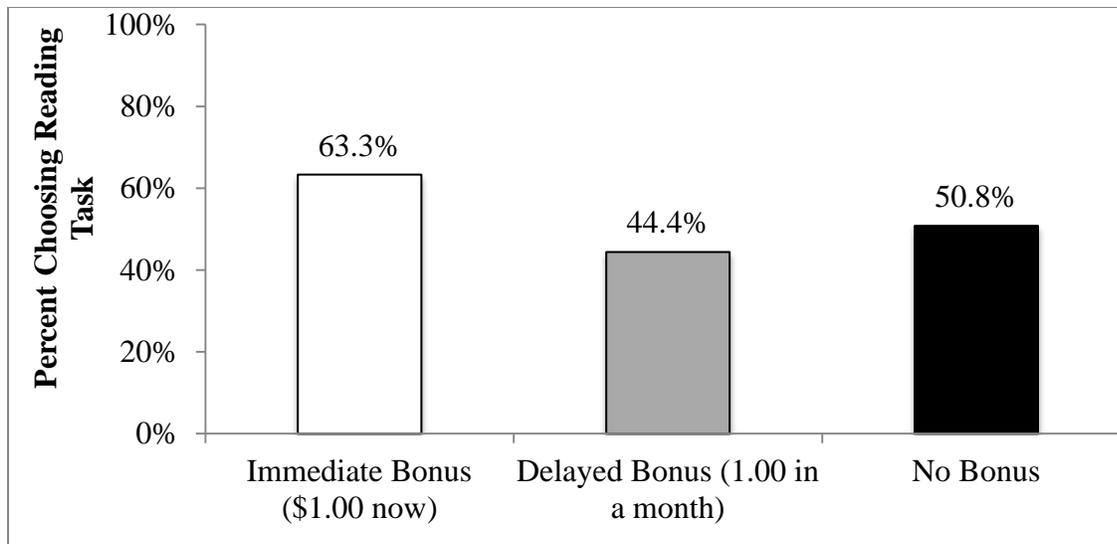


Figure 5. Task engagement during free choice, after the bonus was removed (Woolley & Fishbach, 2017a).

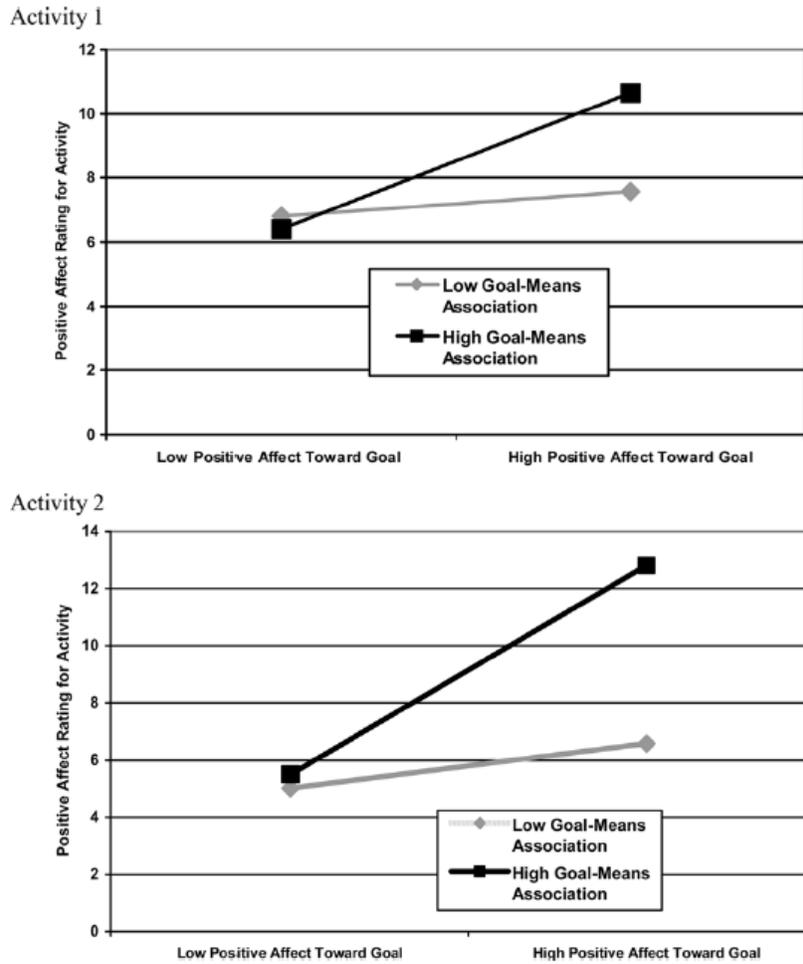


Figure 6. Positive affect toward activity as a function of the strength of the goal–activity association and positive affect toward goal (Fishbach et al., 2004).

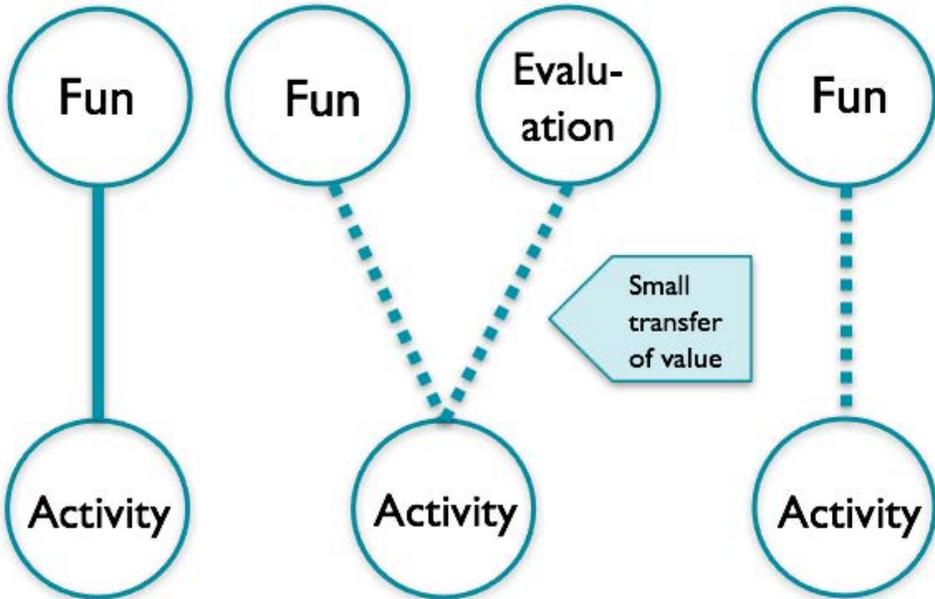


Figure 7a. Undermining intrinsic interest with extrinsic rewards.

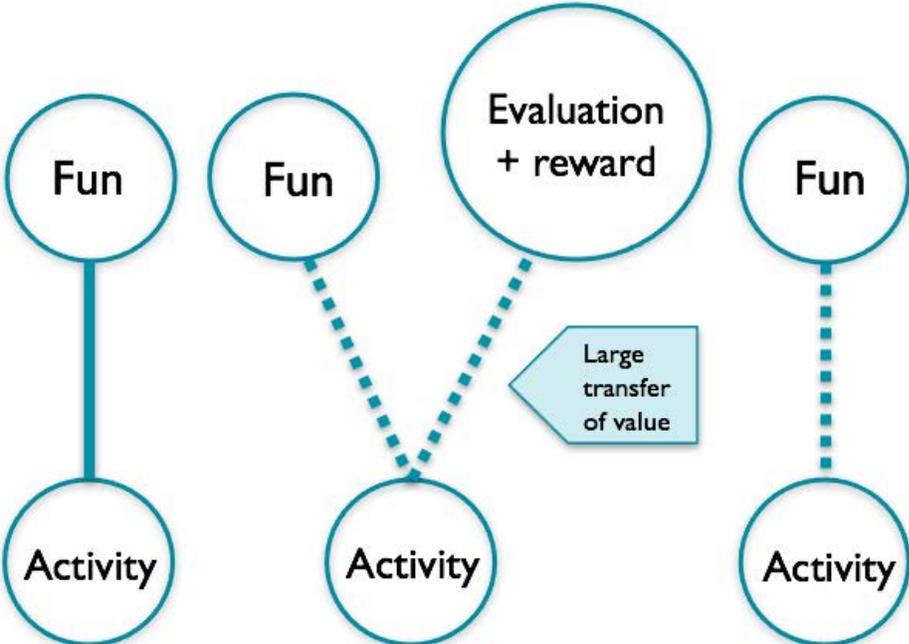


Figure 7b. Augmenting intrinsic interest with extrinsic rewards.