

# Transitions to Safety in “Obsessive” Thinking: Covert Links of Operant Chains in Clinical Behavior Analysis

Robert C. Mellon

Panteion University of Social and Political Sciences

This article provides a tentative account of a much-discussed clinical phenomenon called “thought-action fusion” and its relation to “obsessive” thinking, in terms of a general process interpretation of the determination of operant and respondent behavior derived from experimental analysis. In an interpretive approach originated by B. F. Skinner (e.g., 1953; 1957) the present account invokes plausible, untested causal relations among response-produced private events and behavior occurring in operant chains, and illustrates the more general utility of this practice for the understanding problematic thinking in clinical behavior analysis. Finally, a case is made for the potential therapeutic efficacy of providing simplified versions of such natural science-based interpretations of the determination of troublesome patterns of covert behavior to psychotherapy clients themselves.

*Key words:* private events, automatic reinforcement, aversive control, adventitious reinforcement, safety signal, thought-action fusion

---

A vast number of people worldwide are sorely troubled by what many clinical psychologists call “dysfunctional cognition.” The efforts of behavior analysts in interpreting the determination of the diverse and seemingly discontinuous behavioral phenomena so named have been both disproportionately limited and, of late, generally based in a “post-Skinnerian” contextualism (e.g., Hayes et al., 2006). The present work is an effort to illustrate the heretofore underutilized power of the molecular interpretative approach originated by B. F. Skinner (e.g., 1953; 1957) in understanding problematic

thinking and in the efficient design of conditions that alter its course. This approach might be applied in interpreting a wide range of forms of problematic cognition, from “paranoid ideation” to “depressive rumination” to “distorted body imaging,” but for illustrative purposes the present discussion will focus on the phenomenon known as “thought-action fusion” and its relation to “obsessive” thinking.

## Dimensions of Thought-Action Fusion

In mainstream clinical psychology, the term “thought-action fusion” names a generalized tendency to equate thinking or imagining the performance of socially censured acts with the overt emission of the acts imagined (see Berle & Starcevic, 2005; Shafran & Rachman, 2004 for reviews).

---

This analysis was presented in an invited address to the 6<sup>th</sup> Conference of the European Association for Behavior Analysis in Lisbon, Portugal, in September of 2012.

The author is indebted to Richard L. Shull for his most helpful commentary on a manuscript version of the present article.

Correspondence should be addressed to Robert Mellon, Department of Psychology, Panteion University of Social and Political Sciences, 136 Syngrou Avenue, Athens, Hellenic Republic, 176 71. E-mail: mellon.robert@gmail.com.

Thought-action fusion is said to occur in two highly related forms, termed "moral" and "likelihood" fusion.

In "moral" thought-action fusion, the covert emission of acts regarded as unethical is considered by the actor to be morally equivalent to their overt emission; for example, a person might believe his or her own fantasizing about an illicit sexual relation to be as reprehensible as performing the carnal act. In so-called "likelihood" thought-action fusion, the act of thinking of or imagining harmful or "forbidden" events is held to increase the probability that they will occur; examples would include the belief that imagining someone's untimely death puts the object of the fantasy at risk for a similar tragedy, or the conviction that thinking about oneself performing censured sex acts contributes to their eventual overt occurrence. Given that imagining grievous events or censured actions is believed to cause harm directly or to contribute to the eventual emission of corresponding overt practices, "likelihood" thought-action fusion beliefs also entail a conviction of the immorality of imagining acts that would be subject to social punishment if emitted overtly. Thus, in both "likelihood" and "moral" thought-action fusion, covert and socially-censured overt acts of common form or content are regarded by the actor in question as having equivalent ethical repercussions.

One might expect to observe a low frequency of thinking or fantasizing about harmful or taboo activities and events in persons who believe that covert acts are morally equivalent to (and/or hasten the occurrence of) their overt emission. Surprisingly, the opposite is true: thought-action fusion beliefs occur far more frequently in persons whose behavior satisfies the diagnostic criteria for obsessive-compulsive disorder and other obsessive-compulsive spectrum disorders such as anorexia nervosa (e.g., Coelho, Carter, McFarlane & Polivy, 2008). These criteria include the *repetitive and emotionally disturbing covert emission of*

*taboo acts*, patterns of thinking described as "unwanted" by the persons themselves and termed "obsessive" by clinical psychologists. Such persons frequently report failed attempts at suppressing such vexing thought patterns; indeed, empirical evidence indicates that efforts to control "unwanted" thinking can lead to *increases* in its frequency of emission (see Purdon, 2004, for a review). Why would thinking about prohibited activities occur *more often* in people who believe that there is little or no difference in the character and the consequences of covert and overt misbehavior?

In contemporary cognitive-behavioral clinical psychology, thought-action fusion is regarded as a form of *dysfunctional cognition*; the person is said to "place an undue significance on the meaning of unwanted intrusive thoughts" (Shrafran & Rachman, 2004, p. 106) and such interpretations are themselves treated as a *causal factor* in obsessive-compulsive spectrum disorders (e.g., Salkovskis, 1999). However, as often happens in the causal attribution of action to thought, the determination of these same, supposedly dysfunctional interpretations remains inadequately analyzed. We might well ask: Why is the publically undetectable and presumably harmless covert performance of dangerous or censured actions granted "undue significance" by these persons, and why are these covert acts repeated so often when their performance seems only to frighten, shame or otherwise disturb the actor?

In accord with a long-established tradition in behavior analysis (e.g., Sidman, 1960), the present account rejects the view that the phenomenon of thought-action fusion in "obsessive" thinking is dysfunctional or pathological *per se*; rather, it is regarded as a problematic instance of a more general and familiar adaptive process. More controversially, the account invokes a series of inferred events and processes. Specifically, the present interpretation is based on the assumption that the biobehavioral events produced by covert acts can elicit autonomic

arousal, reinforce and/or punish the acts that produce them, and evoke the subsequent emission of other covert and overt acts (e.g., Skinner, 1953; 1974). Although the processes inferred are all of a kind whose existence has been well established through experimental investigation, some behavior analysts (particularly those viewing operant behavior as determined by molar relations between the frequency of occurrence of acts and reinforcing events) have rejected any appeal to private events in the efficacious interpretation of behavioral phenomena (e.g., Baum, 2011a, b). While such “molar” perspectives offer viable accounts of overt behavior, accounts that are, in general, empirically compatible with “molecular” views (e.g., Dinsmoor, 2001; Hineline, 2001), it will be argued that a disregard for the role of covert response-produced events renders the scientific interpretation of important, predominately private phenomena (such as thought-action fusion in obsessive thinking) unworkable, limiting the utility of such accounts in clinical behavior analysis.

Before introducing a corresponding covert case, the role of response-produced events in the determination of *overt* behavior will be examined; the description of this process will then be extended in the interpretation of privately-observable behavior with analogous features.

### **Response-Produced Events in Behavior Analysis**

As Hefferline (1958; 1962) pointed out a half century ago, the effects of response-produced events in eliciting or evoking subsequent overt responses have long played a central role in behavioral interpretations of the determination of a range of complex phenomena. For example, Ferster and Skinner (1957) introduced their account of the temporal patterning of responding under schedules of reinforcement as follows:

The primary purpose of the present book is to present a series of experiments designed to evaluate the extent to which an organism's behavior enters into the determination of its subsequent behavior. From a formulation of such results we should be able to predict the effects of any schedule (p. 3; see Keller & Schoenfeld, 1950 and Skinner, 1953 for additional early examples).

Response-produced events are also linchpins in behavior-analytic interpretations of cognition. For example, the evocation of *precurrent* behavior, that is, the production of discriminative stimuli which enter into the control of subsequent acts of the individual who produced them, is the central process in behavioral analyses of problem-solving. Skinner's interpretation of such phenomena, in which he develops his conception of “rule-governed” behavior, begins as follows:

The behavior observed when a man solves a problem is distinguished by the fact that it changes another part of his behavior and is reinforced and strengthened when it does so (Skinner, 1966/1969, p. 133; see also Skinner, 1953, pp. 246-252).

In a similar fashion, the (typically unrecorded) self-production of contingency-specifying or instructional stimuli as determinants of subsequent responding is central to some interpretations of “emergent” conditional discriminations as manifestations of more fundamental behavioral processes (e.g., Palmer, 2004 a, b; Lowenkron, 1998; 2004). Behavioral accounts of perceptual phenomena have also made interpretative use of plausible but inferred internal stimulus and behavioral events, as when the stimuli produced by an “observing,” “perceptual,” or “attending” response are said to evoke subsequent discriminative acts such as reporting (e.g., Dinsmoor, 1985; Goldiamond, 1962; Nevin, 2008; Schoenfeld & Cumming, 1963).

## Response-Produced Events in Aversive Control

Of particular significance to an interpretation of problematic cognition, response-produced events have played a prominent role in useful interpretations of behaviors that function to forestall forthcoming presentations of aversive (that is, negatively reinforcing) stimuli. How, otherwise, might the *non-occurrence* of aversive events increase the emission probability of avoidance responding? Except in special circumstances, such non-occurrences would adventitiously follow the emission of so many different acts, on so many different occasions, that their effect on the emission probability of a response form that actually cancels impending aversive events would be negligible (see, e.g., Mowrer & Lamoreaux, 1942).

In a contemporary variant of the traditional two-factor views developed to address this issue, Dinsmoor (1977; 2001a, b) proposed that in avoidance procedures, stimuli automatically produced in the emission of response forms *incompatible* with effective avoidance acquire *negative* reinforcing potency because they are differentially associated with the *presentation* of aversive stimuli (see also Dinsmoor, 1954; Schoenfeld, 1950; Sidman, 1953; Skinner, 1953) whereas stimuli automatically produced in the emission of the *avoidance* response acquire *positive* reinforcing potency due to their differential association with the *termination* of aversive stimuli (see also Dinsmoor & Sears, 1973; Schoenfeld, 1950; but see, e.g., Herrnstein & Hineline, 1966; Hineline, 2001)<sup>1</sup>.

<sup>1</sup>Throughout this article the distinction between positive and negative reinforcement (and, correspondingly, between safety signals and warning signals in avoidance) is employed in the conviction that it provides a useful descriptive convention despite the fact that all stimulus transitions might be described either in terms of the production or of the termination of events, the occurrence and effects of which are not separable (e.g., the production of a safety signal might be described as the termination of its absence, or the production of heat as the termination of cold). With regard to the prediction and control of problematic avoidance, this distinction serves to remind us of the utility of attending to the functional properties of stimulus events on *both* sides of response-produced transitions from threat to safety.

An example of an episode of such a process is diagrammed in Figure 1. A schoolboy rises from his classroom seat without permission and is promptly smacked on the back of his head by his teacher; this latter event unconditionally elicits intense sympathetic arousal and muscular tension, and would be expected to function as a negative reinforcer. Simultaneously with its occurrence, the act of rising from his seat ( $A_1$ ) produces proprioceptive and other response-produced stimuli that, while initially neutral for sympathetic arousal and general muscular tension, acquire conditional eliciting and negative reinforcing potency by virtue of their temporal association with the smack (provided, of course, that the smacks do not often occur, in this context, in the absence of such response-produced stimuli, and vice versa). Therefore, when the boy once again starts to rise from his classroom seat, the proprioceptive and other stimuli mechanically produced by this act will elicit arousal and tension; even if the teacher does not smack him on some occasions, standing up is now punished by its automatic, conditionally aversive consequences. What is the boy to do?

As *sitting down* has often terminated the stimuli produced by standing up, this act ( $A_2$ ) is likely to be evoked in the presence of the negative reinforcers now generated by standing up ( $A_2$  might also initially be evoked by a suggestion provided by the teacher). Sitting down ( $A_2$ ) terminates the conditional aversive stimuli (or *warning signals*) produced by standing up ( $A_1$ ), and this termination itself reduces arousal and tension. The act of sitting down *also* generates distinct proprioceptive and exteroceptive stimulus events which, while initially neutral, directly precede an event that reduces arousal and tension; in a threatening environment, these stimuli might therefore acquire the power to elicit some level of emotional relief or relaxation themselves, and to *positively* reinforce the acts that automatically produce them. When the positive reinforcing potency of events is established by their differential association

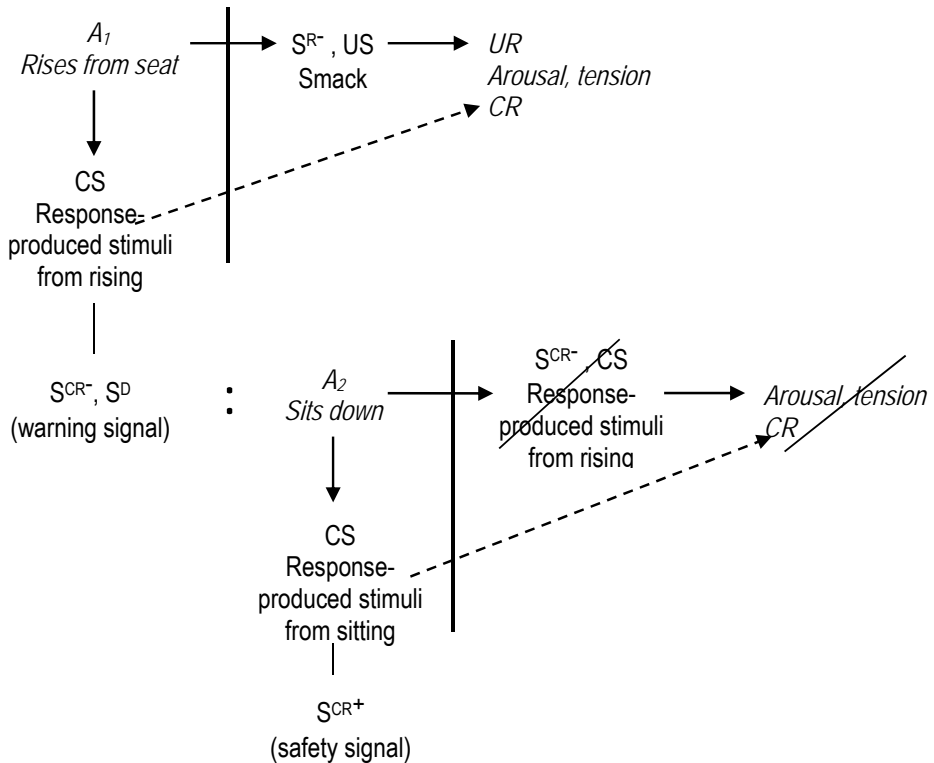


Figure 1. An instance of a typical operant chain establishing the conditional negative reinforcing potency of the stimuli (or warning signals) produced in the emission of an overt, punished response form, as well as the positive reinforcing potency of the stimuli (safety signals) produced by the responses that automatically terminate these warning signals.

with the termination of negative reinforcers, Dinsmoor and others have called them *safety signals*.

Thus, in this view, the warning signal, the avoidance response and the consequent termination of the warning signal (and production of the safety signal) comprise the terminal link of an operant chain initiated in the incipient or inchoate emission of a previously-punished response form. "Avoidance" is viewed as *escape* from *conditional* aversive stimuli, an act that inevitably produces conditional *positive* reinforcers as well. Once this arrangement of events establishes the positive reinforcing potency of the stimuli automatically produced in the emission of the "avoidance" response, they could presumably strengthen this behavior even the absence of a salient warning signal.

That is to say, once established, the positive reinforcing effectiveness of the safety signal could persist for some time in the apparent absence of the warning signal upon which its potency is based. Thus in the short term, neither the elicitation of fear nor the incipient emission of a punished response form would be required for the "avoidance" response (in this case, remaining seated) to be strengthened by the stimulus events that follow upon its emission.

To complete this picture, we must also ask why the boy stands up in the classroom in the first place. Presumably, this act is reinforced by such events as are produced when he stretches his muscles, and also by changes in his point of view and by access to his classmates and to out-of-reach objects. Thus, in this example, standing up is both



*reinforced* and *punished*, where the positively and negatively reinforcing events produced by this behavior act on receptors located on both sides of the individual's skin.

This commonplace example of the social punishment of automatically-reinforced operant behavior would appear to contain all of the physical features of the "conflicts" that psychoanalysts are fond of transferring to an intrapsychic realm; that is, two *incompatible* acts with *functionally opposing* consequences (see Skinner, 1974, pp. 148-166; Sidman, 2001, pp. 149-164). If the boy stands up, he enjoys a stretch and a look around, but also receives the response-produced portent of a smack on the head; sitting down eradicates both the positive and the negative reinforcers at one blow. The positively reinforcing safety signals produced by the non-punished act of sitting down (and remaining seated) are literally a consolation, but the emission of this act is nonetheless a concession to aversive events.

As the biobehavioral private events effected by both the punished and non-punished response forms begin to increase the emission frequency of the latter, this vignette would appear to contain the physical features of a Freudian "defense mechanism," that is, of "unconscious self-control." "self-control" in the sense that response-produced stimuli determine the emission probability of ensuing acts in accord with contingencies initially established by another controlling agent, and "unconscious" in that the boy is presumably unaware of the extent to which his own behavior enters into the determination of his subsequent behavior (cf. Hefferline & Perera, 1963). The transfer of control from events mediated by the teacher to events generated by the boy's own actions would appear to constitute the natural process described by mentalists of many stripes as the "internalization of social norms" (see Hefferline, 1962 for a related discussion).

Thus, common social interventions designed to offset the reinforcement inherent in such activities as stealing, cheating, lying, overeating, injuring, lovemaking, speeding,

neglecting responsibilities, ingesting psychoactive substances, offending or annoying others and being careless, among many others, might be viewed as establishing the negative reinforcing potency of stimuli automatically produced in the emission of the relevant response forms, as well as establishing the positive reinforcing potency of stimuli produced by acts that are incompatible with the emission of censured behaviors. By altering the functions of response-produced events, the changes in behavior produced by these interventions might generalize to situations in which the initial controlling agent is absent or in which the undesired behavior cannot be publically observed.

The response-generated stimulus account of avoidance differs from traditional versions of two-factor theory, in which the reinforcing event for the avoidant response was assumed to be fear reduction itself (e.g., Mowrer, 1960). In the traditional account, in the absence of fear there is no basis for reinforcement. In contrast, Dinsmoor proposed that avoidance is reinforced by the termination of response-generated events that *acquired* their negative reinforcing potency by temporal association with events that elicit fear, and by the production of response-generated events that *acquired* their positive reinforcing potency by temporal association with the termination or absence of these same events. Again, these events might reinforce or evoke operant behavior even on occasions when their emotion-eliciting power is negligible.

Before concluding this description of a contemporary response-generated stimulus account of overt avoidance, it should be noted that the process as diagrammed in Figure 1 might be closer to Schoenfeld's (1950) original conception of the safety-signal hypothesis than it is to Dinsmoor's (1977; 2001a, b) formulation. Schoenfeld specifically attributed the acquisition of the positive reinforcing potency of the stimuli automatically generated by the avoidance response to their temporal pairing with

the termination of the *conditional* aversive stimuli (or warning signals) generated by incompatible, punished response forms. Dinsmoor distanced himself somewhat from this position, stating that the necessary condition for establishing the positive reinforcing potency of a response-produced stimulus in a threatening environment is that it is “paired with the absence of shock” (2001a, p. 311) or “inversely or negatively correlated with the receipt of shock” (2001a, p. 314). However, in his attempt to test the independent, positive reinforcing potency of stimuli produced by avoidance responses (Dinsmoor & Sears, 1973), Dinsmoor was careful to amplify (with supplemental exteroceptive stimuli) these response-generated events *immediately following the emission of the avoidance response, and therefore immediately following the termination of the conditional aversive stimuli generated by punishable response forms*. It seems likely that if the exteroceptive amplification of events produced by avoidance responses were delayed, its effect on the rate of emission of avoidance would be diminished, even if the negative correlation of the avoidance response with shock presentation itself were maintained. At any rate, in cases in which the emission of the avoidance response is incompatible with the emission of the punished response, the proprioceptive, tactile and exteroceptive stimuli generated by acts that automatically terminate response-produced warning signals would *immediately* precede such terminations in the manner diagrammed in Figure 1.

### **Objections to a Response-Generated Stimulus Account of Aversive Control**

More fundamentally, it must also be noted that it is possible to predict the frequency with which the boy of our example stands up in class simply by measuring the more readily observable relative frequency and intensity with which this and other overt response forms are punished by his teacher (e.g., Herrnstein, 1969; Herrnstein & Hineline, 1966; Rachlin, 2000).

These same data would provide useful guidance for interventions intended to increase or decrease the frequency of publically-observable avoidance. Why then would we concern ourselves with response-produced events inconveniently located beneath the individual’s skin?

I subscribe to Dinsmoor’s view that although “response-dependent stimuli are not under the direct control of the experimenter, they are material and observable, rather than hypothetical, and cannot legitimately be ignored in describing the animal’s interaction with its environment” (2001; p. 328). Note that in Dinsmoor’s use of the word, “hypothetical” means “immaterial and therefore unobservable under any circumstance,” as opposed to “material but, in particular cases, unobserved.” In contrast with hypothetical events, events of the *kind* that Dinsmoor’s analysis appeals to *have* been directly observed, and the functions that he attributes to covert stimuli automatically generated by the organism’s behavior do not appear to differ from those of analogous publically-observable stimuli that other agents might freely provide, withhold or terminate.

When covert stimulus and response events are assumed to function in the same manner that overt events of the same kind have been observed to function, then their use is “hypothetical” only in the sense that the physical functions assigned to the covert events are not directly tested; their use is merely *consistent* with experimentally-verified descriptions of causal relations among events with similar properties. But in this sense of the word, *all* interpretations of behavioral phenomena that occur in uncontrolled conditions are “hypothetical,” including interpretations that refer exclusively to publically observable events. Without the plausible yet untested attribution of eliciting, discriminative, reinforcing and motivational functions to classes of physical events, in close accord with an adequately comprehensive, experimentally-established account of the nature and provenance of such functions, clinical behavior analysis is *impossible*, as is any

use of behavioral principles for understanding and ordering our everyday affairs.

Moreover, there are forms of behavior, such as silent reasoning, deliberating, imagining or fantasizing, for which response-produced stimuli might be especially important determinants, as such behaviors are almost entirely covert in form and appear to have no other direct effects on the environment. Why would the inaccessibility of covert acts to other observers render them impervious to their settings and biobehavioral effects, or their causal relation to these events scientifically uninterpretable (see Palmer, 2003)?

Even if it proved to be possible to predict and control overt acts without describing the relation of associated covert acts to their antecedent and consequent events, it would seem that such an omission would be unacceptable to any person who is frequently troubled and perplexed by his or her own cognition, or to any clinician who is entrusted to effectively intervene in such processes, or indeed to anyone who is more than passing curious about the determination of acts that only he or she can observe. Biobehavioral events produced by covert acts are material and observable, rather than hypothetical, and cannot legitimately be ignored in describing a psychotherapy client's interaction with the environment. Let us see then how a response-generated stimulus account of aversive control might help us to interpret the origin and maintenance of supposedly dysfunctional beliefs regarding the "undue significance" of "obsessive" thought processes.

### **Aversive Private Events in "Dysfunctional" Thinking**

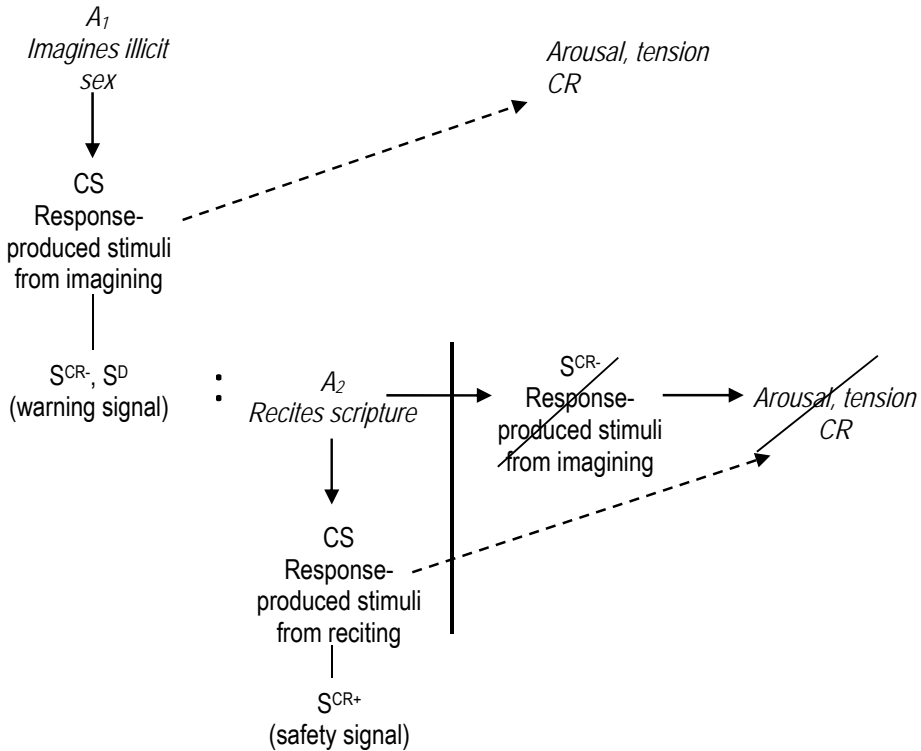
As in the case of higher-amplitude motor acts, lower-amplitude covert acts of corresponding form or content might produce warning signals that may be terminated in the subsequent emission of incompatible acts. That is, when the biobehavioral events automatically produced by a form of thinking or fantasizing elicit fear or shame and function

as negative reinforcers, they might be eliminated by doing something else. Moreover, by virtue of their temporal association with such terminations, the stimuli produced in the act of *escape* from private warning signals might acquire conditional *positive* reinforcing potency. The emission of an act that has terminated covert, response-produced warning signals in the past might therefore be automatically reinforced by the safety signals that it produces even on occasions on which the warning signals are not prominent. As will be detailed below, parties interested in controlling another individual's thought processes (as well as his corresponding overt acts) might arrange special circumstances such that the biobehavioral events that specific forms of covert behavior generate will function as positive or as negative reinforcers, in accord with the respective desirability or undesirability of these acts for the controlling agents.

Covert acts that produce warning signals for punishment might occur concomitantly with the emission of a variety of motor acts; therefore, such activities might not provide effective forms of escape from self-produced conditional aversive stimuli. For example, activities such as smoking or jogging are not necessarily incompatible with thinking about or imagining sexual or violent transgressions. Thus the response forms that might be most effective in terminating conditional aversive stimuli produced by covert acts might be *other covert acts*.

As the sole requirement for the effectiveness of such acts would be their incompatibility with behaviors that produce private warning signals, a wide range of covert response forms would be eligible for reinforcement. For example, while the warning signals mechanically produced when our schoolboy incipiently rises from his desk during a mathematics lesson might be eliminated only by sitting back down, those produced by *thinking* of standing up might be terminated by a much wider range of covert acts, from recollecting a play from yesterday's football match to envisaging the demise of his teacher.





*Figure 2.* An extension of the transition to safety analysis of avoidance to the stimuli automatically produced by covert acts. The stimuli produced in imagining illicit sexual acts are assumed to elicit tension and to function as negative reinforcers due to the punishment of corresponding overt acts and/or the evocation and punishment of similar covert acts (see text). Incompatible covert acts (here, silent recitation of a sacred text) terminate these events, eliciting relief and reinforcing their emission. As the stimuli automatically produced by these “safe” acts are differentially associated with the termination of negative reinforcers, they acquire positive reinforcing potency.

Similarly, the positively reinforcing but simultaneously threatening stimuli produced in the inchoate sexual approach of an attractive stranger might be effectively terminated only by some form of spatial withdrawal, whereas the biobehavioral warning signals produced in covertly *fantasizing* a corresponding act might be terminated by a much wider range of private activities, from performing mental arithmetic to realizing, in alarm, that one might have come in contact with harmful microbes.

Indeed, given that they are incompatible with the production of warning signals, the particular *form* of “safe” acts of cognition

would be irrelevant to the termination of warning signals; *all* covert acts that do not themselves produce warning signals should be equally effective in their termination. However, the covert response forms most *likely* to be emitted and to be, in this sense, adventitiously reinforced by the termination of warning signals that they bring about would be topographies that have been reinforced in other contexts (Mellon, 2009; see also Herrnstein, 1966).

An example is diagrammed in Figure 2, in which, due to a history of punishment of overt forms, the positively reinforcing stimuli produced in thinking about illicit

sexual relations also elicit tension and sympathetic arousal and function as negative reinforcers for the acts that terminate them. If reading or reciting a specific text has been reinforced by other agents in threatening contexts, the emission of a similar form of behavior may therefore be more probable under the threat of private, self-produced stimuli associated with punished acts. As many other forms of cognition would have been equally effective in terminating the warning signals produced by "forbidden" thoughts, in the present instance this particular *form* of cognition (reciting scripture) would be *adventitiously* reinforced. Should they be emitted, more idiosyncratic forms of incompatible ("safe") cognition, such as the counting and recounting of floor tiles, the unvoiced repetition of neologisms, or the act of imagining the aversive consequences of the budding transgression in an afterlife, would be strengthened in the same manner.

While the particular form of a "safe" covert act might be unrelated to its effectiveness in terminating the warning signals produced by "prohibited" covert acts, responses of the specific topography or content that *happened* to terminate the response-produced stimuli of punished forms will be more likely to be emitted when private warning signals are again produced, to be adventitiously reinforced on future occasions when they again terminate these conditional aversive stimuli. Thus while not required by the contingency, some level of stability in the form of responses that terminate private response-produced warning signals would be anticipated. Such relative stereotypy in the form of the covert acts that terminate warning signals would indeed be *required* for the establishment of *safety* signals, for if the form of such acts were sufficiently varied, specific response-produced stimuli would not be systematically associated with the termination of response-produced conditional aversive stimuli.

Thus covert behavior, like overt behavior, might be reinforced when it terminates threatening response-produced stimuli, and

the biobehavioral events thereby produced might acquire positive reinforcing potency as safety signals. Due to less specific response requirements, the form of covert behavior may, in general, be more susceptible to adventitious reinforcement than overt, but its form is not unconstrained; topographies of thinking that have been effective in other contexts would be favored for emission and therefore for adventitious reinforcement, leading to stability in the form of "safe" covert behavior.

### **A Second Source of Adventitious Reinforcement of Covert Acts**

A second source of adventitious reinforcement for the form or content of covert acts that terminate response-generated warning signals issues from the effects of the overt *motor* acts that often accompany them. Operant behavior can, by definition, be emitted in some form in the absence of specific antecedent stimuli, but its emission is of course more likely in contexts in which it has been differentially reinforced in the past. Thus, "forbidden" covert behavior would be more likely to be emitted in contexts in which corresponding *motor* acts have been differentially reinforced. If the emission of an incompatible ("safe") covert act is accompanied by a high-amplitude motor act that physically removes the person from the differential context of reinforcement for the "forbidden" act, the *covert* act would be adventitiously reinforced by the termination of exteroceptive warning signals brought about by the accompanying *motor* act.

For example, the emission probability of *thinking* of extramarital sexual relations (a covert act that might produce both positive and negative reinforcers for a person with the appropriate history of reinforcement and punishment) might well increase as the individual in question approaches a certain hotel at which she has an appointment with her lover; if the woman then begins to covertly recite scripture and subsequently turns away from the hotel, the emission probability of

the *covert* act might be increased both by the termination of the warning signals produced by a “forbidden” form of thinking *and* by the temporally contiguous termination of the warning signals issuing from the *exteroceptive* environment (e.g., the visual stimuli of the hotel), despite the fact these latter stimuli were mechanically terminated by the higher-amplitude *motor* act of turning away.

Similarly, the act of picturing oneself as extremely overweight might be adventitiously reinforced by the simultaneous termination of exteroceptive conditional aversive stimuli brought about when one turns away from the entrance to a pastry shop. Mellon (2009) has provided empirical evidence indicating that covert acts might be adventitiously strengthened when they repeatedly precede the termination of threats, despite the fact that threat termination was actually effected by concurrent by motor acts. Such co-occurrences of “safe” covert and motor acts would also (adventitiously) increase the positive reinforcing potency of the private stimuli automatically produced in reciting scripture or in “distorted body imaging,” respectively, because the comforting termination of exteroceptive threats would follow upon their production.

Having identified two potential sources of reinforcement for “safe” covert acts (the transition to safety from the threatening stimuli produced by “forbidden” covert acts and the termination of associated aversive stimuli in the external environment effected by concomitant motor acts), we are prepared to examine the relevance of these processes to the phenomenon of thought-action fusion in obsessive thinking.

### **An Operational Analysis of the “Moral Equivalence” of Thought and Action**

As noted above, thought-action fusion is, in essence, a generalized belief in the moral equivalence of the covert and overt forms of punishable acts. “Moral codes” might be viewed as descriptions of the consequences

of failures to conform to rules or standards of conduct established by a community to reduce the probability of emission of acts that, while reinforced by their immediate effects, are held to eventually cause harm to the actor or to others (Skinner, 1971; see also Day, 1977). Thus to say that, for example, fantasizing about illicit sexual relations is morally equivalent to performing the carnal act itself is to say that the covert and overt forms have the same consequences. But how can this be when the covert form of the act, by virtue of its inconspicuousness, is far more difficult for the community to punish?

Surely, part of what is meant by the “moral equivalence of covert and overt acts” is the belief that both response forms *should* or eventually *will* have the same aversive consequences, if only in some celestial court of justice which has direct access to our thoughts. For example, Christians are instructed to keep their thoughts pure, and that acts of imagining or desiring to perform sinful deeds are responsible choices to offend the creator deity; acts which have the same terrible implications for the sinner’s eternal future as those to be incurred for corresponding overt transgressions (e.g., Matthew 5:27-28). Indeed, a number of investigators have reported substantial positive correlations between levels of belief about the equivalence of covert and overt offences (i.e., of thought-action fusion) and levels of Christian religiosity (e.g., Berman, Abramowitz, Perdue & Wheaton, 2010; Rassin & Koster, 2003). However, for a person terrified by vivid descriptions of everlasting punishment, such training would be expected not only to affect what they *say* about the relation between thought and action, but also to change the effects of the stimuli *produced* by illicit covert acts on the behaviors that precede and follow them.

Thus, part of what is meant by the “moral equivalence of covert and overt acts” is that the stimuli produced in the covert emission of the act have effects on the actor’s behavior that are *similar to the effects of the consequences*

arranged by the community for corresponding overt emissions. Here again, the "internalization of social norms" might describe the transfer of control from socially-mediated to automatically response-generated stimuli.

Fundamentally, then, in the phenomenon called "thought-action fusion" thought might be neither confused with nor blended with action; rather, the *functions* of the biobehavioral events *produced* by thinking might be similar to the functions of the events produced by corresponding overt acts. Specifically, the same biobehavioral events that positively reinforce the covert emission of censured acts might acquire negative reinforcing potency and thereby the power to evoke the emission of acts that have terminated them in the past, acts such as incompatible forms of thought. Moreover, as we have seen, by virtue of their temporal association with the termination of negatively reinforcing and fear-eliciting events, the stimuli produced by thinking that is *incompatible* with the covert emission of socially-censored response forms might acquire positive reinforcing potency; that is, they might function as safety signals.

As noted above, in cognitive-behavioral theory, thought-action fusion is held to be a form of dysfunctional thinking in which repetitive thoughts are said to be vexing and problematic *because the individual interprets* his or her thoughts as having "undue significance." According to cognitive-behavioral theory, the problem, in the terms of the present analysis, is that the person *mistakenly* believes that covert and overt response forms have common effects (or "significance"). However, to the extent that the stimuli produced by covert and overt forms of punished acts indeed *have* similar eliciting, discriminative, motivating and reinforcing functions, the significance attributed to such thoughts, or rather, to the relationship between such thinking and overt behavior, is both justified and instructive.

It is not the individual's beliefs *per se* but the contingencies of punishment and reinforcement in the social and internal en-

vironments, as implied or explicitly described in the form of beliefs, that give the events produced by the individual's thoughts special significance. To the extent that the relative probabilities of subsequent actions in covert and overt response-produced circumstances are similar, declarations of thought-action fusion beliefs would constitute a form of autoclitic discriminative behavior, or self-awareness (see Thompson, 2008). Thus, belief in the moral equivalence of related response forms might be composed, in part, of the discrimination of the common effects of the biobehavioral events that they automatically produce. What then might be the functional significance of a relative equivalence in the negative reinforcing potency of the stimuli produced by thinking and by acting in corresponding ways?

### **Obsession as a Form of Unconscious Self Control**

When tempted to park illegally on New York City streets, one is likely to encounter a sign reading "Don't even THINK of parking here." These signs (and the catchphrase they inspired) capture an important sequential consistency in the form of human operant behavior: the emission of overt acts, particularly acts that are both reinforced and punished, is often preceded by the emission of corresponding *covert* acts. This common ordering of covert and overt acts of similar form or content is also manifest in the phenomenon termed "likelihood" thought-action fusion, that is, in generalized beliefs that imagining forbidden actions increase the probability of emission of corresponding overt behaviors. However, to the extent that the negative reinforcing potency of stimuli produced in fantasizing "forbidden" acts is similar to that of the stimuli produced by corresponding overt response forms, the response-produced private events would be expected to evoke the emission of *incompatible* response forms, thus derailing, at the covert level, a train of events that might otherwise have led to the emission of a punishable overt act.

Similarly, admonitions such as “*Think before you drink*” and “*Consider the consequences*” would appear to be designed to evoke the emission of covert acts that produce aversive events capable of evoking a change in a hazardous course of action, offsetting the positive reinforcement inherent in continuing that activity.

Indeed, the prohibition of wicked *thoughts*, rather than just the corresponding deeds, seems likely to be based in the recognition that our overt acts are often predicted and preceded by related covert acts. As responding in operant chains is characteristically weaker in its earlier links (e.g., Keller & Shoenfeld, 1950), *covert* acts might be more readily disrupted by aversive consequences than would the *overt* production of more powerful reinforcers later in the chain; the latter events cannot occur if the acts that produce them have been displaced by incompatible activities. Similarly, a person who is troubled when he harms others, and observes that his injurious actions often are preceded by thoughts of doing harm, might monitor his own thinking for indications of covert aggression, such that ill-destined operant chains might be “nipped in the bud,” that is, such that incompatible response forms might be evoked at a safe distance from the more powerful reinforcing and discriminative stimuli produced in subsequent links.

Note however that the events produced in the covert emission of punished acts function as *negative* reinforcers, especially for a person characterized by a high level of thought-action fusion. It can be painful, embarrassing or alarming, for example, to realize that (once again) one is imagining oneself engaged in an immoral or harmful and highly disapproved act. What then maintains the act of generating or observing such aversive stimuli at high frequency in persons characterized by obsessive-compulsive spectrum disorders?

Recall that punished activities such as prohibited erotic and aggressive acts also produce *positively* reinforcing events, which might maintain some probability of emission

despite the negative reinforcing potency of the events that they produce, as observed when a hungry rat continues to press a lever which gives it access to food despite the fact that the same act occasionally produces painful shock. Moreover, although they are aversive, the stimuli produced in the early, covert links of chains leading to punishable overt acts are also discriminative stimuli for the emission of acts *incompatible* with their continuation. If warning signals are terminated by the emission of incompatible acts *earlier* more readily than *later* in operant chains, the (simultaneously aversive) stimuli of low-amplitude, covert forms of punished acts will acquire additional *positive* reinforcing potency; the timely, if disturbing, discovery of incipient wayward behavior will reinforce observing one’s own thought-produced events.

Of course, the individual in question need not wait and watch for the emission of covert forms of punished acts; as they are automatically generated, private warning signals might be produced *whenever* the initiation of a punishable operant chain is probable. Due to their concurrent negative reinforcing potency, the response-generated stimuli that evoke the emission of incompatible response forms might elicit discomfort and might be described by the actor as “unwelcome” or “vexing” (and by clinical psychologists as “ego-dystonic”) but they are both self-produced and self-serving; a lesser “transgression” initiates a series of events incompatible with the eventual emission of a greater one.

When the stimuli produced by covert and overt transgressions have similar negative reinforcing potency, obsessions might, in the present sense, enter into self-control relations. For example, the stimuli produced when a mother fantasizes of harming her child might disturb and frighten her, but they might concomitantly evoke the comforting observation that she is not in fact injuring the child at the moment, as well as the emission of other acts that render such aggression



impossible or less probable, such as physical withdrawal from the child, self-reproach and repentance, support seeking, ritualized movement, or acts of kindness toward the child. Similarly, an anorexic person might imagine the consumption of highly caloric food, an act that, while distressing, might also be readily disrupted by its aversive consequences in a manner that precludes contact with the powerful gustatory consequences of corresponding overt acts. Thus while retaining their negative reinforcing potency, the stimuli produced in the incipient, covert emission of punished response forms might also acquire positive reinforcing potency, when they evoke the emission of acts that "lead us not into temptation, but deliver us from evil."

Note that there is nothing pathological about this process *per se*; the stimuli produced, for example, when one recalls or reminds oneself of an upcoming dental appointment might also be unpleasant, but they are presumably produced because such stimuli tend to initiate series of actions incompatible with those that would result in missed appointments and, ultimately, in much greater misery. Similarly, covert deliberation of whether or not to engage in any overt activity that has been both reinforced and punished might produce events that, while disagreeable, initiate an incompatible chain of action at a safe distance from the more powerful reinforcers of the overt acts. The process becomes problematic when the generation and subsequent termination of private warning signals is not only incompatible with the emission of harmful or censured response forms, but also with the emission of acts that would enrich the life of the actor.

### **The Social Establishment of the Negative Reinforcing Potency of Private Events**

Given that chains of corresponding covert and overt operant behavior are more easily disrupted in their earlier links, in which the private discriminative, motivating, eliciting and reinforcing events tend to be less potent than the events produced in later, overt links,

controlling agents who wish to reduce the likelihood of omission of specific forms of *overt* behavior might wish to establish the negative reinforcing potency of the stimuli produced in their earlier-appearing *covert* emission. However, a controlling agent who wishes to arrange for the punishment of forms of thinking in another person faces a technical challenge: as the behavior in question is not publically observable, it is difficult to arrange a contingency between its emission and some aversive event (e.g., Skinner, 1953; 1974).

One solution to this problem would be to punish *overt* acts that share stimulus properties with the covert acts, as the punishment of motor acts of a specific topography might increase the negative reinforcing potency of the stimuli produced by corresponding covert acts. This solution, though frequently employed as a temporary expedient, would not be entirely satisfactory to controlling agents, as efforts to coerce thinking or imagining appear to be designed to forestall *entirely* (or to severely limit) the emission of corresponding (and automatically reinforced) overt acts. Once the full-blown motor act is emitted and produces its powerful automatic reinforcers (once "the genie is out of the bottle") its emission probability may be much more difficult to restrict with aversive measures.

An alternative solution would be for the controlling agent to first arrange circumstances such that the emission of "undesired" *covert* acts in the controlee is highly probable, and subsequently to introduce aversive events. This method would appear to be frequently used in moral and in safety training, in which even a single emission of a form of overt behavior is considered to be highly detrimental. Using acoustic and iconic discriminative and eliciting stimuli (e.g., narration, texts, illustrations, dramatic presentations, films, songs, etc.) the controlee is induced to behave in a manner corresponding to that of a tragic hero, a behavior which subsequently has dreadful consequences for

the protagonist (see Schlinger, 2008 for a related analysis of listening as socially-mediated behavior). To the extent that the rendition of the hero's fall is frightening or painful to the observer, the stimuli produced in observer's own subsequent covert or overt production of events corresponding with the errors depicted will then trouble him and give him pause. In an effort by the controlling agent to forestall the emission of a greater transgression, a lesser one is evoked and punished.

The observer might also be induced to participate in a tragic hero's subsequent struggles to redeem himself, and when specific forms of behavior bring his troubles to an end, the stimuli produced by corresponding acts of the observer are paired with the relief-eliciting termination of negative reinforcers. The subsequent covert or overt emission of similar safe, redeeming or penitent acts might then be strengthened by their automatic, positively reinforcing consequences.

In the event that such training fails to forestall the emission of a "prohibited" act, a public confession of guilt might take a similar form. The suppliant is first induced to describe (that is, to verbally repeat) his transgression, then receives aversive consequences. He may ultimately be granted forgiveness as a consequence of the emission of expressions of remorse and resolve to change (that is, as a consequence of indications that the eliciting and negative reinforcing potency of the events produced by such acts is sufficient to preclude their subsequent emission, rendering further punishment unnecessary) as well as for acts of reparation incompatible with the emission of the punished response form.

The contingencies of punishment and negative reinforcement that inhere in such training are often summarized in the form of morals or rules of conduct. Due to their conciseness, such rules might then be conveniently introduced to increase the negative reinforcing potency of the stimuli automatically produced by "forbidden" acts in contexts in which they are more likely to be emitted. Rules specifying contingencies of punish-

ment and negative reinforcement might be introduced on such occasions by another agent, but their production might also be evoked by response-produced warning signals of an incipient offense, supplementing a "self control" relation. The recitation of rules specifying the punishment of "forbidden" acts would presumably be incompatible with the covert emission of the acts themselves and would increase the aversiveness of the stimuli produced in the incipient emission of such acts. Among these rules might be the statements of relations between covert acts and their consequences (e.g., "*impure thoughts incur God's wrath*") that constitute defining features of the phenomenon called thought-action fusion.

### **Equivalence Relations as a Source of Functions of Covert Response-Produced Events**

Experiments involving interlocking conditional discriminations with human subjects have suggested that stimuli can acquire negative reinforcing potency without being directly paired with other negatively reinforcing stimuli. For example, in an arbitrary matching to sample preparation, if choosing comparison stimulus B is differentially reinforced in the presence of stimulus A, and choosing comparison stimulus C is differentially reinforced in the presence of stimulus B, then in the absence of B, subjects will tend to choose C in the presence of A as well as A in the presence of C, despite the fact that these conditional discriminations have never been explicitly trained (e.g., Sidman, 1994). If stimulus C is paired with an aversive stimulus D, stimulus A might also begin to elicit autonomic arousal and function as a negative reinforcer, despite the fact that it was never been directly paired (by the experimenter) with stimulus D (e.g., Dougher et al., 1994). Moreover, in other conditional discriminations that include stimulus A, still other stimuli might acquire some of the eliciting and negative reinforcing potency of D, despite the fact that they have

never been paired (by the experimenter) with D or even with stimulus B (see Dymond & Roche, 2009 for a review of such studies).

The establishment of multiple conditional discriminations with shared stimulus elements may be yet another source of the social establishment of the conditional negative reinforcing potency of the stimuli produced by certain forms of thought. For example, if a person is taught that sin (A) marks the soul (B) and that those who die with such marks go to hell (C), where "hell" has been paired with gruesome phrases such as "eternal agony in a lake of fire" (D), it seems likely that not only the word "sin" but also unknown words subsequently associated with it (e.g., "sodomy is a sin," "hip hop is sinful") might acquire something of the aversive properties of stimulus D even without being directly paired with it. The aversive properties of "sodomy" or "hip hop" thus acquired would presumably remain if the words were then privately self-produced; these stimuli might then evoke the emission of covert acts incompatible with their production, initiating a series of "permitted" or "safe" acts.

Much has been made recently of the potential clinical significance of such bidirectional relations and transferences of functions in the apparent absence of a direct learning history, a process which is said to result in patterns of transformation of "almost unimaginable complexity" (Dymond & Roche, 2009, p. 11) held to provide the basis for avoidant behavior said to be inexplicable by a "purely contingency-based analysis" (Barnes-Holmes, Barnes-Holmes, McHugh, & Hayes, 2004, p. 355). However, one's ability to imagine the role of emergent relations in human thinking and psychopathology processes might be restored by a consideration of the fact that the new stimulus functions thus induced are both generated and constrained by the prevailing contingencies of reinforcement and punishment.

For example, in accord with the principle of extinction, emergent A-C and C-A con-

ditional discriminations (as well as their derived eliciting and reinforcing functions) will be weakened if they are not subsequently differentially reinforced, even if the reinforcement of the "if A, then B" and "if B, then C" conditional discriminations that was responsible for their emergence is maintained. The fact that, by a process that is not yet well-specified (e.g., Spradlin, 2003; Tapaeru Minster, Elliffe, & Muthukumaraswamy, 2011), stimulus functions emerge from combinations of conditional discriminations with shared stimulus elements does not mean that such functions inevitably materialize and are then free to proliferate without regard to circumstances. Useful clinical case formulations based upon such "derived" stimulus relations must carefully specify, on an individual basis, both the arrangement of contingencies of reinforcement and punishment (and of the stimulus pairings inherent in these contingencies) that might cause these functions to emerge *and then to be maintained*.

The present analysis might be useful in interpreting such maintenance. When a stimulus acquires both negative and positive reinforcing potency due to the punishment of a "pleasurable" act, both functions should "emerge" in neutral stimuli interlocked with it in conditional discriminations. If the termination of such "derived" (thus weak) evocative stimuli for punished acts effectively forestalls wrongdoing, they would contribute to self control.

### **Therapeutic Significance of Understanding Transitions to Safety in Problematic Thinking**

The present description of conditions that might establish the negative reinforcing potency of private events automatically produced in the emission of early-appearing, covert behaviors of operant chains leading to forbidden reinforcers, such that acts incompatible with the continuation of ill-fated or otherwise undesired activities might be evoked and automatically reinforced at a safe distance from the more powerful reinforcers of later-appearing overt acts, is obviously metaphorical.

However, as B. F. Skinner once remarked of the interpretation of the determinants of problematic patterns of thinking, "If it is impossible to be wholly nonmetaphorical, at least we may improve upon our metaphors" (1954/1972, p. 247). While not directly testable, the present account consists exclusively of generalizations of familiar, experimentally-derived principles of behavior determination, principles that provide a conceptual framework for interpreting the provenance and maintenance of repetitive and vexing patterns of thinking in psychotherapy clients, a phenomenon that resists direct experimental analysis. As such it would appear to be of potential value to, and subject to improvement by, clinical behavior analysts.

A related advantage of the present interpretation is that the processes invoked are simple enough to be understood by laymen who complain of obsessive patterns of thinking, given appropriate instruction adapted to their personal histories. An interpretation of the origin and maintenance of thought-action fusion in obsessive thinking as a wholly natural and generally adaptive process might well be as valuable to the client as it is to the clinical behavior analyst.

This view is in accord with an analysis of the efficacy of a variety of forms of "talking cures," in which Charles Ferster (1972) noted that verbal psychotherapy might be usefully conceived as having two types of effects. First and most obviously, therapeutic interactions might change the eliciting, reinforcing, discriminative and motivational effects of classes of events on important features of the client's behavior, leading to new forms of extratherapeutic interactions that augment changes in stimulus functions initiated in therapy sessions. But successful psychotherapy also appears to increase the client's *observation* or *awareness* of systematic relationships between his or her own behavior and other events (see also Mellon, 1998).

The first class of effects consists of changes directly related to the client's "complaint" or "disorder;" it is these changes that are

typically assessed in outcome measures of treatment effectiveness. The second class of effects, which Ferster termed the development of "outsight," might be construed as a repertoire of application of general principles of determination of psychological phenomena in interpreting the provenance and maintenance of one's own tendencies to feel, think and take action in significant ways; a cohesive set of discriminations comprising the client's understanding of where his or her problematic behavior might have come from, why it has changed during the therapeutic period, and how he or she might now act to forestall its return in a similar or different form. The client's emerging propensity to observe his or her own overt and covert behavior as an orderly constituent of natural processes, rather than a mysterious caprice of mind, brain, character or soul, might initiate beneficial changes in perception, in emotion and in action tendencies, and reduce or eliminate the need for subsequent professional intervention when psychotherapy is concluded.

As a source of such understanding and in stark contrast with the dominant theoretical approaches in clinical psychology, the behavior-analytic perspective has the significant advantage of limiting its analysis of the provenance of psychological phenomena to the familiar and observable realm of physical events. With appropriate instruction adapted to their personal histories, persons undergoing psychotherapy could presumably discriminate the essential features of the processes of respondent and operant conditioning in the world around them, and apply these same principles in understanding important aspects of their own affairs, including their vexing thought processes, as readily as any undergraduate university student could.

Note, however, that not all clinical behavior analysts believe that the provision of such a general process account of problematic covert behavior would be helpful for verbally skilled persons whose behavior satisfies the

diagnostic criteria for psychological disorders. For example, in the highly popular therapeutic approach known as "Acceptance and Commitment Therapy" (ACT), persons undergoing psychotherapy are taught that, above all, they must *not* try to evaluate, understand or attempt to control their thoughts and feelings, as "verbal problem solving and reasoning is based on some of the same cognitive processes that can lead to psychopathology, and thus it is not practically viable to eliminate these processes" (Hayes et al., 2006, p. 5). Indeed, in ACT covert behavior processes are described to psychotherapy clients in explicitly mentalistic terms; one is instructed, for example, to abstain from the harmful (for psychotherapy clients) practice of evaluating or analyzing private events by turning off a "struggle switch in the back of one's mind."

The use of metaphor is a time-honored practice for generating discriminations of physical processes, but in ACT it is used to *suspend* the client's efforts to observe order between covert behavior and its conditions. While described as an empirically-grounded, principle-focused extension of a behavior analysis that failed to adequately account for cognition, ACT shares nothing of its complex, abstract, contextually-dependent, relational-frame interpretation of the determination of these processes with persons undergoing psychotherapy, as the establishment of the relevant discriminations would entail a good deal of the kind of verbal problem solving and reasoning that is held to be psychopathogenic. Exactly how psychopathology was avoided in the generation of ACT itself, which presumably also entailed extensive verbal problem solving and reasoning related to the determination of troublesome cognitive and emotional processes, is not clear; what *is* clear is that even in an approach to psychotherapy said to issue from and to extend behavior analysis, afflicted persons' understanding of the provenance of their psychological difficulties remains studiously uninformed by

naturalistic, general principles of behavioral determination.

In accord with the present account, it is not possible for psychotherapy clients (or for anybody else) to suspend control of their own thoughts, in the sense that thinking is inevitably evoked and inhibited by biobehavioral and other events produced by prior overt and covert behavior in the form of operant chains. It is only possible to change *how* thinking is controlled by the products of our own behavior, and the key to this transformation is a sufficiently detailed analysis of the functional significance of self-produced covert events. It would seem to be worth considering that the needs of people who are baffled and disturbed by their own repetitive patterns of cognition might be best served neither by violent struggles to suppress their own thoughts, nor by the passive observance and acceptance of covert behavioral processes as uncontrollable or inconsequential, but by a basic understanding of their physical determination and their relation to overt behavior.

People suffering with emotional disruption and debilitating action tendencies associated with their own thinking have been, and continue to be, deprived of the benefits of an adequately detailed, yet comprehensible, natural science interpretation of their troubles. While dismissed both by molar behaviorism and by psychological contextualism, a tentative, conceptually rigorous, point-by-point "molecular" interpretation the effects of biobehavioral events produced by covert activity, adapted to the client's current discriminative repertoire, might provide comfort and useful direction for individuals who regard their own thinking as threatening, alien and uncontrolled, or as issuing from illness or inherent flaws of character. As an alternative to efforts to reduce problematic cognition by encouraging additional self-produced aversive stimulation ("thought suppression") or a Zen-like resignation to its occurrence, the generation of a scientific understanding of the origin and maintenance of such behavior would provide the client



with a set of discriminations the acquisition of which seems likely to produce beneficial change in the functions of the events observed, and guidance in the production of circumstances likely to offset the pernicious side-effects of social punishment.

## References

- Barnes-Holmes, Y., Barnes-Holmes, D., McHugh, L., & Hayes, S. C. (2004). Relational frame theory: Some implications for understanding and treating human psychopathology. *International Journal of Psychology and Psychological Therapy*, *4*, 303-324. Retrieved from <http://eprints.nuim.ie/402/1/psyc21.pdf>
- Baum, W. M. (2011a). Behaviorism, private events, and the molar view of behavior. *The Behavior Analyst*, *34*, 185-200. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3211378/pdf/bhan-34-02-185.pdf>
- Baum, W. M. (2011b). No need for private events in a science of behavior: Response to commentaries. *The Behavior Analyst*, *34*, 237-244. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3211384/pdf/bhan-34-02-237.pdf>
- Berle, D., & Starcevic, V. (2005). Thought-action fusion: Review of the literature and future directions. *Clinical Psychology Review*, *25*, 263-284. doi: 10.1016/j.cpr.2004.12.001
- Berman, N. C., Abramowitz, J. S., Pardue, C., & Wheaton, M. G. (2010). The relationship between religion and thought-action fusion: Use of an in vivo paradigm. *Behaviour Research and Therapy*, *48*, 670-674. doi: 10.1016/j.cpr.2004.12.001
- Coelho, J. S., Carter, J. C., McFarlane, T., & Polivy, J. (2008). Just looking at food makes me gain weight: Experimental induction of thought-shape fusion in eating-disordered and non-eating-disordered women. *Behaviour Research and Therapy*, *46*, 219-228. doi: 10.1016/j.brat.2007.11.004
- Day, W. (1977). Ethical philosophy and the thought of B. F. Skinner. In Krapfl, J. E., & Vargas, E. A. (Eds.), *Behaviorism and ethics*. (pp. 7-28). Kalamazoo, MI: Behaviordelia.
- Dinsmoor, J. A. (1954). Punishment: I. The avoidance hypothesis. *Psychological Review*, *61*, 34-46. doi: 10.1037/h0062725
- Dinsmoor, J. A. (1977). Escape, avoidance and punishment: Where do we stand? *Journal of the Experimental Analysis of Behavior*, *28*, 83-96. doi: 10.1901/jeab.1977.28-83
- Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior*, *43*, 365-381. doi: 10.1901/jeab.1985.43-365
- Dinsmoor, J. A. (2001a). Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. *Journal of the Experimental Analysis of Behavior*, *75*, 311-333. doi: 10.1901/jeab.2001.75-311
- Dinsmoor, J. A. (2001b). Still no evidence for temporally-extended shock-frequency reduction as a reinforcer. *Journal of the Experimental Analysis of Behavior*, *75*, 367-378. doi: 10.1901/jeab.2001.75-367
- Dinsmoor, J. A., & Sears, G. W. (1973). Control of avoidance by a response-produced stimulus. *Learning and Motivation*, *4*, 284-293. doi: 10.1016/0023-9690(73)90018-0
- Dougher, M. J., Augustson, E. M., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior*, *62*, 331-351. doi: 10.1901/jeab.1994.62-331
- Dymond, S., & Roche, B. (2009). A contemporary behavior analysis of anxiety and avoidance. *The Behavior Analyst*, *32*, 7-28. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2686994/pdf/bhan-32-01-7.pdf>
- Ferster, C. B. (1972). An experimental

- analysis of clinical phenomena. *The Psychological Record*, 22, 1-16. Retrieved from <http://psycnet.apa.org/psycinfo/1972-29251-001>
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York, NY: Appleton-Century-Crofts.
- Goldiamond, I. (1962). Perception. In Bachrach, A. J. (Ed.), *Experimental foundations of clinical psychology* (pp. 280-340). New York, NY: Basic Books.
- Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and Commitment Therapy: Model, processes and outcomes. *Behaviour Research and Therapy*, 44, 1-25. doi: 10.1016/j.brat.2005.06.006
- Hefferline, R. F. (1958). The role of proprioception in the control of behavior. *Transactions of the New York Academy of Sciences, Series II*, 20, 739-764. doi: 10.1111/j.2164-0947.1958.tb00635.x
- Hefferline, R. F. (1962). Learning theory and clinical psychology: An eventual symbiosis? In Bachrach, A. J. (ed.), *Experimental foundations of clinical psychology* (pp. 97-138). New York, NY: Basic Books.
- Hefferline, R. F., & Perera, T. B. (1963). Proprioceptive discrimination of a covert operant without its observation by the subject. *Science*, 139, 834-835. doi: 10.1126/science.139.3557.834
- Herrnstein, R. J. (1966). Superstition: A corollary of the principles of operant conditioning. In W. K. Honig (ed.), *Operant behavior: Areas of research and application* (pp. 33-51). New York, NY: Appleton-Century-Crofts.
- Herrnstein, R. J. (1969). Method and theory in the study of avoidance. *Psychological Review*, 76, 49-69. doi: 10.1037/h0026786
- Herrnstein, R. J., & Hineline, P. N. (1966). Negative reinforcement as shock-frequency reduction. *Journal of the Experimental Analysis of Behavior*, 9, 421-430. doi: 10.1901/jeab.1966.9-421
- Hineline, P. N. (2001). Beyond the molar-molecular distinction: We need multi-scaled analyses. *Journal of the Experimental Analysis of Behavior*, 75, 342-347. doi: 10.1901/jeab.2001.75-342
- Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology: A systematic text in the science of behavior*. New York: Appleton-Century-Crofts.
- Lowenkron, B. (1998). Some logical functions of joint control. *Journal of the Experimental Analysis of Behavior*, 69, 327-354. doi: 10.1901/jeab.1998.69-327
- Lowenkron, B. (2004). Meaning: A verbal behavior account. *The Analysis of Verbal Behavior*, 20, 77-97. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2755434/pdf/anverb-behav00025-0079.pdf>
- Mellon, R. (1998). Oversight: Radical behaviorism and psychotherapy. *Journal of Psychotherapy Integration*, 8, 123-146. doi: 10.1023/A:1023284228809
- Mellon, R. C. (2009). Superstitious perception: Adventitious reinforcement and punishment as determinants of repetitive eccentric interpretations. *Behaviour Research and Therapy*, 47, 868-875. doi: 10.1016/j.brat.2009.06.016
- Mowrer, O. H. (1960). *Learning theory and behavior*. New York, NY: Wiley.
- Mowrer, O. H., & Lamoreaux, R. R. (1942). Avoidance conditioning and signal duration—a study of secondary motivation and reward. *Psychological Monographs*, 54, (5, Serial No. 247). doi: 10.1037/h0093499
- Nevin, J. A. (2008). Stimuli, reinforcers and private events. *The Behavior Analyst*, 31, 113-126. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2591752/pdf/bhan-31-02-113.pdf>
- Palmer, D. C. (2003). Cognition. In K. A. Lattal & P. N. Chase (eds.), *Behavior theory and philosophy* (pp. 167-185). New York, NY: Plenum Press.
- Palmer, D. C. (2004a). Data in search of a principle: A review of Relational frame

- theory: A post-Skinnerian account of human language and cognition. *Journal of the Experimental Analysis of Behavior*, *81*, 189-204. doi: 10.1901/jeab.2004.81-189
- Palmer, D. C. (2004b). Generic response classes and relational frame theory: Response to Hayes and Barnes-Holmes. *Journal of the Experimental Analysis of Behavior*, *82*, 225-234. doi: 10.1901/jeab.2004.82-225
- Purdon, C. (2004). Empirical investigations of thought suppression in OCD. *Journal of Behavior Therapy and Experimental Psychiatry*, *35*, 121-136. doi: <http://dx.doi.org/10.1016/j.jbtep.2004.04.004>
- Rachlin, H. (2000). *The science of self-control*. Cambridge, MA: Harvard University Press.
- Rassin, E., & Koster, E. (2003). The correlation between thought-action fusion and religiosity in a normal sample. *Behaviour Research and Therapy*, *41*, 361-368. doi: 10.1016/S0005-7967(02)00096-7
- Salkovskis, P. M. (1999). Understanding and treating obsessive-compulsive disorder. *Behaviour Research and Therapy*, *37*, 29-52. doi: 10.1016/S0005-7967(99)00049-2
- Schlinger, H. D. (2008). Listening is behaving verbally. *The Behavior Analyst*, *31*, 145-161. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2591755/pdf/bhan-31-02-145.pdf>
- Schoenfeld, W. N. (1950). An experimental approach to anxiety, escape, and avoidance behavior. In P. H. Hoch & J. Zubin (Eds.), *Anxiety* (pp. 70-99). New York, NY: Grune & Stratton.
- Schoenfeld, W. N., & Cumming, W. W. (1963). Behavior and perception. In S. Koch (Ed.), *Psychology: A study of a science*, Vol. 5 (pp. 213-253). New York, NY: McGraw Hill.
- Shafran, R., & Rachman, S. (2004). Thought-action fusion: A review. *Journal of Behavior Therapy and Experimental Psychiatry*, *35*, 87-107. doi: 10.1016/j.jbtep.2004.04.002
- Sidman, M. (1953). Two temporal parameters of the maintenance of avoidance behavior by the white rat. *Journal of Comparative and Physiological Psychology*, *46*, 253-261. doi: 10.1037/h0060730
- Sidman, M. (1960). Normal sources of pathological behavior. *Science*, *132*, 61-68. doi: 10.1126/science.132.3419.61
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston, MA: Authors Cooperative.
- Sidman, M. (2001) *Coercion and its fallout*. Boston, MA: Authors Cooperative.
- Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Macmillan.
- Skinner, B. F. (1954/1972). A critique of psychoanalytic concepts and theories. *Scientific Monthly*, *79*, 300-305. Reprinted in B. F. Skinner, *Cumulative Record: A Selection of Papers (3rd ed.)*, (pp. 239-438). New York, NY: Appleton-Century-Crofts.
- Skinner, B. F. (1966/1969). An operant analysis of problem solving. In B. Kleinmuntz (Ed.), *Problem solving: Research, method, and theory* (pp. 225-257). New York, NY: Wiley. Reprinted in B. F. Skinner, *Contingencies of reinforcement: A theoretical analysis* (pp. 133-157). New York: Appleton-Century-Crofts.
- Skinner, B. F. (1971). *Beyond freedom and dignity*. New York, NY: Alfred A. Knopf.
- Skinner, B. F. (1974). *About behaviorism*. New York, NY: Knopf.
- Tepaeru Minster, S., Elliffe, D., & Muthukumaraswamy, S. D. (2011). Emergent stimulus relations depend on stimulus correlation and not on reinforcement contingencies. *Journal of the Experimental Analysis of Behavior*, *95*, 327-342. doi: 10.1901/jeab.2011.95-327
- Thompson, T. (2008) Self-awareness: Behavior analysis and neuroscience. *The Behavior Analyst*, *31*, 137-144. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2591754/pdf/bhan-31-02-137.pdf>