CONCEPTUALIZING ADDICTION

Cognitive biases and addiction: an evolution in theory and method

CHRISTOPHER G. MCCUSKER

Department of Clinical Psychology, The Royal Hospitals Trust, Belfast, Northern Ireland, UK

Abstract
An evolution in theoretical models and methodological paradigms for investigating cognitive biases in the addictions is discussed. Anomalies in traditional cognitive perspectives, and problems with the self-report methods which underpin them, are highlighted. An emergent body of cognitive research, contextualized within the principles and paradigms of cognitive neuropsychology rather than social learning theory, is presented which, it is argued, addresses these anomalies and problems. Evidence is presented that biases in the processing of addiction-related stimuli, and in the network of propositions which motivate addictive behaviours, occur at automatic, implicit and pre-conscious levels of awareness. It is suggested that methods which assess such implicit cognitive biases (e.g. Stroop, memory, priming and reaction-time paradigms) yield findings which have better predictive utility for ongoing behaviour than those biases determined by self-report methods of introspection. The potential utility of these findings for understanding “loss of control” phenomena, and the desynchrony between reported beliefs and intentions and ongoing addictive behaviours, is discussed. Applications to the practice of cognitive therapy are considered.

Introduction
Current psychological interventions in the addictions, such as those pertaining to relapse prevention (Marlatt & Gordon, 1985) or motivational interviewing (Miller, 1985), have arguably had their roots in social learning theory. Such approaches are underpinned by theory which suggests that the addictive behaviour is maintained by a biased belief system, with interventions consequently aimed at cognitive restructuring. Although the nature of these biases may vary, both across addictions and in terms of how erroneous vs. merely “different” they are from controls, abstraction of principles suggests that inflated “positive outcome expectancies” for engaging in the behaviour, together with minimized negative expectancies and poor self-efficacy or beliefs about one’s ability to cope without the drug/activity, maintains the addictive behaviour and predicts relapse.

In support of this position, levels of alcohol consumption have been shown to vary as a function of positive belief biases (e.g. regarding enhancement of personal and social functioning), in both cross-sectional and prospective studies (Christiansen et al., 1989; Stacy, Newcomb & Bentler, 1991; Sher et al., 1996). Moreover, Connors, Tarbox & Faillace (1993) noted a relationship between the extent of such biases and post-treatment outcome in problem drinkers.
More recently, Jones & McMahon (1996) have highlighted a particularly important role for negative expectancies, and especially the “value” ascribed to those expectancies, in predicting abstinence survivorship following treatment in problem drinkers. Similar predictive utility for biased expectancies has been documented in smokers (Brandon & Baker, 1992), marijuana and cocaine users (Schafer & Brown, 1991) and gamblers (Griffiths, 1994). Discriminative differences in the balance of positive vs. negative outcome expectancies was demonstrated by McCusker et al. (1995) across four subgroups of UK teenagers who varied, not only in their actual drug-using behaviours, but also in their “vulnerability” to using illegal drugs.

Cognitive therapies have thus been aimed fundamentally at restructuring the belief systems assumed to be of motivational significance. Emphases and protocols have varied but inherent in most approaches is the promotion of cognitive dissonance by gradually highlighting and strengthening awareness of negative effects, challenging positive expectancies, and then facilitating more adaptive behavioural change as a resolution of this dissonance. The heuristic and clinical utility of these perspectives, together with promising outcomes in treatments (Moor, 1989) should not, however, obscure some very fundamental problems and anomalies in the theory.

Problems with traditional cognitive perspectives

Historically, much of the work on “outcome expectancies” appears to emphasize patterns of between group differences (e.g. positive expectancy differences between users vs. non-users of a drug). They have, however, paid insufficient attention to more anomalous within group patterns of expectancies which have been demonstrated for many addictive behaviours. Although smokers may endorse more positive outcome expectancies when compared to non-smokers, consistent with a positive bias, within group analyses suggest they endorse just as many, or indeed more, negative compared to positive expectancies (Litz, Payne & Colletti, 1987; Leung & McCusker, 1999). This is at best inconsistent with a positive bias hypothesis and may, indeed, suggest a negative bias. Similarly inconsistent with ongoing behaviour, Curran (1999) highlighted apparent negative biases in problem drinkers based on their endorsements of negative vs. positive alcohol-outcome word pairs. Moreover, while the balance or differences between negative vs. positive expectancies about illegal drug use were discriminating between groups in the study reported above (McCusker et al., 1995), within-group patterns none the less suggested that all groups, from “resistant” non-users to “repeated” users, reported a greater number of negative than positive outcome expectancies per se.

Such research may not be surprising to the clinician. Users of illegal drugs and problem drinkers may be more aware than most that the negative effects of their behaviour outweigh the positive (Plant & Plant, 1992; McCusker, Leung & Armstrong, 1999). Many young people start smoking while being cognizant of the dangers (Loken, 1982), and educational initiatives predicated on the assumption that greater awareness of negative outcomes would prevent illicit drug use have had at best equivocal success and at worst have been counterproductive (Plant & Plant, 1992).

A related problem for traditional cognitive approaches has been their equivocal ability to account for the desynchrony between cognitive intentions and ongoing behaviour, (i.e. the “loss of control” phenomenon). Current models of social cognition (e.g. Theory of Planned Behaviour, Ajzen, 1991) emphasize cognitive beliefs, behavioural intentions, subjectively held norms and self-efficacy as antecedents and mediators of subsequent behaviour. This model has been shown to have good predictive utility for many health-related behaviours (Zimmerman & Vernberg, 1994; Connor & Sparks, 1996), with cognitively espoused “behavioural intentions” having a robust and proximal influence on predicting subsequent behaviour. This relationship is compromised in the addictive behaviours. Indeed, a defining feature of the addictions is the continuation of the behaviour, and/or excessive engagement in it, despite consciously expressed intentions to abstain from, or moderate, the behaviour. This may suggest there are processes governing the behaviour which are outside conscious awareness and volitional control (McCusker & Gettings, 1997; McCusker et al., 1999), a proposition which has been levelled elsewhere as a limitation in applying social cogni-
Cognitive biases and addiction

49
tive models to other health-related behaviours (Joffe, 1996).

Such a proposition questions the utility and validity of methods of cognitive research which have relied on self-report strategies to ask people to retrospectively introspect about the cognitive features of their addictive behaviour. Typical methods have used questionnaires or checklists in which participants are asked to endorse or rate the extent to which certain propositions (e.g. “smoking is relaxing” or “smoking will give me cancer”) apply to themselves. However, it has long been questioned whether such methods can actually reflect the sorts of cognitive processes and propositions which actually motivate ongoing behaviour (Nisbett & Wilson, 1977; Feldman & Lynch, 1988).

In these procedures the individual is arguably engaged in a “cued-judgement” exercise (Leung & McCusker, 1999). Such judgements may be confounded by dissonance reactions—for example, the individual may be deciding that because they drink so much alcohol they must agree with a proposition that drinking is relaxing (Stacy, 1997). Judgements made may reflect socially held “facts” (e.g. that smoking causes cancer), rather than personally held and motivating propositions, and they may be mediated by demand characteristics of the test situation or by what the individual would “like” to think they thought. Moreover, responses made to earlier test items may prime, or artificially bias, responses made to later items (Bargh et al., 1986). Most fundamentally, even if such judgements do reflect personal and motivating propositions, they may say little about whether these are the sorts of propositions which are accessible and retrievable from memory in everyday situations pertaining to the addiction (Leung & McCusker, 1999). It has been a growing appreciation of this latter point—that investigation of cognitive biases in the addictions must be contextualized within what we know of memory processes and structures (Stacy, 1997; Goldman, 1999; Leung & McCusker, 1999)—that a second wave of cognitive perspectives is emerging.

**Alternative methods of enquiry**

A shift in the conceptual and methodological reference base, used to study cognitive biases in the addictions, has begun to occur during this last decade. It is known, for example, that “knowledge”, which may simply be a word which describes beliefs or expectancies assumed to be true, may mediate ongoing behaviour “automatically”—i.e. without need for conscious awareness or intentional control (Shiffrin & Schneider, 1977; Logan, 1988). Thus, self-reported introspection may be telling us only part of the story about the sort of cognitive processes and beliefs involved in an addictive behaviour. Tiffany (1990) has likened addictive behaviours to other highly skilled behaviours (such as playing the piano or driving a car). He draws persuasive parallels between the features of the addictive experience and automatized behaviours (“stereotyped … loss of control … stimulus bound … difficulty regulating by intentional processes”: Tiffany & Carter 1998). If such behaviours are optimally under the control of automatic cognitive processes and judgments, which are assumed to be outside of conscious awareness, they may not easily be open to introspective accounting. What may emerge will be as useful, but as hopelessly limited, as the professional golfer simply “telling” the club player how to execute the perfect drive.

It is also known that accessibility of information from memory is highly cue- and situation-dependent. Individuals may indeed be aware of automatically triggered propositions within the addiction-related context, but the ability to report these propositions at a later time may become obscured by the other dissonance and demand influences of the research or clinical situation as discussed above. Thus the sort of information accessible or reported when completing a questionnaire, involving cued-judgements in a research context, may bear little similarity to information most readily accessed in drug-using contexts. Cooney et al. (1987) have demonstrated that self-reported “expectancies” do change pre- and post-exposure to addiction-related cues.

In this context, a number of cognitive researchers have introduced methods for investigating cognitive biases in the addictions which are drawn from cognitive science and cognitive neuropsychology (Litz et al., 1987; Sayette & Hufford, 1994; Weingardt, Stacy & Leigh, 1996; Armstrong, 1997; McCusker & Gettings, 1997; Leigh & Stacy, 1998; Tiffany & Carter, 1998; Weinstein et al., 1998; Leung & McCusker, 1999). Such methods do not rely on what people “say” about what they think, but rather make
inferences about cognitive processes and structures based on behavioural responses (e.g. on memory, priming, reaction time or perceptual tasks). Drawing typically from concepts of “implicit cognition” (Schacter, 1992), or those aspects of recollective experience which are not in conscious awareness but which may nevertheless influence ongoing behaviour, such methods (a) directly assess the propositions and processes of memory which putatively motivate ongoing behaviour rather than the individual’s self-perception of behaviour, (b) are more demand-free since direct enquiry about reasons for behaviour are not made, and (c) access those aspects of cognition which are not open to introspective accounting (Stacy, 1997). What do such methods of cognitive enquiry tell us?

**Automaticity of cognitive biases**

A number of studies across different addictive behaviours have supported the proposition that cognitive biases pertaining to an addictive behaviour may operate at automatic levels. In one such group of studies modified Stroop tasks (Warren, 1972) have been employed. In these tasks participants are asked to indicate the colour of ink in which a series of words are printed as fast as they can. Successful performance requires attention to the perceptual characteristics of the word and suppression of semantic processing. Despite the task incentives, problem drinkers, smokers and compulsive gamblers have shown a selective interference effect (i.e. increased colour-naming times) for words whose semantic elements pertain to their addictive behaviour (Gross, Jarvik & Rosenblatt, 1993; Johnson et al., 1994; McCusker, McClements & McCartney, 1995; McCusker & Gettings, 1997). Figure 1 illustrates the effect in a study with compulsive gamblers. In this study (McCusker & Gettings, 1997) gamblers showed selectively elevated colour-naming times for gambling-related words, compared to both neutral words and words pertaining to other addictive behaviours. That this effect was maintained, not only in relation to a control group with little gambling experience, but also in comparison with their own spouses (who the authors argued would have had familiarity and emotional experiences with gambling-related constructs), suggests that the effect was not simply due to familiarity with, or a non-specific reaction to the emotional salience of, the stimuli.

Interference on the Stroop task is generally interpreted as due to automatic cognitive processes (Reingold & Toth, 1996). The findings reported above suggest, therefore, that some
form of automatic preoccupation with addiction-related information, which overrides or compromises the intentional behaviour of colour-naming, is occurring. In other words a pre-conscious and non-volitional bias for processing, or attending to, addiction-related information appears to be associated with the addictive behaviours.

Griffiths’ (1994) field studies, involving verbal protocol analyses of cognitive activity when gamblers were actually engaged in gambling behaviour, is consistent with these experimental studies. He found that the minds of heavier gamblers tended to go “blank” during significant periods of gambling behaviour. At such times they seemed unable to report on internal, gambling-related, cognitive activity. After considering alternative explanations for this phenomenon, the author invoked concepts of “automaticity” of gambling-related cognitive processes.

Other studies have arrived at similar conclusions, using different paradigms. Sayette & Huffman (1994) and Sayette et al. (1994) had smokers and problem drinkers engaged in a reaction time task, while exposed to task-irrelevant auditory stimuli which, for optimal task performance, should be intentionally ignored. However, they found that participant reaction times were selectively increased if the irrelevant stimuli were related to their addictive behaviour. The authors interpreted their findings as indicative of an automatic diversion of cognitive resources from the intentional cognitive activity and behaviour, towards stimuli and information related to their addictive behaviour.

Findings that cognitive processes related to an addictive behaviour can operate at automatic, pre-conscious, levels of awareness, and that such processes negatively affect intentional behaviour, may have implications for understanding the clinical phenomenon of loss of control. A related question is posed, however. That is whether, and what, automatic biases might exist in terms of the behaviourally motivating propositions, beliefs or “expectancies” associated with the addictive behaviour. The next group of studies reviewed address these questions.

**Biases in memory structures and processes**

It is assumed that all propositions or expectancies which motivate behaviour reside in neocognitive structures which may generally be referred to as “memory”. Drawing, therefore, from the paradigms of cognitive neuropsychology to investigate memory structures and processes pertaining to addictive behaviours, cognitive biases have been reported which could not be elucidated by traditional self-report methods. Litz et al. (1987) found that while smokers endorsed just as many negative as positive outcome associations, their incidental recall of the positive outcomes used in the task was better. This memory bias for positive information about smoking was more consistent with their actual smoking behaviour than what they said they believed. Armstrong (1997) assessed endorsements and endorsement reaction times for positive vs. negative alcohol associations in two groups of heavy and light drinkers. She argued that endorsements might reflect the “availability” of certain propositions in memory, but that the speed at which affirmative endorsements were made would indicate how “accessible” these constructs were from memory. Such an argument is consistent with long-standing methods of evaluating “distance” between concepts in semantic memory (Meyer & Schvaneveldt, 1971; Warren, 1977, Yaniv & Meyer, 1987). Typical between-group differences were found in the number of positive endorsements made. However, of greater interest was the finding that the light drinkers, despite endorsing more positive than negative alcohol-associations, were significantly faster in their endorsements of the negative outcomes. Heavy drinkers, however, endorsed more positive than negative propositions and were faster in the speed at which these positive endorsements were made (see Fig. 2a and b). The author argued that these findings imply that when stimuli pertaining to alcohol are presented, positive and behaviourally appetitive information is most accessible in the memories of heavy drinkers, while negative and behaviourally inhibiting information is more accessible in light drinkers. Such patterns of within-group differences in the accessibility of alcohol-related associations appear more consistent with behaviour than those elucidated by self-reported endorsements.

Studies utilizing other methods appear to confirm this accessibility bias for positive associations. Leung & McCusker (1999) used a free-association task with smokers and non-smokers. Both groups generated more negative than posi-
Figure 2. Mean number of positive vs negative endorsements and response times to make those endorsements in heavy and light drinkers. Both groups endorse more positive than negative outcomes (a). However, while light drinkers are faster in their endorsements of negative outcomes, the reverse is true for heavy drinkers (b) (Armstrong, 1997).

Positive associations to a smoking cue. However, while the ratio of positive/negative associations was constant across free-association time intervals in non-smokers, smokers generated proportionately more of their positive associations in the early time interval and proportionately more of their negative associations in the later time period. The authors suggested that associations generated in the early time period were of an “automatic” nature, whereas those generated later reflected more effortful and intentional search processes, and interpreted these findings as evidence for an “accessibility bias” for positive smoking associations in smokers.

Goldman and his colleagues (Rather et al., 1992; Goldman & Rather, 1993; Rather & Goldman, 1994) used a psychometric approach to arrive at similar conclusions. They used cluster analysis and multi-dimensional scaling techniques to compare the relative endorsement ratings given to positive/negative outcomes for drinking behaviour among heavy and light drinkers. Results suggested “tightly packed” clusters of positive associations in the heavy compared to the light drinkers. Moreover, while scaling techniques suggested that negative and positive propositions were closely related in “Euclidean distance” for the light drinkers, such
distance was significantly greater in the heavy drinkers. From this statistical model they suggested positive/appetitive alcohol-related information quickly triggered more negative and behaviourally inhibiting information in the memories of light drinkers. Such a moderating influence was, however, reduced in the “cognitive architecture” of heavy drinkers, who showed a bias in the accessibility of positive information “spreading” most immediately to further positive associations.

Related research has focused on “implicit” memory for addiction-related information. As noted above, implicit memory refers to aspects of recollective experience which may be “activated” and which may mediate ongoing behaviour, but in a way which is not consciously known to the individual (Schacter, 1992). One experimental way of measuring implicit memory has been to use “word-stem completion” tasks (Graf & Mandler, 1984). Thus words “primed”, but not consciously recalled from previous exposure to a word list (e.g. “Ban-”), may be offered as the first word triggered, rather than the many other possible alternatives, when the individual is asked to complete the word stem “Ban-” with the first word which comes to mind. This “priming” effect has been interpreted as indicative of “implicit” memory for the word (Graf & Mandler, 1984). Biases in implicit memory for addiction-related words, and for positive vs. negative outcomes related to the addiction, have been observed in gamblers, heavy drinkers and smokers (Armstrong, 1997; McCusker & Gettings, 1997; Leung, personal communication). Moreover, within-group differences for primed positive vs. negative words in these studies were more consistent with actual behaviour than within-group differences in endorsements of, or explicit (conscious) memory for, the words.

Finally, Stacy and his colleagues (Stacy et al., 1996; Weingardt et al., 1996; Stacy, 1997; Ames & Stacy, 1998) have used a “semantic priming” measure of implicit cognition in which participants were asked to generate activities associated with positive/negative “outcomes” or states of being (e.g. “relaxed” “happy”, etc.). Their research suggested that positive outcomes, not linked explicitly in the task to alcohol, nevertheless automatically primed representations of alcohol as a function of drinking history and behaviour. A similar priming effect was found in marijuana users (Stacy et al., 1996). These implicit memory biases not only varied as a function of current or previous alcohol/drug use but also predicted future usage in a prospective design (Stacy, 1997). Moreover, inspection of the statistical pathway analyses presented (Stacy, 1997) suggests that such measures of implicit cognition predicted a greater proportion of the variance in later behaviour than did “explicit” self-reported expectancies.

Discussion
In explaining the findings presented in this paper, authors have tended to draw from spreading activation models of memory (Collins & Loftus, 1975; Anderson, 1983), schema theory (Turk & Salovey, 1985), implicit cognition (Tulving & Schacter, 1990) and neural network theory (Grossberg, 1995; Masson, 1995). A detailed review of each of these theories is beyond the scope of this review. It is probably also fair to say that a specific model of cognitive biases in the addictive behaviours, derived from this class of theories, is still at an emergent stage. In terms of a working model of aetiology, however, representations of the behaviour (e.g. drinking, gambling, smoking, etc.) are assumed to be “linked” in long-term or semantic memory with propositions about outcome (e.g. relaxing, risky, etc.). Such links may be created by direct experience but are not likely to be solely determined by this, and may be formed by abstraction of information from the environment. The motivational significance of these associations is likely to be positive and appetitive, consistent with experience in the early stages of an addiction career and the initial effects of the substance/behaviour. These “semantic” links between representations of the behaviour and outcome become strengthened and more tightly connected with repeated behavioural “practice”. Over time, activation of one part of the “network” (e.g. alcohol-representations) comes to automatically trigger propositional links in other parts (e.g. relaxation concepts) and, importantly, vice versa. Thus an accessibility bias for positive information about the behaviour develops. Negative and behaviourally inhibiting information may be available, and perhaps even comes to be more available than positive information, to the individual. However, since this information is less accessible and relies more on effortful and non-automatic cognitive processes, a moderating im-
pact on behaviour is compromised. Moreover, if representations of the addictive behaviour are neurally connected to positive motivational brain systems (Grossberg, 1995), the automatic diversion of cognitive resources towards processing addiction-related stimuli to the detriment of intentional cognitive activity (cf. the Stroop studies), may have adaptive evolutionary significance.

Automatic and implicit cognition is known to operate at pre-conscious levels of awareness, and indeed appears to be subserved by neural pathways which are at least partially independent of those related to explicit cognition (Schacter, 1992). The anomalous desynchrony between self-reported beliefs and intentions and ongoing addictive behaviour, discussed earlier in this paper, becomes more understandable if many of the cognitive biases mediating the behaviour are operating at automatic and pre-conscious levels of awareness. Clinical and research assessments, which rely only on self-report methods of introspection, conducted out of the normal social context for the addictive behaviour, will not easily access these. Moreover, counteracting the effects of such automatic biases may not be easy. Neuromotivational pathways, once triggered, may be difficult to “turn off”, consistent with Gray’s (1990) theory of mutually inhibitory motivational systems in the brain. Implications for understanding “loss of control” phenomenon begin to occur. The Stroop studies certainly show that it is difficult to inhibit such automatic processing of addiction-related stimuli and that this detracts from intentional, but non-automatic, cognitive activity. Sayette & Hufford (1994) suggest that the very fact that automatic processes override and deplete purposeful cognitive resources may detract from the mental effort required to cope with “high risk” situations in individuals attempting abstinence. Tiffany & Carter (1998) have suggested that it is primarily when attempts are made to stop the automatized “drug use action schema”, or when it is blocked, that the discomfiting experience of “craving” emerges. Such a state may be most immediately relieved by lapsing back in the addictive behaviour, which reinforces the whole vicious cycle further.

The new programme of research on cognitive biases in the addictions discussed in this paper is in its infancy. Work to date has been largely of a cross-sectional nature. An important direction for future research will be to clarify the mechanisms of action by which automaticity of cognitive processes and judgments directly mediate an “automaticity” of behaviour (e.g. the “loss of control”, compulsive, element of addictive behaviour). Stacy’s (1997) longitudinal study, which appears to suggest that automatic cognitive biases have statistically better predictive power for drug and alcohol behaviour across time than self-reported expectancies, confirms the importance of this endeavour. Roehrich & Rather (1995) have also shown in an experimental study that the priming of positive alcohol expectancies increases drinking behaviour in the absence of conscious recollection of such exposure. Such studies strengthen the case for causal associations between implicit drink-related cognitions and subsequent behaviour. Secondly, it may be important to clarify whether available negative information might actually be inhibited or “avoided”, rather than simply not be accessible, if for example the individual is in an appetitive motivational state. The greater “accessibility” for positive vs. negative alcohol-associations in heavy vs. light drinkers reported above (Armstrong, 1997) was not, for example, found to be generalized to problem drinkers in treatment vs. social drinkers (Armstrong, personal communication). Rather, Armstrong found a trend to be evident for problem drinkers in an inpatient setting, who were currently motivated to attain abstinence, to show greater availability and accessibility for negative alcohol-related information. “State” vs. “trait” aspects of cognitive biases warrant further clarification.

Finally, clinical implications of this emergent body of work should be considered. As well as having demonstrated diagnostic significance, there are implications for the protocols of cognitive therapy in addiction. Most ostensibly, “activating” implicit positive propositions which motivate the addictive behaviour, rather than simply rehearsing the negative associations which are already likely to be in conscious awareness, would seem an important first step to modifying them. Incorporating the methods of imaginal or in vivo cue-exposure, while attending to and exploring the cognitive elements of the experience, may facilitate this, given the situational and cue-dependent nature of memory processes. Clinical studies of this nature, as well as further basic experimental studies, will undoubtedly play a role in advancing this promising avenue of new research.
words more accessible, but not necessarily more retrievable, Journal of Verbal Learning and Verbal Behavior, 23, 553–568.


