The Status of Rule-Governed Behavior as Pliance, Tracking and Augmenting within Relational Frame Theory: Middle-Level Rather than Technical Terms

Harte, C., & Barnes-holmes, D. (2021). The Status of Rule-Governed Behavior as Pliance, Tracking and Augmenting within Relational Frame Theory: Middle-Level Rather than Technical Terms. *The Psychological Record*, *72*(2), 145-158. https://doi.org/10.1007/s40732-021-00458-x

Link to publication record in Ulster University Research Portal

Published in:
The Psychological Record

Publication Status:
Published online: 11/03/2021

DOI:
10.1007/s40732-021-00458-x

Document Version
Author Accepted version

General rights
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.

Download date: 01/02/2023

The status of rule-governed behavior as pliance, tracking and augmenting within relational frame theory: Middle-level rather than technical terms

Colin Harte[1] and Dermot Barnes-Holmes[2]

[1]Department of Psychology, National College of Ireland, Dublin, Ireland

[2]School of Psychology, Ulster University, Coleraine, Northern Ireland, UK

Correspondence concerning this article should be sent to Colin.Harte@ncirl.ie

Abstract

A recent systematic review has highlighted that the terms pliance, tracking, and augmenting have rarely been used as the basis for conducting systematic experimental-analytic research since their conception in 1982, despite their theoretical centrality to the study of rule-governed behavior and their presumed impact on psychological suffering. Given that some time has passed since the review article, it may be useful to reflect again upon their place within the literature on the experimental analysis of human behavior, and relational frame theory in particular. As such, the current article constitutes a "position piece" rather than another formal systematic review. In reviewing (informally) the literature since the systematic review, the recent emergence of psychometric research involving these concepts could be seen as reinforcing the original conclusions, in that researchers are recognizing that pliance, tracking and augmenting may be of limited value in the experimental analysis of human behavior. Instead, the concept of rule-governed behavior itself, as well as the sub-categories of pliance, tracking and augmenting, should be considered middle-level terms, which lack the relative precision of more technical terms within the literature on relational frame theory.

KEYWORDS: PLIANCE; TRACKING; AUGMENTING; RULE-GOVERNED BEHAVIOR; RFT/ACT

Research into rule-governed behavior has long been a focal point of experimental inquiry within behavior-analysis, and has formed the conceptual basis of a prominent third-wave psychological therapy, Acceptance and Commitment Therapy (ACT; Hayes et al., 1999). Almost 40 years ago, three functional classes of rule-governed behavior were proposed: pliance, tracking, and augmenting (Zettle & Hayes, 1982), which were used in producing a potential theoretical framework for cognitive-behavior therapy.

In brief, *pliance* refers to rule-governed behavior that is controlled predominantly by speaker mediated consequences for a correspondence between behavior and the rule (e.g., "you can only watch television after you finish your homework" where doing one's homework is under the control of the speaker mediated consequence of being able to watch television). The consequence (watching television) does not need to be explicitly stated in the rule but may be implied by the speaker and inferred by the listener (in this case, the listener "understands" that the speaker will allow access to the television only when the homework is complete). *Tracking* refers to rule-governed behavior that is controlled predominantly by the correspondence between environmental contingencies and the rule. That is, behavior under the control of this type of instruction is mediated by (potentially) contacting the consequences specified (explicitly or implicitly) in the rule itself, for example "study hard to do well in your exams and you will feel great afterwards" (in this example, the reference to "feeling great" may be explicitly stated or implied by the speaker, but in both cases the consequence must be inferred by the listener).

Finally, *augmenting* refers to rule-governed behavior that can occur together with pliance or tracking to alter the extent to which rule-specified consequences have reinforcing or punishing properties. Augmentals were suggested to take two forms: *motivative and formative*. A motivative augmental was argued to momentarily increase or decrease the

extent to which a previously established consequence functions as a reinforcer or punisher. For example, “it’s very cold outside today. You should wear a hat to keep you warm” could be considered a motivative augmental if it temporarily increased the reinforcing value of putting on a hat. A formative augmental, on the other hand, was argued to establish reinforcing or punishing functions for a previously neutral stimulus. For example, “this piece of paper is a voucher that will get you a free hat” could be considered a formative augmental if it established the piece of paper as a reinforcer. In general, when defining pliance, tracking and augmenting, it is important to emphasize that they are defined as antecedent verbal stimuli, and thus they influence the behavior of a listener because they refer to (either explicitly or implicitly) "apparent" consequences (i.e., they actualize specific functions in the stimuli for the listener).

While these functional classes of rule-governed behavior have sat within behavior analytic discourse since their inception almost 40 years ago, a recent systematic review highlighted that these terms have rarely been used as the basis for conducting systematic experimental psychological research in that time (Kissi, et al., 2017). Specifically, Kissi et al. first identified articles using eleven search terms (i.e., “rule governed behavior”, “rule-governed behavior”, “verbal regulation”, “instructional control”, “verbal rule”, “instructed behavior”, “instructed learning”, “instruction following”, “instruction-following”, “rule following”, “rule-following”) across five online databases (i.e., “Web of Science”, “PsychINFO”, “PsychArticles”, “PubMed (Medline)”, “Google Scholar”). Experts were also contacted for suggestions of other potentially relevant sources. From the articles identified, records were included if they: 1. used an experimental design, 2. focused on operant human learning, 3. stated clearly in the abstract or introduction that a primary aim was investigating pliance and/or tracking and/or augmenting, 4. were written in English, and 5. were published in a peer-reviewed journal. These criteria resulted in identifying only nine experimental

studies since 1982 (the year in which the terms were coined by Zettle and Hayes). Of these nine studies, two explored the differential impact of pliance versus tracking, two (comprising four experiments in total) examined pliance alone, one examined tracking alone, two explored formative augmenting, and none focused on motivative augmenting.

It is evident from the limited number of articles that met Kissi and colleague's (2017) criteria, that these concepts have rarely been used in experimental research since their conception. It was also noted in the review that when the terms have been the focus of experimental inquiry, they have sometimes given rise to contradictory findings. For example, a study by Baruch et al. (2007) found no differential impact of plys versus tracks on persistent rule-following in dysphoric participants, while a study by McAuliffe et al. (2014) did find an effect for pliance versus tracking, but only in participants who reported high dysphoria.

In their Discussion, Kissi et al. (2017) suggested that one of the main reasons for this tendency for inconsistent findings, and general lack of experimental use, may be that pliance, tracking and augmenting are not sufficiently well-defined, distinct functional-analytic concepts. For example, when a particular instance of rule-governed behavior is considered a track, it is often possible to argue that it also has some of the properties of a ply. Imagine, for example, someone is told that the best way to get from place X to destination Y is to follow route Z. The extent to which they follow that instruction may be considered a good example of tracking (assuming that the individual who provided the direction is not involved in reinforcing arrival at destination Y). On balance, imagine that the degree to which the rule is followed depends on the so-called credibility of the person who gave the directions. If the rule-giver was a police officer, for instance, rule-following may be more likely than if the rule-giver was perceived to be unreliable in some way (e.g., the rule-giver was perceived to

be slightly drunk). In this instance, although the rule-following could be defined as tracking, the tracking may not be entirely independent of the rule-giver (i.e., we are more likely to follow an instruction issued by an authority figure than a drunk). In this sense, it could be argued that there is some element of (implicit) pliance in this instance of tracking (i.e., in general listeners would be more likely to follow rules provided by police officers than random drunks). Of course, one could argue that other principles in behavior analysis are not necessarily binary. For example, the distinction between respondent and operant conditioning is not entirely dichotomous (see Rehfeldt & Hayes, 1998, for a review of these arguments). Nevertheless, the pragmatic use of such terms is immediately evident from the wealth of experimental research that they have given rise to over the past 70 or so years. Pliance, tracking, and augmenting, however, cannot claim this same heritage and utility within behavior-analysis, nor indeed, within psychology more generally.

One potential conclusion arising from the Kissi et al. review is that the concepts of pliance, tracking and augmenting should be completely abandoned. To reach such a conclusion would, in our view, involve failing to appreciate the value of its findings. One of the key aims of the current article, therefore, is to reflect on how the research community might best view the concepts of pliance, tracking, and augmenting (and the generic concept of rule-governed behavior itself) in light of the Kissi, et al, review. In engaging in this exercise it seems important to recognize that it is now five years since Kissi et al. conducted their final literature review (i.e., although the article was officially published in January 2017, the articles for the review were retrieved in July 2015). Given the time that has passed since then, it may be wise to ascertain if there has been any evidence of an increase in the number of articles published in that time that have been driven by these concepts. In doing so, we are not proposing to undertake another systematic review similar to that reported by Kissi, et al. because it could be seen as somewhat premature. Also, it could be argued that a second such

review would be of limited value because the original review was quite narrow, in the sense that no effort was made to connect to related literatures on compliance and obedience in the social psychology literature, for example. While recognizing this limitation, addressing it in a follow-up review, would miss the point of the purpose of the original review. Specifically, did these behavior-analytic concepts (pliance, tracking and augmenting) serve to generate a rich vein of experimental analyses of behavior? The fact that the terms are broadly or closely related to terms used in other areas of psychology is irrelevant to the question that the first review sought to address. The current article should thus be seen as a "position piece" rather than a review (and certainly not a systematic review). To reiterate, our aim here then, is to attempt to clarify how these terms would best be viewed within behavior analysis (or contextual behavioral science; CBS), focusing in particular on relational frame theory (RFT; Hayes et al., 2001).

The status of the concept of rule-governed behavior (and by implication, pliance, tracking and augmenting) in RFT

When discussing the concept of a rule within the seminal RFT volume (Hayes, et al., 2001) it was argued that, "it seems more useful to use terms like "rule" in a less technical way, as a method of ensuring contact with a domain of events. In unpacking this domain we need to note the functional processes involved and to collect data that might reflect on those processes. This will allow progress without creating blind spots." (Barnes-Holmes et al., 2001, pp. 108). Since this was written, the concept of the 'middle-level term' has emerged within CBS (Barnes-Holmes et al., 2016). Such terms are said to lack the precision of technical terms, but do serve to orient the researcher towards a domain. Indeed, many of the concepts associated with ACT have been deemed middle-level concepts, used for the purposes of communication among applied researchers, clinicians, and clients. Interestingly,

the argument that middle-level terms may serve as orienting stimuli for researchers toward a domain is almost exactly what was expressed in the original RFT volume without using the term 'middle-level'. As such, there should be no expectation that middle-level concepts will participate *directly* in the relatively precise experimental analyses typically associated with basic research in RFT studies. By extension, as subsets of the term rule, it follows that there should be no expectation that the terms pliance, tracking, and augmenting will participate in (basic lab-based) RFT analyses in any substantive or meaningful way. At this point, it also seems important to note that although the terms pliance, tracking, and augmenting appear in the seminal RFT volume (Hayes, et al., 2001), and could be argued to be reliable RFT concepts on that basis, they actually pre-date the emergence of the theory itself. Specifically, the terms first appeared in a paper published in 1982 by Zettle and Hayes, while the first public presentation of RFT is dated as 1986, four years later (Hayes, 2016). In this sense, it could be argued that these terms were backward-engineered into RFT in the absence of reasonably informative experimental analyses of derived relations and rule-governed behavior at that time (see Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020).

To be clear, we are not arguing that the concepts of pliance, tracking, and augmenting are without value *per se.* We are simply suggesting that these terms should be treated as middle-level concepts. As shall be seen below, a number of psychometric instruments have been developed in recent years, which aim to measure or quantify the behaviors associated with these terms (pliance and tracking). Interestingly, the recent focus on the development of such psychometric instruments serves to highlight, in our view, that these concepts do not share the same level of precision enjoyed by those RFT concepts that are regularly employed in (lab-based) experimental analyses, such as mutual entailment, combinatorial entailment, transformation of functions, relational networks, relating relations, and so forth (see O'Hora et al., 2004, 2014, for examples of highly technical experimental analyses of rules as

relational networks). In arguing that these latter concepts are relatively precise, we simply mean there appears to be broad agreement about how to employ these technical terms in the experimental analysis of behavior of individual participants. For example, the precise steps involved in producing a transformation of functions (in a lab-based study) in accordance with the relational frame of comparison would be relatively well defined.

Of course, we recognize that even these latter concepts are not *absolutely* precise (very few terms in behavior analysis, and even science generally, have absolute precision). Nevertheless, it seems important to strive for maximal clarity in determining which of the concepts that appeared in the original RFT volume (Hayes, et al., 2001) may be considered relatively precise lower level, bottom-up concepts or less precise middle-level terms. And we would also argue that this should be seen as a work-in-progress. For example, one might reason that the distinction between Crel and Cfunc concepts in RFT lacked clarity in the original volume but have since gradually acquired greater experimental precision in recent research (e.g., Finn et al., 2018, 2019). Finally, it should also be recognized that incorporating psychometric instruments, when they become available, into experimental analyses may be useful at times, but the mere development of those psychometric measures should not be seen as a *substitute* for the basic experimental analysis of behavior, which, in our view, provides the bedrock of RFT itself.

Insofar as RFT seeks to develop concepts that have clear utility in experimental analyses, it seems important to consider whether this criterion applies to the use of pliance, tracking, and augmenting in the literature within the last 5 years. Before proceeding, it is important to note that we are not suggesting that RFT concepts be restricted *solely* to basic experimental analyses, but rather that they need to be rooted there. Thus, it seems important to reflect upon the extent to which pliance, tracking, and augmenting have proven useful in the experimental analysis of human language and cognition as seen through the lens of RFT,

in an attempt to clarify how these terms would best be viewed within the field. In the next section, we will therefore first turn to research that has focused on pliance, tracking, and augmenting since Kissi et al.'s final literature search.

Before continuing to the next section, we should clarify that we are not entirely comfortable in designating terms as either strictly technical or middle-level. As noted by Barnes-Holmes, et al. (2016) "describing something as a middle-level term is a way of placing it *on a continuum* between the analytic units of the basic science. . . and folk psychological terms. . . within a given domain." (emphases added; pp. 367). In suggesting that the concepts of rule, pliance, tracking, and augmenting should be considered middle-level terms we are simply arguing that they apparently lack, at the current time, the experimental precision evident in the literature for other RFT-related concepts. Thus, we are simply arguing that in terms of a continuum, the concept of rule and any sub-categories (pliance, etc.) appear to be less precise than some other RFT concepts, but may well be more precise than folk-psychological terms.

To elaborate, as stated above, technical terms enjoy broad agreement about how to employ these terms in the experimental analysis of behavior of individual participants. To this we would add that relatively precise technical terms should guide researchers in exactly what to do in experimental research to produce specific effects without relying upon other ill-defined concepts. Thus, the original concept of rule-governed behavior lacked precision because it invoked the term 'specify' (as in specify a reinforcement contingency; Skinner, 1966) without clarifying, technically, the functional definition of 'specify' itself. Subsequently, the work of Sidman (1994) and RFT increased the precision of the term 'rule' by providing a technical definition of 'specify' (see Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020) and also defining and demonstrating experimentally that a rule could be

considered a type of derived relational network (e.g., O'Hora, et al., 2004). Critically, however, the concept of a rule within RFT still lacks precision because although all rules may be considered examples of relational networks, not all relational networks are necessarily rules. Thus, RFT can define what a rule is (a relational network) but cannot, at least currently, distinguish clearly (in an experimental analysis) between networks that are considered rules from those that are not. It is for this very reason that Barnes-Holmes, et al. (2001) argued that "it seems more useful to use terms like "rule" in a less technical way. . ." (pp. 108).

Having gone through the process of writing the current article, and addressing some of the reviewers' concerns, we are convinced even more so of the wisdom of this position. In this regard, the following points seem most salient to us. First, one reviewer asked that we modify the definitions of pliance and tracking contained in the first submission of the current article. Upon doing so, a second reviewer of the revised submission then objected to the modified definitions. Furthermore, and perhaps somewhat ironically, the second author of the current manuscript co-authored the chapter defining those very terms within RFT in the 2001 seminal volume (Barnes-Holmes et al., 2001). The apparent inability of so-called experts (current authors and reviewers alike) to agree readily on the definitions of pliance and tracking only serves to highlight their current lack of precision within the field.

Second, throughout the review process, conspicuous by its absence were concrete suggestions on how RFT researchers could, in fact, subject the concepts of pliance and tracking to rigorous experimental analyses. An entirely reasonable argument suggested that pliance and tracking should be distinguished ". . . according to the apparent consequences actualized in the listener for following the rule. . ." Interestingly, however, the same reviewer admits that "measuring the actualized functions of the participants or manipulating pliance-tracking functions seems quite tricky." And of course, that is the very point we are making --

the terms pliance and tracking do not readily yield to an experimental analysis because manipulating the variables that distinguish the two concepts is "quite tricky." Indeed, as mentioned previously, and discussed subsequently in more detail below, recent research on pliance and tracking has tended to adopt a psychometric rather than an experimental approach, which perhaps reflects the "tricky" nature of using these very concepts experimentally. And it seems worth noting that even the psychometric approach is limited in terms of clearly identifying these concepts (Stapleton & McHugh, 2020a). For example, these authors argued that although one of the psychometric measures that was designed to measure pliance would be

> undoubtedly useful (particularly in clinical contexts), it may not always effectively identify generalised pliance (see Waldeck, Pancani, & Tyndall, 2019 for further discussion). To illustrate this point, take item seven of the eight-item quantitative self-report measure; "It is very important for me that others have a good impression of me" (see Salazar et al. (2018) for the full questionnaire). If a child responds "frequently true" to this item, then they are not necessarily reporting generalised pliance. For example, the child could have a learning history where they previously experienced positive treatment when others had a good impression of them, leading the child to derive that "if others have a good impression of me, then I am treated kindly". Therefore, in this instance, the child is tracking the non-arbitrary consequence of positive treatment arising from good impressions (learned via direct experience), rather than adhering to a ply (p. 2).

The authors (i.e., Stapleton & McHugh) then go on to argue that one solution to this problem would be to adopt 'qualitative explorations' of childrens' rule-following. The success or

value of doing so remains to be seen, but once again this solution does not involve experimental analyses. At some point, the concepts of pliance and tracking will need to yield to an experimental analysis if they are to contribute directly to the scientific goals of prediction-and-influence, with precision, scope and depth. If meeting those goals with these concepts requires a mixed-methods approach, almost by definition an experimental analysis will also be required (i.e., psychometric, qualitative, *and* experimental).

In moving forward with an RFT-based analysis of rule-governed behavior one possible strategy could involve pursuing experimental analyses that do *not rely* on the concepts of pliance, tracking, and augmenting, and indeed we will outline a recent program of research, later in the current article, that has adopted this very strategy (see also Harte, Barnes-Holmes, Barnes-Holmes, & Kissi, 2020). As mentioned above, however, it seems important first to consider the extent to which recent research (i.e., published since 2015, the final year covered by the Kissi. et al., review) on rule-governed behavior, as pliance, tracking, and augmenting, has progressed experimental analyses within RFT. And it is to this issue we now turn.

Recent research focused on pliance, tracking and augmenting

For current purposes we broadly employed the same criteria as Kissi et al. (2017) in an ad-hoc manner to search for additional records that had been published since July 2015. As mentioned above, the reader should note that this search should not be considered as rigorous and systematic as that conducted in the original review. For example, we did not contact relevant experts for articles. Rather, the aim of this informal search was to get a general sense of how these terms had been employed in this period, focusing on the nature of their application, and whether they demonstrated clear utility in the experimental analysis of behavior. This search (conducted in April 2020) resulted in the retrieval of five experimental

papers. It is also worth noting that we did not apply the criterion by which the papers must be experimental in nature, and as a result four additional papers were retrieved. These papers could be classified as psychometric in orientation rather than experimental; psychometric publications targeting pliance, tracking, and augmenting were not present in the literature at the time of the original Kissi et al. review (ascertained through correspondence with the lead author of the review). First, we will turn to the three experimental papers published since the Kissi et al. review to examine their key aims and findings.[1]

Experimental research. As mentioned above, five experimental papers were identified since the publication of the systematic review. Three of these focused on pliance and/or tracking, with one also including augmenting: the first made an experimental distinction between pliance and tracking in the context of persistent rule-following (Kissi et al., 2018), the second focused purely on pliance in the context of rule persistence and mindfulness (O'Connor et al., 2019), while the most recent paper explored pliance, tracking, and augmenting within the context of its application to pedestrian traffic rules (Ruíz-Martínez, et al., 2019). The other two experimental studies focused primarily on augmenting: one assessed the impact of formative augmentals on consumer purchasing behavior (Fagerstrøm et al., 2015), while the other examined the impact of motivative augmentals on individual and cooperative behavior in the workplace under different financial performance contingencies (Rafacz et al., 2019). We will first turn to the former three studies.

The first study was conducted by Kissi, et al. (2018), the lead authors of which also authored the systematic review paper discussed above. Specifically, participants were divided

[1] We would also note that other papers have been published since our informal review of the literature (in April 2020) that could be seen as broadly relevant (e.g., Ghezzi et a., 2020; Ruiz et al., 2020). However, the general argument we are advancing in the current "position" article, that rule-governed behavior, and thus pliance, tracking and augmenting, are best considered middle-level terms, are not fundamentally challenged by the publication of these additional articles.

into three experimental groups: a pliance rule group, a tracking rule group, and a no instructions group. Participants in each group were then required to complete a task within which responding in accordance with the task contingencies was initially reinforced via points. For the ply and track groups, this instruction also initially corresponded with the initial task contingencies. For the no-rule group, participants had to learn the correct way to respond by trial-and-error. Mid-way through the task, an un-cued contingency reversal occurred after which previously accurate responding became inaccurate. Results showed that 1. participants in the pliance and tracking groups were more likely to persist with the rule in the face of reversed reinforcement contingencies than those in the no-rule group, and 2. those responding in accordance with a ply persisted for longer than those in the tracking group. The authors also reported that, while ultimately behaving in a more 'contingency sensitive' manner than the two active rule groups, participants in the no-instructions group did persist for longer than expected when the contingencies reversed. The authors suggested that this may be a result of participants constructing and following a 'self-rule'.

At first glance, this publication appears to yield data that looks broadly consistent with the concepts and their proposed functional independence. Upon closer inspection of the literature as a whole, however, this is only the second published paper to show a difference in rule-following between pliance and tracking, the first of which was by McAuliffe et al. (2014). However, it is interesting to note that in the McAuliffe et al. study, a differential effect for pliance and tracking was only recorded for participants who were classified as high in self-reported depressive symptomatology. Those from the normative group failed to show a significant difference between pliance and tracking. In this sense, therefore, the contrast between the study by McAuliffe et al. and Kissi et al. (2018) provide another example in which inconsistencies have emerged across published studies in the literature. This of course is compounded by the findings reported by Baruch et al. (2007) in which no differential

impact of pliance and tracking was found between high and low dysphoric participants. Thus, while the Kissi et al. study reports a difference between pliance and tracking conditions in a normal population, other studies have failed to find a similar difference. The lack of *consistent* cumulative experimental data for pliance and tracking effects in a similar direction or population thus further calls into question the utility of these concepts in experimental analyses.

The second study, conducted by O'Connor et al. (2019), employed a self-report measure of what they called 'generalized pliance' to explore the extent to which this predicted sensitivity to changing contingencies and self-reported mindfulness. In this sense, pliance was operationalized through a psychometric instrument rather than a direct experimental manipulation. Specifically, participants completed two contingency-switching tasks (i.e., the contingency-switching variant of the Iowa Gambling Task and Wisconsin Card Sorting Task) as well as the two self-report measures (for generalized pliance and mindfulness). Results showed that higher levels of pliance predicted greater levels of insensitivity to changing contingencies on both tasks, and lower levels of mindfulness. While this study obtained a correlation between a psychometric measure of pliance and contingency sensitivity (and mindfulness), it should be noted that there was no experimental (functional-analytic) manipulation of pliance. Furthermore, the research did not attempt to measure, psychometrically or otherwise, the other putative categories of rule-following (tracking and augmenting).

The third study was conducted by Ruíz-Martínez and colleagues (2019). In this study, participants were exposed to three experimental phases. In phase 1, all participants were first presented with a hypothetical pedestrian road crossing scenario followed by two options: one that adhered to correct pedestrian safety rules, and one that did not (e.g., waiting for the lights

to change so that crossing is safe versus crossing and trying to avoid cars) to assess general rule compliance or transgression. In phase 2, all participants were presented with another hypothetical situation and presented with 4 options for responding; three options described a situation in which rules were followed but each specified different consequences for rule-following, which were labelled as involving a ply, a track, or an augmental, while the fourth option described a transgressional response. Finally, in phase 3, participants were divided into three experimental groups (pliance, tracking, and augmenting). They were again presented with some hypothetical situations and asked whether or not they would comply with appropriate rule-following. If participants complied, they were reinforced differentially depending on the group to which they had been assigned. That is, in the pliance group, participants were presented with a positive consequence for rule-following that involved social reinforcement such as praise; in the tracking group, by following the rule participants could continue in their current behavior uninterrupted; in the augmenting group, by following the rule participants successfully dealt with a putative event that likely increased the reinforcing properties of getting to the other side of the road safely.

The results of the study showed that each rule condition produced increased rule compliance but clear differences among the rule conditions were not forthcoming. Thus, although this study had the advantage of including each of the three different types of rules in the experimental design, the results unfortunately did not yield clear evidence that the way in which the three rules had been operationalized impacted in a clear and systematic way on the extent to which participants engaged in rule-following. Once again, therefore, the concepts did not appear to function differentially and thus their utility remains somewhat opaque.

Having first reviewed these three experimental papers that have been published since Kissi et al.'s systematic review (2017), it seems that the literature is no further along in

demonstrating the functional-analytic value of the concepts of pliance, tracking, and augmenting in the experimental analysis of rule-governed behavior. While some evidence demonstrated a difference between pliance and tracking conditions, overall results continue to prove inconsistent across studies, and in some studies no clear differences emerged at all. As mentioned previously, the remaining two studies to which we will now turn focused primarily on augmenting.

The first, conducted by Fagerstrøm, et al. (2015), assessed the impact of formative augmentals on consumer purchasing behavior in the context of corporate social responsibility (CSR). Specifically, participants were presented with a number of scenarios in which they were provided with information about a piece of exercise clothing. Each scenario gave information about the quality of the product, the washing information, the brand, the price, and a formative augmental that specified consequences of purchasing the product; the consequences varied across the augmentals. Each augmental specified support for either (i) general ethical trading initiatives to assess the impact of general CSR activity, (ii) a near militant environmental group to assess the potentially negative impact of CSR activities, or (iii) an international movement supporting breast cancer awareness to assess whether specific positive CSR activity could differentially impact on consumer gender. Participants were asked how likely they were to buy the product described in each scenario and to rate how much each element influenced their decisions.

Results showed that the formative augmentals employed were generally less impactful on self-reported purchasing likelihood than product price, brand, and wash information, but more impactful than product quality. The nature of the consequence specified by the augmentals (i.e., general, specific negative, or specific positive) did differentially impact purchasing likelihood. Specifically, support for ethical trading (general) had a small positive

impact, breast cancer awareness (specific positive) had a relatively larger positive impact, and differentially impacted upon participant gender, while support for the near militant environmental group (specific negative) had a negative impact on purchasing likelihood. The authors concluded that the findings demonstrated the benefits that a behavior-analytic approach to rule-governed behavior can provide for a functional analysis of CSR on consumer purchasing behavior.

The next study, conducted by Rafacz et al. (2019), sought to explore the impact of motivative augmentals on individual and cooperative behavior in the workplace under different pay-for-performance contingencies. In an initial experiment, a multiple-trial implicit relational assessment procedure (MT-IRAP) was used to identify individual (e.g., "self-reliant") and cooperative (e.g., "partnership") stimuli that evoked, on average, the strongest positive responses for participants. These stimuli would be used within the motivative augmentals employed later. In a second experiment, a different group of participants were exposed to an organizational data entry analogue task in which participants could choose to work for themselves or to help a fictitious partner under piece-rate and profit-share conditions in a counterbalanced reversal design. During some trials, motivative augmentals were presented that specified consequences intending to increase individual responding despite profit-share contingencies (e.g., "remember that self-reliance is highly valued"). In other trials motivative augmentals were presented that specified consequences intending to increase cooperative responding despite piece-rate contingencies (e.g., "remember that being in a partnership is highly valued"). In the remainder of trials, statements employed were neutral with respect to pro-individual or pro-cooperative consequences (e.g., "remember that paying attention is highly valued").

The results showed that, in general, the financial contingencies at play had a greater impact upon participant performance than did the motivative augmentals. However, the augmentals did somewhat influence participant behavior when the contingencies were financially neutral (i.e., profit share contingencies), and there was no financial cost for doing so. The authors concluded that organizations could employ motivative augmentals in memos or speeches to increase worker performance, but that these would need to be consistent with financial contingencies for maximal benefit.

In reflecting upon the latter two studies, which involved focusing only on augmenting, both appear to indicate relatively clear effects for motivational variables. However, it could be argued that the concept of augmental rule control as employed in these studies adds very little beyond the common sense concept of motivation *per se*. Although this may seem like a harsh criticism, it is interesting to note that in the article by Rafacz et al. (2019), the phrase "motivational statement" is used far more frequently than the term augmental. Of course, one could argue that a "motivational statement" comprises the very definition of an augmental. In doing so, however, the question then emerges, what value is the concept of an augmental, as a technical term, if it can be substituted readily with an everyday common sense concept? Let us be clear, we are not questioning the value of these studies, the importance of the results they report, nor even the authors' use of the term augmental. We are simply asking to what extent the concept of augmental influenced the design and analysis of the research in some unique manner that made a clear difference from what the researchers would have done had they simply chosen to employ the concept of motivation (or motivational statement) alone?

Overall, therefore, the studies reviewed thus far either failed to provide clear and consistent effects, at least across studies, when the terms pliance and tracking were employed. Furthermore, the type of research that has employed these terms, and also

augmenting, do not tend to test the concepts themselves in a functional-analytic abstractive manner. For example, the term compliance could simply be used in place of pliance, and "motivational statement" could be substituted for augmental, without making any discernible difference to the research, at least as far as we can tell. To put it another way, the studies appeared simply to employ the concepts of pliance, tracking, and augmenting rather than testing them as functional-analytic abstractive concepts in and of themselves.

Psychometric-focused research. As mentioned above, four psychometric studies have appeared in the literature since the final retrieval of records was conducted by Kissi et al.'s review (2017). One focused on developing a measure of what the authors referred to as generalized pliance (Ruiz et al., 2019); another sought to examine the construct validity of this measure against two measures with which it is suggested to overlap conceptually (Waldeck et al., 2019); a third developed a children's version of the first measure (Salazar et al., 2018); while the final paper primarily sought to explore the convergent validity of the children's measure in adolescents (Stapleton & McHugh, 2020b).

The Generalized Pliance Questionnaire (GPQ; Ruiz et al., 2019) was designed, in part, as a response to the lack of empirical research in the literature that has focused on functional classes of rule-governed behavior (i.e., pliance, tracking, and augmenting). Specifically, the authors argued that one of the main reasons behind the limited experimental research employing these concepts, and the inconsistent findings produced where such research has occurred, may be due to the lack of a valid behavioral or self-report measure of pliance or tracking. They suggested that such a measure could thus prove useful with added potential to contribute to experimental psychopathology, longitudinal, and clinical research (interestingly, no reference is made directly to its potential for use in basic experimental-analytic RFT research).

Generalized pliance itself is suggested to occur if pliance as a class of rule governance over-generalises and social approval becomes the main source of reinforcement for an individual (Ruiz et al., 2019). Thus, the GPQ measure comprises 18 items (note that a shorter 9-item version is also available) that pertain to instances in which socially mediated consequences are the primary source of control over an individual's behavior (e.g., "My decisions are very much influenced by other people's opinions"; "My self-worth depends on what other people think and say about me"). Both the 18- and 9-item versions were reported to have excellent internal consistency and correlated with a range of other self-report measures that are associated with ACT such as, for example, cognitive fusion, mindfulness, and experiential avoidance. Subsequent research conducted by Waldeck et al. (2019) aimed to further validate the GPQ-9 against two other self-report measures that assess social reinforcement. The researchers reported good construct validity and psychometric properties of the measure. They also reported, however, that structural equation modelling revealed that the GPQ-9 was *not* significantly associated with measures of psychological flexibility and psychological distress (one of the other measures employed in the study, the Brief Fear of Negative Evaluation Scale, did significantly correlate).

The need for a children's version of the GPQ was argued by Salazar et al. (2018) on the conceptual basis that pliance is the first type of rule-governed behavior established in the behavioral repertoire of children. Consequently, it is argued that its overgeneralization could disrupt successful contact with environmental contingencies in favour of social approval and facilitate forms of psychological distress thought to be underpinned by rule-based contingency insensitivity. In principle, therefore, measurement of generalized pliance in children could aid in research exploring its relationship with the development and establishment of psychological distress. The GPQ-Children (hereafter referred to as the GPQ-C) adapted the 18-item adult version of the GPQ by rephrasing some of the statements so that

they could be more readily understood by children (e.g., “In order to be happy, I need people to value me” to “I need people to like me to feel happy”) and by removing some items that purely pertained to adult life (e.g., “If other people don’t value my work, I feel as though it was not worth the effort”). The study reported adequate internal consistency in a population of children aged 8-13 years, that girls obtained higher scores than boys, and that scores seemed to decrease with age. Finally, the authors reported that this measure correlated with a number of other self-report measures of, for example, psychological inflexibility and repetitive negative thinking. A subsequent study by Stapleton and McHugh (2020b), aimed to further explore the relationship between the GPQ-C and various self-reported measures (i.e., psychological inflexibility, positive and negative affect, empathic concern, and capacity to enjoy interpersonal interactions) in a population of adolescents aged 14-17 years. The authors reported support for the convergent validity of the GPQ-C and significant correlations with the other self-report measures.

Overall, the findings of the psychometric approach to research into pliance and its overgeneralization are encouraging. The measures are validated against measures of other concepts posited as core to the analysis and treatment of psychological suffering, particularly within ACT (e.g., psychological inflexibility and cognitive fusion). At this point, therefore, it could be argued that the concept of pliance has some validity in the wider context of ACT. On balance, ACT is not RFT, and the original Kissi et al. (2017) review correctly considered pliance, tracking, and augmenting as concepts directly relevant to RFT-based experimental analyses rather than ACT per se. Indeed, in the latter half of the quotation provided previously from the seminal RFT volume, the authors stated that in order to unpack a particular domain, such as rule-governed behavior, data that might reflect the functional processes involved in this domain would need to be collected. In the service of this aim, we will now consider a series of recently published studies that have attempted to explore a more

technical (i.e., basic experimental analysis) of rules as relational networks rather than instances of pliance, tracking, and augmenting. In doing so, we recognize that this research program is still very much in its infancy. Nevertheless, it seems useful to present it here because the approach employed, and the results obtained, serve to illustrate the type of distinction we are attempting to draw between basic lab-based RFT research and the middle-level driven and/or psychometric studies reviewed above.

Some examples of basic lab-based RFT studies of rule-governed behavior as relational networking

Within RFT, a conceptual analysis of rule-governed behavior involves defining rules as relational networks. Indeed, this analysis initially gained empirical support from two studies by O'Hora and colleagues (2004, 2014) that successfully modelled rules as derived relational networks in the experimental lab. In attempting to further explore a basic experimental analysis of rule-following, and in particular persistent rule-following, a series of studies have emerged over the past number of years investigating potential variables involved in derived rule-following in the face of competing reinforcement contingencies.

The first study: direct versus derived rule-following. Harte et al., (2017) aimed to investigate the impact of giving participants a direct rule versus a rule that involved a novel derivation on persistent rule following. Specifically, the researchers provided half of the participants with a direct rule, that did not involve responding in accordance with a novel derived relation generated within the experiment. That is, they were instructed to choose the *least similar* comparison stimulus presented in an MTS task. In contrast, the other half of the participants were presented with the same instruction except that the phrase "least similar" was replaced with a nonsense word that participated in a derived coordination relation with that phrase. In other words, participants were required to treat the new nonsense word as

meaning the same as the phrase “least similar”. The critical question was whether there would be a difference in persistent rule-following between these two groups of participants. For example, would participants who had been provided with a “direct” rule (i.e., that did not require any novel derivation within the experiment) produce higher persistence in rule-following than participants who were required to partially derive the meaning of the rule.

Persistent rule-following was assessed through an MTS task in which the rule was initially consistent with task feedback contingencies for responding (i.e., the rule required the participant to choose the least similar comparison and points were awarded whenever this comparison was chosen). After a certain number of MTS trials, however, the feedback contingencies reversed, after which points were awarded for choosing the most similar comparison. The researchers took a number of measures of the extent to which participants persisted with following the rule despite the fact that doing so now resulted in the loss of points on the MTS task. In general, the results showed that participants who had been provided with a direct rule persisted for longer than those provided with a rule that required deriving a relation within the experiment, but only when the opportunity to follow the rule before the contingencies reversed was relatively protracted (Experiment 2).

Manipulating derivation. In a subsequent study, Harte et al. (2018) explored the impact of a feature of derived relating highlighted by a framework that has emerged within RFT in recent years; the details of the framework are beyond the scope of this article but the reader is referred to Barnes-Holmes et al. (2017, 2020) for extensive treatments. Harte et al. explored the impact of level of derivation, which refers to the extent to which a particular pattern of relational responding has been emitted previously; derivation is defined as *reducing* each time a relation is derived. In other words, the first time a relation is derived, it is derived directly from the initial relation (high in derivation); as the relation is derived over

and over again, it acquires its own history making it less and less derived from the original relation (relatively low in derivation).

In the study by Harte et al. (2018), the researchers experimentally manipulated level of derivation and assessed its impact on a similar contingency switching MTS task as employed by Harte, et al (2017); that is, the task contingencies supported the derived rule initially but later reversed. In this study, all rules provided to participants now required a novel derivation within the experiment (i.e., there was no direct rule as in Harte et al., 2018). In one condition, participants had many opportunities to derive this novel relational response (low derivation), while in another condition participants had relatively few opportunities (high derivation). The researchers also explored the impact of mutual and combinatorial entailment in the study. That is, in one experiment, participants were required to derive that the phrase 'least similar' was equivalent to a nonsense word (i.e., A=B) many versus few times, while in a second experiment, participants were required to derive that 'least similar' was equivalent to a nonsense word through a middle node (i.e., A=B=C), also many versus few times. This relation was subsequently inserted into the rule required for responding on the contingency-switching MTS task. The results broadly indicated that lower levels of derivation produced greater persistence in rule-following. This was the case at both mutually and combinatorially entailed levels. That is, the more opportunities participants had to derive the rule, the more they persisted with rule-following when the MTS task contingencies no longer supported the rule.

Manipulating coherence (and derivation). In another two studies, researchers explored the impact of a second variable on persistent rule-following. Specifically, the researchers manipulated the coherence of the derived rule (Harte, Barnes-Holmes, Barnes-Holmes, McEnteggart, et al., 2020; Harte, Barnes-Holmes, Barnes-Holmes, & McEnteggart, 2020). In brief, coherence refers to the extent to which a pattern of derived relating is

consistent with or coheres with previously established patterns of such responding. In both studies, participants were again provided with rules that required a novel, within-experiment derivation, but the coherence of these rules was manipulated here through the provision versus non-provision of performance feedback for deriving the relations between the nonsense word and phrase 'least similar'. The researchers also manipulated level of derivation, but across experiments. That is, while all participants had the same number of opportunities to derive the relation within each experiment, they had relatively more or less opportunities across experiments (e.g., five blocks of training trials in one experiment compared to only one block of training trials in a second experiment). In other words, *within* each experiment, performance feedback was manipulated while the level of derivation of the derived relation remained constant, but *across* each experiment level of derivation varied. Results showed that feedback for deriving significantly impacted upon persistence in rule-following when derivation was high (i.e., fewer opportunities to derive), but not when relatively low (i.e., more opportunities to derive). That is, it appeared that the less derived the rule became, the less impact feedback for deriving had on rule-following within the MTS task. When the rule was relatively high in derivation, however, feedback for deriving significantly impacted upon MTS rule persistence.

It is worth noting at this point that in both of the foregoing studies, the researchers also employed the GPQ. In both cases, the GPQ successfully correlated with a specific aspect of rule-following, thus orienting the authors to some speculative conclusions in the General Discussion sections. We believe this provides an example of the point we made earlier in the article. That is, in the current example, incorporating a psychometric instrument into the experimental analyses yielded an interesting result certainly worthy of discussion. However, the development of the psychometric instrument, *per se,* was not a *substitute* for the basic experimental analysis of behavior that was the focus of these studies. Specifically, the

experimental analyses provided data that potentially reflected the functional processes involved in persistent rule-following as derived relational networks.

Manipulating coherence in a non-critical section of the relational network. The potential for exploring relatively subtle functional-analytic processes involved in persistent rule-following (as relational networking) was illustrated in a recent study reported by Bern et al. (2020). Specifically, these researchers further explored the impact of coherence on derived rule-following, but the researchers manipulated the coherence of a component of a relational network that was not directly involved in the derived rule following on the MTS task. Specifically, participants were first trained on a novel relational network that was either maximally coherent (i.e., train A=B=C=D=E=F; reinforce the derived F=D relation) or partially incoherent (i.e., train A=B=C=D=E=F; punish the derived F=D relation) before being exposed to a similar contingency switching MTS task as employed in the Harte et al. studies above. Critically, the coherence of the part of the network that was involved in deriving the rule for completing the subsequent MTS task was restricted to the A=B=C part of the network, and as such the coherence in this part of the network was essentially the same across the two conditions of the experiment. Nevertheless, results showed that undermining a "remote" part of the relational network (i.e., the D=E=F section) impacted upon persistent rule-following on the MTS task. Specifically, in the condition in which the coherence of the D=E=F part of the network was undermined, participants persisted with rule-following for significantly more trials following the contingency reversal than the maximal coherence group.

At this point it is important to note that in the Bern et al. (2020) study greater persistence in rule-following was observed in the condition deemed to be low in coherence. In contrast, in the Harte and colleagues studies outlined above (Harte, Barnes-Holmes, Barnes-Holmes, & McEnteggart, 2020; Harte, Barnes-Holmes, Barnes-Holmes,

McEnteggart, et al., 2020), the high derivation conditions that were also deemed to be high in coherence (i.e., with Feedback) produced more persistence (in rule-following). One possible explanation offered by Bern et al. (2020) is that attempting to reduce coherence by punishing coherent relational responding with regard to a non-critical part of the network may have reduced the coherence properties of feedback itself. Thus, when feedback was used to punish rule-consistent responding on the MTS task, following the contingency reversal, it was less effective in suppressing rule-following in the low coherence group, and therefore persistent rule-following became more likely in this group. More informally, the feedback presented during the MTS task was undermined somewhat because it had proved to be unreliable in an earlier part of the study when it was used to punish coherent relational responding. Of course, this explanation was speculative and post-hoc, but it does serve to highlight that systematic experimental analyses of persistent rule-following are required to more fully understand the functional-analytic processes that may be at play in this domain.

Exploring flexibility and coherence in relational networking with individual participants. While all of the research described thus far was conducted with group designs, a very recent study attempted to explore persistent rule-following as derived relational networking with individual participants. In doing so the research aimed to achieve the relatively high levels of behavioral prediction-and-influence, with precision, scope and depth, that are the hallmarks of functional-analytic research (Harte, et al., 2021). The researchers again focused on coherence but also flexibility. Specifically, the study explored the extent to which flexibility in reversing derived relations would control responding when the reversed relations cohered to relative degrees with the contingencies for rule-following. The first experiment focused entirely on flexibility in derived relational responding. Specifically, a relational network consisting of two combinatorially entailed relations were first established for 3 participants (i.e., A1=B1=C1 and A2=B2=C2). Next, the B and C relations in each

network were reversed, and the new relations were tested (i.e., A1=B1=C2 and A2=B2=C1). The experiment involved a number of these reversals, within participants, to determine the extent to which each participant would demonstrate flexible relational responding in the formation of a derived relational network. Results showed that each participant successfully demonstrated flexible relational responding (i.e., each participant produced a test performance that was in accordance with the most recently trained baseline relations).

In a second experiment, 3 additional participants were again first trained and tested in a novel derived relational network identical to the first experiment, immediately after which they were exposed to an MTS task that required responding in accordance with the derived rule. Next, participants were trained and tested in a similar relational network that again involved reversing the trained B-C and tested A-C relations (identical to Experiment 1 above). They were then re-exposed to the MTS task, which now provided contingency feedback that was *inconsistent* with the reversed derived rule. The goal here was to explore the extent to which a rule that contained a derived relation, *which had been reversed,* would override competing reinforcement contingencies. Results showed that all participants generally responded in accordance with the MTS task contingencies, rather than the derived rule. Thus, in this case, reversed derived relational responding (i.e., relational flexibility) did not necessarily control behavior in the face of competing MTS contingencies.

A third and final experiment partially replicated the second, but the MTS procedure this time involved a reversal in task contingencies. Specifically, following the same training and testing of the reversed relational network as in the first two experiments, the first half of MTS trials were also reversed (i.e., consistent with the reversed derived rule), but for the final half of MTS trials, the task contingencies reversed again (i.e., now inconsistent with the reversed derived rule). The goal here was to determine whether initially increasing the coherence between the reversed derived rule and MTS task contingencies would produce

greater persistence in rule-following after the contingencies reversed. In this case, evidence of rule persistence emerged in all 3 participants.

In their Discussion of the results, the authors argued that two competing relational networks were at play in the experiment: the derived rule emerging from the training and testing procedure and a rule that likely emerged through direct contact with the MTS task contingencies. In Experiment 2, the MTS contingencies remained unchanged throughout, and therefore this network could be defined as maximally coherent. As a result, it provided a relatively strong source of behavioral control. In contrast, because the network that emerged as a result of the derived relations training and testing procedure reversed multiple times throughout the experiment, the coherence of this network relative to the MTS network was greatly reduced. In other words, participants responded in accordance with the most coherent network (i.e., the one that never changed), and therefore failed to show any evidence of rule-persistence in the face of competing MTS contingencies.

In Experiment 3, however, the MTS contingencies also reversed within the experiment. In this case, therefore, the coherence of the MTS-generated network was much reduced relative to Experiment 2. As such, relative to Experiment 2, the behavioral control properties of the reversed derived rule network may have been greater, and as such was now competing with an MTS network that also contained some level of incoherence. Additionally, the authors argued that the coherence of the reversed derived rule network in Experiment 3 increased, relative to Experiment 2, because it initially cohered with the MTS contingencies following the reversal. Thus, not only did the MTS network *decrease* in coherence, the reversed derived rule network *increased* in coherence, which resulted in rule-persistence in Experiment 3. Once again, while the explanation provided by the authors was speculative and post-hoc, it nonetheless illustrates the potential complexities involved in (persistent) rule-

following that seemingly extend beyond the rather "blunt" middle-level concepts of pliance, tracking and augmenting (as middle-level terms).

Concluding comments. The foregoing experimental analyses of rule-following highlight what appear to be highly complex and subtle effects. In an attempt to grapple with these subtleties, research has involved exploring the impact of experimentally definable, relational variables like coherence, derivation, and flexibility, for both mutually and combinatorially entailed relations (contained within a 'rule' as a relational network), on responding in the face of competing reinforcement contingencies. The findings illustrate the many behavioral complexities likely involved in rule-following, complexities that may remain largely hidden when researchers rely too heavily on middle-level concepts like rules or its related subsets (i.e., pliance, tracking, and augmenting) in the experimental analysis of behavior.

The foregoing series of studies also serve to illustrate the distinction we are attempting to make between basic lab-based RFT research, driven by highly technical concepts (e.g., derived relational networks), and the research reviewed earlier in the paper, which was driven by the concepts we are arguing are best considered middle-level terms (i.e., pliance, tracking, and augmenting). In response, perhaps it could be argued that the experimental research conducted on pliance and tracking thus far has simply not been sophisticated or creative enough in measuring whether participants, irrespective of rule content, are following a ply or a track. In other words, perhaps the researchers were not differentiating clearly enough between the stimulus (rule) and stimulus functions (the pliance or tracking functions actualized in the listener when following the rule). Indeed, Ruiz et al. (2019, 2020) cite this as one of the reasons for creating the Generalized Pliance and Tracking Questionnaires. While recognizing that the development of psychometric instruments may be

of some benefit in this regard[2], once again we would stress that they should not be treated as a substitute for the basic lab-based experimental analyses that are needed to better understand rule-following (persistent or otherwise). In this respect, we would argue that the problem in developing a productive program of experimental research on rule-governed behavior within RFT may not be lack of creativity or sophistication on behalf of the researchers, but the lack of appropriate stimulus control over the behavior of the researcher afforded by the concepts of pliance, tracking, and augmenting. In short, the problem may be more likely to be in the concepts, not the researcher (i.e., to use a well-worn behavior-analytic maxim -- the rat is always right).

General conclusion

Kissi et al.'s systematic review (2017), of the terms pliance, tracking, and augmenting as functionally independent classes of rule-governed behavior within the experimental analysis of human language and cognition suggested that relevant research had been scarce and largely conflicting in nature, at least up until July 2015 (i.e., the date of the final retrieval of articles relevant for the review). In reviewing the literature (informally) since then, the same conclusions seem to apply. Indeed, the recent emergence of traditional psychometric research involving these concepts could be seen as reinforcing the original conclusions, in that researchers may be recognizing that pliance, tracking and augmenting may be of limited value in the *experimental* analysis of behavior. In drawing this conclusion, again we are not suggesting that the terms are without value, but it is important to recognize the boundaries and limitations of the scientific terms we employ as basic and applied researchers, and as

[2] Parenthetically, it is also worth noting that some of the leading researchers who have argued for the value of middle-level terms in behavioral psychology have recently questioned the use of traditional psychometry itself in favor of idiographic, process-based research (Hayes et al., 2019). Thus, the use of nomothetic psychometrics in support of the concepts of pliance, tracking and augmenting could be challenged further in the newly emerging idiographic, process-focused era.

practitioners. Recognizing such boundaries will always be important because as Skinner noted in 1953, “We divide behavior into hard and fast classes and are then surprised to find that the organism disregards the boundaries we have set” (p. 94). Insofar as the concepts of pliance, tracking, and augmenting have failed to yield the levels of prediction-and-influence (with precision, scope, and depth) we would typically seek in the experimental analysis of behavior, we would argue it seems best to treat them as clear examples of middle-level concepts. As such, it is now time to develop and employ additional concepts, not necessarily to the exclusion of pliance, tracking and augmenting, but which are wrought from the crucible of the basic experimental research laboratory. It would be unfortunate if RFT researchers were reticent in recognizing the potential limitations of concepts that appear in the seminal text (Hayes, et al., 2001), when those concepts fail to yield readily to experimental analyses. Indeed, this would be particularly ironic given that early advocates of RFT criticized Skinner’s (1957) volume, *Verbal Behavior,* for failing to produce a vibrant and progressive program of basic experimental research.

Declarations

Conflict of Interest: The authors declare they have no conflict of interest

Funding: Not applicable

Ethics approval: Not applicable

Consent to participate: Not applicable

Availability of data an material: Not applicable

Code availability: Not applicable

References

Barnes-Holmes, D., Barnes-Holmes, Y., Luciano, C., & McEnteggart, C. (2017). From IRAP and REC model to a multi-dimensional multi-level framework for analyzing the dynamics of arbitrarily applicable relational responding. *Journal of Contextual Behavioral Science, 6*(4), 473-483. https//:doi.org/10.1016/j.jcbs.2017.08.001

Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2020). Updating RFT (more field than frame) and its implications for process-based therapy. *The Psychological Record*. https//:doi.org/10.1007/s40732-019-00372-3

Barnes-Holmes, Y., Hussey, I., McEnteggart, C., Barnes-Holmes, D., & Foody, M. (2016). Scientific ambition: The relationship between relational frame theory and middle level terms in acceptance and commitment therapy. In R.D. Zettle, S.C. Hayes, D. Barnes-Holmes & A. Biglan (Eds.), *The Wiley Handbook of Contextual Behavioral Science* (pp. 365-382). Wiley-Blackwell.

Barnes-Holmes, D., O'Hora, D., Roche, B., Hayes, S.C., Bissett, R.T., & Lyddy, F. (2001). Understanding and verbal regulation. In S.C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 103-118). Plenum.

Baruch, D.E., Kanter, J.W., Busch, A.M., Richardson, J.V., & Barnes-Holmes, D. (2007). The differential effect of instructions on dysphoric and nondysphoric persons. *The Psychological Record, 57*, 543-554. https://doi.org/10.1007/BF03395594

Bern, R., Persdotter, T., Harte, C., & Barnes-Holmes, D. (2020). Relational coherence and persistent rule-following: The impact of targeting coherence in a 'non-critical' component of a relational network. *The Psychological Record.* Advance online publication. https//:doi.org/10.1007/s40732-020-00414-1

Fagerstrøm, A., Stratton, J.P. & Foxall, G.R. (2015). The impact of corporate social responsibility activities on the consumer purchasing situation. *Journal of Organizational Behavior Management, 35,* 184-205. https://doi.org/10.1080/01608061.2015.1093053

Finn, M., Barnes-Holmes, D., & McEnteggart, C. (2018). Exploring the single-trial-type-dominance-effect in the IRAP: Developing a differential arbitrarily applicable responding effects (DAARE) model. *The Psychological Record, 68*(1), 11-25. https://doi.org/10.1007/s40732-017-0262-z

Finn, M., Barnes-Holmes, D., McEnteggart, C., & Kavanagh, D. (2019). Predicting and influencing the single-trial-type-dominance-effect: The first study. *The Psychological Record, 69*(3), 425-435. https://doi.org/10.1007/s40732-019-00347-4

Ghezzi, E.L., Houmanfar, R.A. & Crosswell, L. (2020). The motivative augmental effects of verbal stimuli on cooperative and conformity responding under a financially competing contingency in an analog work task. *The Psychological Record, 70*, 411–431. https://doi.org/10.1007/s40732-020-00400-7

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.* https://doi.org/10.1007/s40614-020-00256-w

Harte, C., Barnes-Holmes, Y., Barnes-Holmes, D., & McEnteggart, C. (2017). Persistent rule-following in the face of reversed reinforcement contingencies: The differential impact of direct versus derived rules. *Behavior Modification, 41*(6), 743-763. https//:doi.org/10.1177/0145445517715871.

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2018). The impact of high versus low levels of derivation for mutually and combinatorially entailed

relations on persistent rule-following. *Behavioural Processes, 157,* 36-46. https//:doi.org/10.1016/j.beproc.2018.08.005.

Harte, C., Barnes-Holmes, D., Barnes-Holmes, Y., McEnteggart, C., Gys, J., Hassler, C. (2020). Exploring the potential impact of relational coherence on persistent rule-following: The first study. *Learning and Behavior, 48,* 373-391. https//:doi.org/10.3758/s13420-019-00399-0

Harte, C., Barnes-Holmes, D., Moreira, M., de Almeida, J.H., Aparecida-Passarelli, D., & de Rose, J.C. (In press). Exploring a Training IRAP as a single participant context for analyzing reversed derived relations and persistent rule-following. Journal of the Experimental Analysis of Behavior.

Hayes, S.C. (2016). Why contextual behavioral science exists: An introduction to part I. In R.D. Zettle, S.C. Hayes, D. Barnes-Holmes & A. Biglan (Eds.), *The Wiley Handbook of Contextual Behavioral Science* (pp. 9-16). Wiley-Blackwell.

Hayes, S. C., Strosahl, K., & Wilson, K.G. (1999). *Acceptance and Commitment Therapy: An experiential approach to behavior change.* Guilford Press.

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., Hofmann, S.G., Stanton, C.E., Carpenter, J.K., Sanford, B.T., Curtiss, J.E., & Ciarrochi, J. (2019). The role of the individual in the coming era of process-based therapy. *Behavior Research and Therapy, 117,* 40-53. https://doi.org/10.1016/j.brat.2018.10.005

Kissi, A., Hughes, S., De Schryver, M., De Houwer, J., & Crombez, G. (2018). Examining the moderating impact of plys and tracks on the insensitivity effect: A preliminary investigation. *The Psychological Record, 68,* 431-440. https://doi.org/10.1007/s40732-018-0286-z

Kissi, A., Hughes, S., Mertens, G., Barnes-Holmes, D., De Houwer, J., & Crombez, G. (2017). A systematic review of pliance, tracking, and augmenting. *Behavior Modification, 41*(5), 683-707. https://doi.org/10.1177/0145445517693811

McAuliffe, D., Hughes, S., & Barnes-Holmes, D. (2014). The dark-side of rule governed behavior: An experimental analysis of problematic rule-following in an adolescent population with depressive symptomatology. *Behavior Modification, 38*(4), 587-613. https://doi.org/10.1177/0145445514521630

O'Connor, M., Byrne, P., Ruiz, F.J., & McHugh, L. (2019). Generalized pliance in relation to contingency sensitivity and mindfulness. *Mindfulness, 10*, 833-840. https://doi.org/10.1007/s12671-018-1046-5

O'Hora, D., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2004). Derived relational networks and control by novel instructions: A possible model of generative verbal responding. *The Psychological Record, 54*, 437-460. https://doi.org/10.1007/BF03395484

O'Hora, D., Barnes-Holmes, D., & Stewart, I. (2014). Antecedent and consequential control of derived instruction-following. *Journal of the Experimental Analysis of Behavior, 102* (1), 66-85. https://doi.org/10.1002/jeab.95.

Rafacz, S.D., Houmanfar, R.A., Smith, G.S., & Levin, M.E. (2019). Assessing the effects of motivational augmentals, pay-for-performance, and implicit verbal responding on cooperation. *The Psychological Record, 69,* 49-66. https://doi.org/10.1007/s40732-018-0324-x

Rehfeldt, R.A. & Hayes, L.J. (1998). The operant-respondent distinction revisited: Toward and understanding of stimulus equivalence. *The Psychological Record, 48*, 187-210. https://doi.org/10.1007/BF03395266

Ruiz, F.J., García-Martín, M.B., Suárez-Falcón, J.C., Bedoya-Valderrama, L., Segura-Vargas, M.A., Peña-Vargas, A., Peña-Vargas, A., Henao, A.M., Ávila-Campos, J.E. (2020). Development and initial validation of the Generalized Tracking Questionnaire. *PLoS ONE, 15*(6): e0234393. https://doi.org/10.1371/journal.pone.0234393

Ruiz, F.J., Suárez-Falcón, J.C., Barbero-Rubio, A., & Flórez, C.L. (2019). Development and initial validation of the generalized pliance questionnaire. Journal of Contextual Behavioral Science, 12, 189-198. https://doi.org/10.1016/j.jcbs.2018.03.003

Ruiz-Martínez, Y., Hernández-Buitrago, P., & Quiroga Baquero, L.A. (2019). Effects of differential training in rule-governed behavior on compliance with traffic rules by pedestrians. *Interamerican Journal of Psychology, 53*(3), 399-416. https://doi.org/10.30849/rip/ijp.v53i3.883

Salazar, D.M., Ruiz, F.J., Flórez, C.L., & Suárez-Falcón, J.C. (2018). Psychometric properties of the generalized pliance questionnaire-children. *International Journal of Psychology and Psychological Therapy, 18*(3), 273-287.

Sidman, M. (1994). Equivalence relations and behavior: A research story. Authors Cooperative.

Skinner, B.F. (1953). *Science and Human Behavior.* Macmillan.

Skinner, B.F. (1957) Verbal Behavior. Appelton-Century-Crofts.

Skinner, B.F. (1966). An operant analysis of problem solving. In B. Keinmuntz (Eds.), Problem-solving: Research, method, and therapy (pp. 225-257). Wiley.

Stapleton, A. & McHugh, L. (2020). Adolescent correlates of the generalized pliance questionnaire - children. *Journal of Contextual Behavioral Science, 15,* 131-134. https://doi.org/10.1016/j.jcbs.2019.12.006

Stapleton, A. & McHugh, L. (2020). Just ask! Importance of qualitative explorations of children's rule-following. Children's Research Digest "Children Should be Seen AND Heard", 6(2), 1-5.

Waldeck, D., Pancani, L., & Tyndall, I. (2019). An examination of the construct validity of the generalized pliance questionnaire. *Journal of Contextual Behavioral Science, 11,* 50-54. https://doi.org/10.1016/j.jcbs.2018.12.003

Zettle, R.D. & Hayes, S.C. (1982). Rule-governed behavior: A potential theoretical framework for cognitive-behavior therapy. In P.C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (Vol. 1: pp. 73-118). Academic.