

**A science of behaviour still in progress: The implications of rule-governed behaviour
and derived stimulus relations for ‘Beyond Freedom and Dignity’**

Colin Harte^{1,2} & Dermot Barnes-Holmes³

¹Department of Psychology, Federal University of São Carlos, Brazil

²Paradigma – Centre for Behaviour Science and Technology, Brazil

³School of Psychology, Ulster University, Northern Ireland

Author’s note: Preparation of this manuscript was supported by a research fellowship awarded to the first author by the São Paulo Research Foundation (FAPESP, Grant #2019/24210-0).

In 'Beyond Freedom and Dignity', Skinner argued for the use of a technology of behaviour to make life less punishing for the human species. That is, to arrange contingencies of reinforcement in such a way as to remove or limit putative and coercive control, and in doing so enhance the opportunities for reinforcing human activities. The potential benefits of this powerful and revolutionary idea are elegantly outlined in the book as well as compelling arguments for the dangers of classic notions of freedom and dignity, and indeed autonomous man. But as Skinner pointed out, "human behaviour is a particularly difficult field" (1971, p.12). Success in arranging contingencies to limit punishment and maximize opportunities for reinforcement has been observed in domains where behaviour can be specified relatively easily, such as with interventions for individuals with learning disabilities. However, this seems to be a far more difficult task when dealing with the often highly abstract and increasingly complex nature of the behaviour of verbally sophisticated humans. We certainly agree that there would be undoubted benefits to arranging contingencies of reinforcement in such a way as to deliver the reinforcers that should provide the basis for a 'happy' life, such as shelter, food, warmth, stimulation, play, and physical health. Yet the experience of verbally sophisticated humans tells us that even those who have access to these reinforcers may still live in psychological worlds filled with despair (see Hayes et al., 1999).

One domain that has been argued to be potentially crucial here, and that seems to be lacking from Skinner's treatment, is the impact of human language on designing and creating a culture based primarily on reinforcement rather than punishment and coercion. The problem is that through language, humans seem to have the unique ability to punish themselves in complex and abstract ways that are increasingly removed from *direct-acting* reinforcement contingencies. Informally, we can create abstract realities of which we have no direct experience -- literally, heavens and hells here on earth. Thus, although a community built on providing increased access to reinforcers, such as those listed above, could surely create greater feelings of freedom and dignity and potentially more productive and 'happy' societies, it is nonetheless possible that the insidious side of human language could still lead to myriad unperceived problems that were not reflected upon in Skinner's book.

It is important to stress that this should not be seen as a criticism of the text when considered in its time and place in history. Skinner was writing largely from a pre-appreciation of the impact of

human language on behaviour and the capacity it provides for deriving increasingly complex relations between arbitrary stimuli. Indeed, Murray Sidman's seminal study on stimulus equivalence was only published in the same year as 'Beyond Freedom and Dignity', and the implications for its elaboration into the generic account of language and cognition that would become known as relational frame theory (RFT; Hayes et al., 2001), and new forms of therapy, such as acceptance and commitment therapy (ACT book 1999), had not yet been conceived. Now equipped with what has been learned about human behaviour from this type of work, which has emerged in the 50 years since the publication of 'Beyond Freedom and Dignity', we know that the human ability to engage in highly abstract and complex derived relating is an issue of central importance when it comes to predicting and influencing human behaviour.

To Skinner's credit, one of the primary areas that seems to be of particular relevance here is a phenomenon Skinner himself introduced only a few years earlier: instructional control or rule-governed behaviour (1966). Some 10 years after the publication of 'Beyond Freedom and Dignity', an explosion of research would emerge exploring the conditions under which behaviour under the control of rules or instructions seemed to persist in the face of changing reinforcement contingencies. Indeed, another 40 years later, we are only now beginning to appreciate the sheer complexity involved in rule-governed behaviour and the seemingly crucial role played by derived relating in this behaviour. Furthermore, excessive, or persistent, rule-following seems to be intricately related to the ubiquitous nature of human misery but exactly how still remains unclear. Empirical work conducted by the current authors in recent years on persistent derived rule-following in the face of reversed reinforcement contingencies has only begun to unveil the complexity of this issue (see Harte et al., 2020, and Harte & Barnes-Holmes, 2021, for recent overviews of this work). For example, a derived rule or network of arbitrarily related stimuli may undermine the impact of changing reinforcement contingencies when this network, or part of it, has been derived many versus a few times in the past (low in derivation), and when it is highly consistent with the learning history that has preceded it (high in coherence). Furthermore, these effects themselves interact in subtle and sophisticated ways with each other to co-determine the relative flexibility of rule following or sensitivity to changing contingencies (e.g., Harte et al., 2021a, 2021b). Current evidence, albeit limited, suggests that

excessive rule following is implicated in psychological distress, or human misery, but in a much more complicated way than originally thought. For example, some empirical evidence shows that participants formerly presenting with persecutory delusions (Monestes et al., 2014) and self-reported depression (McAuliffe et al., 2014) persist with rule-following in the face of reversed reinforcement contingencies far more readily than ‘healthy’ control participants. In contrast, other studies have found that participants with self-reported depression are *more* sensitive to changes in task contingencies than their nondepressed counterparts (e.g., Rosenfarb et al., 1993).

All of this work on rule-governed behaviour, and how it impacts upon direct contingencies of reinforcement, clearly has important implications for our understanding of how operant contingencies shape and control human behaviour. That is, rules, and human language generally, when viewed as arbitrarily applicable derived relational responding, often appear to fundamentally change how direct reinforcement contingencies impact upon human behaviour. As a simple example, imagine you won \$1000 in the lottery. In one moment, this highly reinforcing event could be transformed into an aversive one if you were then told that someone you dislike intensely had just won \$10,000 in the lottery. You would not even need to see that person receive the winnings -- just the arbitrary sounds (words) informing you that your winnings compared unfavourably with those of another individual whom you disliked could greatly diminish the reinforcing value of your \$1000. Such is the power of human language (relating in arbitrary and abstract ways) to undermine the power of direct-acting reinforcement contingencies.

Thus, it could be argued that our understanding of contingencies of reinforcement when it comes to complex human behaviour was, in hindsight, limited and somewhat simplistic. Indeed, empirical and conceptual analyses we have conducted in recent years seem to suggest that persistent rule-following in the face of competing reinforcement contingencies may be better considered as evidence of competition between two separate classes of generalized operant behaviour (i.e., a relational network that involves a derived rule versus a relational network generated from direct interaction with direct reinforcement contingencies; Harte et al., 2021a, Zapparoli et al., 2021). Considering behaviour and rules as instances of complex relational networking that interact with contingencies of reinforcement and differentially control behaviour may help unpack this complex

area of study. That is, under what conditions do one set of competing operant contingencies control behaviour over another. With the benefit of this emerging scientific knowledge, therefore, it could be argued that while the impressive treatment provided by Skinner in 1971 is of course to be applauded, it may be seen as somewhat simplistic and limited in 2021. But this is a good sign – it shows that the science of behaviour analysis has not stagnated in that time.

Fifty years later, it remains the case that “we need to make vast changes in human behaviour” (p.10) to create a world that truly harnesses the benefits of a technology of behaviour for the good of the species. This seems just as salient in the present day as it did in 1971, given the continued ubiquity of human suffering and distress, and indeed other vital issues such as the current climate crisis and the rate at which we are destroying the planet and making parts of it inhospitable. Thus, although we undoubtedly applaud Skinner’s book, its crucial message and all that it stands for, 50 years later it seems unavoidably limited in terms of the sophisticated understanding we now have of human psychology and the impact that excessive control by verbal rules can have on behaviour. The complete implications of that more sophisticated understanding for ‘Beyond Freedom and Dignity’ and the design of a culture are not immediately evident to us, and indeed considering them in any great detail would be beyond the scope of the current article. But one thing is clear: the capacity to create misery even when there is no apparent reason to be miserable is something that would have to be addressed in any modern version of Skinner’s book. That is, in conjunction with its obvious phenomenal advantages, how do we deal with the insidious problems that language also creates for verbally able humans? While we are certainly in a better position now than in 1971 to begin to grapple with this issue, it still seems to be the case that “a science of behaviour is not yet ready to solve all our problems, but it is a science in progress” (Skinner, 1971, p. 158) and even more so that “the real mistake [would be] to stop trying” (Skinner, 1971, p. 153).

To close, although the impact and contribution of ‘Beyond Freedom and Dignity’ is still very much felt today (evidenced by this special issue), we unfortunately still seem some distance away from the world that Skinner envisaged, in which punishable behaviour is a rare occurrence and in which blame and censure are not employed as forms of control. Our subject matter is still highly complex. However, it seems wise to keep Skinner’s message to us in mind: “The important thing is

not so much to know how to solve a problem as it is to know how to look for a solution” (1971, p. 158). While we still may not know exactly how to solve the problem, perhaps approaching “the problem” of human language and cognition as involving derived relations and rule-governed behaviour could help us look for the solution.

References

- Harte, C. & Barnes-Holmes, D. (2021). The status of rule-governed behaviour as pliance, tracking and augmenting within relational frame theory: Middle-level rather than technical terms. *The Psychological Record*. Advanced Online Publication.
<https://doi.org/10.1007/s40732-021-00458-x>
- Harte, C., Barnes-Holmes, D, Barnes-Holmes, Y., & Kissi, A. (2020). The study of rule-governed behavior and derived stimulus relations: Bridging the gap. *Perspectives on Behavior Science*, 43, 361-385. <https://doi.org/10.1007/s40614-020-00256-w>
- Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & McEntegart, C. (2021b). Exploring the impact of coherence (through the presence versus absence of feedback) and levels of derivation on persistent rule-following. *Learning and Behaviour*, 49, 222-239.
<https://doi.org/10.3758/s13420-020-00438-1>
- Harte, C., Barnes-Holmes, D., Moreira, M., de Almeida, J.H., Aparecida-Passarelli, D., & de Rose, J.C. (2021a). Exploring a Training IRAP as a single participant context for analyzing reversed derived relations and persistent rule-following. *Journal of the Experimental Analysis of Behaviour*, 115(2), 460-480.
<https://doi.org/10.1002/jeab.671>
- Hayes, S. C., Barnes-Holmes, D, & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. Plenum.
- Hayes, S. C., Strosahl, K., & Wilson, K.G. (1999). *Acceptance and Commitment Therapy: An experiential approach to behaviour change*. Guilford Press.
- McAuliffe, D., Hughes, S., & Barnes-Holmes, D. (2014). The dark-side of rule governed behaviour: An experimental analysis of problematic rule-following in an adolescent population with depressive symptomatology. *Behaviour Modification*, 38(4), 587–613. <https://doi.org/10.1177/0145445514521630>.

- Monestes, J. L., Villatte, M., Stewart, I., & Loas, G. (2014). Rule-based insensitivity and delusion maintenance in schizophrenia. *The Psychological Record, 64*(2), 329–338. <https://doi.org/10.1007/s40732-014-0029-8>.
- Rosenfarb, I. S., Burker, E. J., Morris, S. A., & Cush, D. T. (1993). Effects of changing contingencies on the behaviour of depressed and nondepressed individuals. *Journal of Abnormal Psychology, 102*(4), 642–646. <https://doi.org/10.1037/0021-843X.102.4.642>.
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech, Language, and Hearing Research, 14*, 5-13. <https://doi.org/10.1044/jshr.1401.05>
- Skinner, B.F. (1966). An operant analysis of problem solving. In B. Keinmuntz (Eds.), *Problem-solving: Research, method, and therapy* (pp. 225-257). Wiley.
- Skinner, B.F. (1971). *Beyond Freedom and Dignity*. Pelican Books.
- Zapparoli, H.B., Marin, R., & Harte, C. (2021). Rule-governed behavior: An ongoing RFT-based operant analysis. *Perspectivas em Análise do Comportamento, 12*(1). <https://doi.org/10.18761/PAC.2021.v12.RFT.09>