Title
What *Behavioral Variability* Means in Molecular, Molar, and Unified Analyses

Journal
International Journal of Comparative Psychology, 27(2)

ISSN
0889-3675

Author
Shimp, Charles Patterson

Publication Date
2014

DOI
10.46867/ijcp.2014.27.02.08

INTERNATIONAL JOURNAL OF COMPARATIVE PSYCHOLOGY

2014, 27 (2), 224-247
A. Blaisdell & W. D. Stahlman
Special Issue
Peer-reviewed

How Molecular, Molar, and Unified Analyses Change the Meaning of *Behavioral Variability*

Charles P. Shimp
University of Utah, USA

What effects reinforcement is assumed to have and what data are collected depend on what *behavioral variability* means. It has extremely different meanings in molecular, molar, and unified behavior analyses. In molecular analyses the term relates reinforcement and moment-to-moment behaving of an individual organism, as when hand shaping creates new complex patterns extended in time or as when cumulative records show complex patterns. Molecular behavioral variability is easy to see, as in these two examples, but is hard to describe quantitatively. Behavioral variability in the context of molar analyses requires first aggregating behaviors, then counting them or finding their cumulative durations, and finally quantitatively summarizing the aggregate by a statistic, usually an average rate of occurrence of, or an average time allocated to, the aggregated behaviors. The statistic can also be a measure of variability, like the U statistic, rather than of central tendency. Molar behavioral variability can also be quantitatively defined as the variability of a statistic describing some property of an aggregate as a function of time, individuals, or, most commonly, experimental parameters. Some molar accounts interpret the aggregate statistic itself (average rate, time allocation, or variability) as an operant response. Quantitative theories account for over 90 percent of this kind of variability in thousands of molar analyses. Molar variability, however, seldom describes or explains any molecular variability, and a common molar interpretation of free-operant behaving is that molecular behavior varies only randomly over time with a constant probability. This interpretation assumes the absence of any shaping effect on the temporal organization of behavior. There is little, if any, evidence for this interpretation and a considerable literature that suggests it is incorrect. A unified analysis combines automated shaping of molecular, quantitative patterns of behaviors, a molar aggregate of those patterns, and one or more statistics descriptive of the aggregate. A unified analysis involves both kinds of quantitative behavioral variability: moment-to-moment variability of shaped patterns resembling target patterns, and molar variability of a statistic defined over an aggregate of such shaped patterns, such as the variability of the average rate of, or time allocated to, a shaped pattern. Only simulation theories seem sufficiently powerful to produce a general and unified theory to account for both moment-to-moment behaving and statistics that describe molar aggregates.

B. F. Skinner promoted two sharply different ideas about behavior under control of its reinforcing consequences. One idea is conceptual and speculative; the average rate of occurrence of a free operant estimates response probability which in turn estimates theoretical response strength (Skinner, 1950) and thereby puts the Law of Effect (Thorndike, 1911) on a quantitative foundation (Herrnstein, 1970; Killeen & Hall, 2010). The other idea is methodological and practical; the experimental method of successive approximations, or shaping, creates new, complex, responses that are unified behavioral patterns extended in time (Peterson, 2004). These two contrasting ideas have led to radically different *molecular* and *molar* perspectives on what behavioral variability is and on what the effects of reinforcement are. That is, what behavioral variability means depends critically on whether it occurs as part of a molecular or a molar analysis. My first goal must therefore be to clarify what molecular and molar analyses are. To anticipate what follows below, I recommend thinking of molecular analyses as creating new moment-to-moment, temporally-extended patterns of behavior. These analyses involve how reinforcement reorganizes behavior in time (Morse, 1966; Peterson, 2004). I suggest molar analyses are usefully thought of in terms of how reinforcement generates more or less of certain activities over a period of time (Baum & Rachlin, 1969; Herrnstein, 1970); they aggregate these activities and then compute a statistic to summarize how much or how little of them occur over a period of time. My second goal is to describe implications of these two analyses for the meaning of

Please send correspondence to Charles P. Shimp, Psychology Department, 380 South 1530 East, Room 502, University of Utah, Salt Lake City, Utah 84112-0251. (charlie.shimp@psych.utah.edu)

behavioral variability. My third goal is to describe a third, rarer kind of analysis that shares some of the characteristics of both molecular and molar analyses, and in that sense connects or partially unifies them; this kind of analysis generates both molecular and molar kinds of behavioral variability. The overarching purpose of this paper is to clarify how the meaning of the term *behavioral variability* depends on the kind of analysis in which it appears, and to clarify how the meaning of reinforcement therefore also depends on the kind of analysis that is conducted.

Skinner's advocacy of two very different ideas about the effects of reinforcement, in terms of shaping and strengthening for a time fell on entirely receptive ears; both were seen as necessary components of a general theory of behavior (Morse, 1966). A shift took place in the 1960's in the Harvard *pigeon lab*, however, when its supervision passed from him to R. J. Herrnstein, who for the first time put into practice the molar, quantitative possibilities implicit in Skinner's idea about the relations among response rate, response probability, and response strength. Herrnstein's influence was so widespread that it may be only fair to call the analysis he initiated *The Herrnstein School of Molar Behaviorism*. This shift has been described several times with different kinds of emphases (e.g., Baum, 2002a; Catania, 2002; Skinner, 1976) but to my knowledge, how this shift changed the broader conceptual and empirical meanings of behavioral variability *per se* has not been explicitly examined. As we will see, this shift altered the balance of scholarly efforts away from how reinforcement shapes behavior and toward how it strengthens it.

The terms *molecular* and *molar* have many different meanings in the literature, and many different empirical interpretations. I do not review them all here because my goal is to shed light on the terms' core meanings and origins rather than to review myriad variations. I hope some will find the idea of a *unified* analysis attractive because it may facilitate showing how molecular and molar analyses can mutually inform each other rather than motivate claims that one approach is simply better than the other. Let us begin by examining what molecular analyses are and the sense in which they create and display behavioral variability.

Molecular Analyses of Behavior

A good example of a molecular analysis is Skinner's hand shaping of *pigeon ping pong* (see Shimp, 2013). Some readers may blink after reading the previous sentence. How can an episode of shaping that creates new behavior be an analysis of anything? Some readers may see shaping as that which an analysis should explain. Consider, however, that shaping is an experimental method to create new behaviors and it adapts to a subject's changing behavior as the approximation between behavior and target improves. An experimenter using shaping observes the subject, judges how well his method is narrowing the gap between a subject's behavior and his target pattern (which in principle can be extremely general – see Skinner, 1979), and judges whether and how to change his method to improve his outcome. It is, or at least includes, a moment-to-moment analysis. Shaping is an example of an *interaction* between a behavior analyst and his experimental subject; it displays variability in the behavior of behavioral scientists interacting with their subjects. The practice of behavior analysis obviously includes many different kinds of moment-to-moment behaviors (Skinner, 1959), and Skinner often emphasized that the science of behavior applies to behavior analysts (Skinner, 1948, 1959, 1979) and to scientists in general. He did not draw a sharp distinction between scientific analysis and the behavior of scientists. An excellent example is his *A Case History in Scientific Method* (Skinner, 1959), which appeared in an American Psychological Association collection of chapters written by eminent psychologists who were asked to describe their approach to science. It is instructional to compare his chapter in this collection (Koch, 1959) with that of Estes (1959), who tried to formalize the science of learning and made a sharp distinction between abstract theory, on the one hand, and behavior by experimenters and theorists, on the other hand. There are critical philosophical differences between Skinner's and Estes's approaches. Estes' approach might not have upset many positivist philosophers. Skinner's, however, might have, because he saw science as a social project that cannot be understood without understanding how shaping creates the behavior of behavior analysts and of scientists in general (Shimp, 2001). I agree with Skinner's epistemology and think shaping is a kind of real-time, qualtitative analysis where an experimenter keeps making observations, judgments, and discriminations to revise his methodology as a result of the behavior he has so far generated. For these and other reasons, in this paper, I therefore call shaping an analysis. See Hackenberg (2013), Palmer (2013), and Schlinger (2013) for related views.

The end product of Skinner's hand shaping of pigeon ping-pong can be seen on the web (http://www.youtube.com/watch?v=vGazyH6fQQ4). In this demonstration, Skinner trained each of two pigeons to strike a ping pong ball with its beak in a manner that prevented an oncoming ball from rolling off its end of a table and made it roll back toward the other pigeon on the other side of the table. A reinforcer was given a pigeon when the other pigeon failed to return the ball rolling toward its edge of the table. Skinner's hand shaping produced two interacting, continuous moment-to-moment *behavior streams* (Schoenfeld & Farmer, 1970), which together defined pigeon ping-pong. Consider this example in terms of Skinner's two ideas. Did he aggregate different instances of a behavior on behalf of his notion that amount of an activity over a period of time estimated response probability? Did he count anything, plot any functions, or develop a quantitative theory to explain response strength? All the answers are, of course, no; it is not a molar analysis. As a result, is his demonstration of shaping by today's scientific standards a dismal failure? Does it only reflect a primitive, pre-scientific nature of behavior analysis before the implications for quantification inherent in his other idea, about response probability, were put into practice? Some might think so. If quantification of aggregated response strength or aggregated time allocation is thought to be at the heart of modern behaviorism, and many molar behaviorists appear to believe it is (Baum, 2002a, b; Killeen & Hall, 2010; Rachlin & Laibson, 1997; Wilson & Herrnstein, 1985), and if moment-to-moment behaving is seen as too idiosyncratic and too variable to be lawful, and again, some believe it is (Jensen, Miller, & Neuringer, 2012), then Skinner's demonstration of pigeon ping pong might seem conceptually outdated. It might even seem too much like a video game to be taken as serious science.

There are still more reasons why a molar position might find hand shaping problematic and difficult to describe or explain. Not only did pigeon ping pong not lead to a quantitative analysis, hand shaping in general seems qualitative (Galbicka, 1994). It might very loosely be described as having a quantitative feature by describing it as increasing the occurrence of behavior similar to the target goal; hand shaping would then satisfy the textbook definition of the quantitative strengthening effect of reinforcement according to which a reinforcer increases the occurrence, or probability, of the behavior that precedes it. However, this use of *probability* fails to satisfy the relative frequency interpretation of probability (Feller, 1950) because hand shaping reorganizes the temporal patterning of behavior in ways difficult to describe quantitatively. This *probability* does not satisfy the transition to quantitative behavior analysis that took place when Herrnstein replaced Skinner as the director of the Harvard pigeon lab. Lastly, despite its very general applicability, hand shaping is certainly not universally applicable. The general constraints on shaping have been extensively discussed but are not yet well understood; they define one of the pressing issues in behavior analysis (Athens, Vollmer, & Pipkin, 2007; Breland & Breland, 1966: Galbicka, 1994) and imply that our understanding of shaping is definitely provisional. Hand shaping might therefore look relatively unimportant to a molar theorist already predisposed to see it as a tangential method, perhaps best suited only to creating in advance of an experiment the relatively simple and not very temporally extended key pecks or lever presses necessary to the actual experiment.

For these and other reasons, the use of hand shaping to create new complex patterns extended in time became less frequent in behavior analytic theory during the 1970's as Herrnstein's emphasis on quantitative theory of aggregated behavior became more prominent. This shift closely followed a broader shift in experimental psychology toward greater quantification (Bush & Estes, 1959; Bush & Mosteller, 1955; Catania & Reynolds, 1968; Estes, 1950). The founding in 1978 of the Society for the Quantitative Analyses of Behavior reflected and further promoted this growing emphasis on numbers.

If hand shaping appears so scientifically feeble from a molar perspective, is there anything that can be said on its behalf? Let us first remind ourselves of the purpose of shaping. It is to create a desired new response, a new unified, hierarchically-organized pattern extended in time (Morse, 1966; Peterson, 2004; Shimp, 2013). Because of the power of shaping to create so many different kinds of temporal patterns, Skinner (1976) argued that moment-to-moment, molecular analyses were actually the most important and useful kinds of behavioral analyses. In this he was saying something very different from his idea that the average rate of a free operant estimates response probability and therefore estimates response strength, thereby providing the basis for a quantitative analysis. Hand shaping seems irrelevant to that kind of quantification, and the difficulty in quantifying it only deepens when we grant that it presumably involves many different interacting component behavioral processes (Morse, 1966). For present purposes it is convenient and sufficient to write simply of the shaping effect of reinforcement to reorganize behavior extended in time as a process, but it should not be forgotten it presumably is a complex collection of processes.

Let us continue to ask, if shaping is so difficult to quantify, and has these other limitations, why should anyone view it as scientifically useful? Is there an adequate response to a molar critic who is doubtful that shaping is worthy of intensive study? There are several reasons why I believe it is. First, it is about powerful experimental control of behavior. Control of phenomena has been key to improved measurement methods and to progress elsewhere in science and there seems little reason to expect control will be less so in the development of a science of behavior. I am not saying control is the essence of science but I am saying that it would be strange to ignore its usefulness if it were available. Second, it can create and control everyday behavior of individuals, and many people spend most of their time trying to create new patterns of behavior extended in time. This includes teachers, parents, coaches, and mentors of all sorts. In reorganizing everyday behavior streams and in creating new, everyday moment-to-moment behavior, the shaping process contrasts with how the strengthening process increases or decreases the quantitative amount of the same, already-established activity: people pay to see shaped behavior (Breland & Breland, 1966) but seldom pay to see a statistic that describes how an aggregate of identical performances, like button pressing or standing on one side of a chamber or the other, increases or decreases. (This is not to say such statistics are irrelevant to the development of a science of behavior. They can be very relevant if we want to increase or decrease the amount of some previously established activity that can be meaningfully aggregated.) Third, pigeon ping pong illustrates that a molecular, moment-to-moment analysis is not restricted to brief periods of responding; the term *molecular* refers to moment-to-moment behaving extended for *any* period of time, not just short times. Similarly, cumulative records can picture moment-to-moment behaving extended for any duration of time, including very long durations. However, just because shaping and cumulative records can deal with behaviors extended in time does not mean they qualify as molar analyses because they do not directly generate or picture numbers that describe aggregated behaviors.

Behavioral Variability in Molecular Analyses: Concepts and Constraints

It is important not to underestimate the broader historical and conceptual justification for studying behavioral variability in molecular contexts. The justification ranges far afield, and includes discussions about the sequential and logical structure of natural language, music composition and performance, speech recognition, production, and synthesis, and so on and on. It is hard to think of un-aggregated operant behaviors that do not involve the shaping effect on behavioral variability. Let us therefore look more closely at our two examples of moment-to-moment variability.

Hand Shaping

Anyone reading this paper is likely to be able to develop effective hand shaping methods, to teach clearer pronunciation by a third grader, to improve observation of the ball by a high school baseball player who strikes out too often, and even to teach more musical singing by members of a choir. In imagining how shaping might work in such cases, some readers might think about the potential applicability of modern technology. Voice recognition technology might aid in the case of teaching pronunciation, high speed photography might aid in changing how a batter watches the ball, and so on, and this in turn might encourage thinking about automated shaping (Galbicka, 1994, and see below). It is easy to start imagining how to apply shaping and its practical utility seems established beyond any doubt.

Problems arise, however, when we try to specify hand shaping quantitatively because behavior constantly changes when being shaped to more precisely resemble target behavior. Another problem is that we do not know how an experimenter judges whether to deliver a reinforcer after a particular pattern because we do not know how to describe or quantify the proximity between an emitted pattern and a target pattern. That is a conceptual and methodological problem in the discrimination of complex performances that is beyond present understanding, especially for multidimensional target behaviors. More generally, as I observed above, shaping can be viewed not as a single process but as a complex, dynamic combination of several, including stimulus control, categorization, and memory (Morse, 1966; Zeiler, 2006; Marr, 2004, 2012). Quantification of hand shaping is all the more difficult because this already complex system relies also on an experimenter's idiosyncratic skills. There have been attempts to address some of these problems (Athens, Vollmer, & Pipkin, 2007; Blough, 1966; Galbicka, 1994; Galbicka, Kautz, & Jagers, 1993; Galbicka & Platt, 1986; Hawkes & Shimp, 1975, 1998; Machado, 1989; Platt, 1973), but Marr (2012) has convincingly described past efforts as too limited to unusually simple cases, compared to the complexity of much every day behavior.

From the perspective of behavioral variability, molecular analysis in the form of hand shaping may be, as Skinner (1976) claimed it is, the most useful behavioral control method, and it is technologically simple and available to many. To repeat, however, from the perspective of quantitative theory it is not very sophisticated or scientific. It is also limited to a midrange of complexity. In the case of extreme complexity, we do not know, for example, how even to conceptualize teaching someone to play the violin in terms of shaping, although I believe considerable progress would be made if more concerted efforts, using modern technology, were focused on that goal. At the opposite extreme, as we will see below, we do not know how to shape what might be viewed as the simplest of all performances, random behavior; shaping the variability of an individual's behavior over time that corresponds to temporally random behavior has proven to be elusive.

In summary, hand shaping is an excellent method to create new moment-to-moment patterns of behavior, and it clearly illustrates Skinner's idea that reinforcement shapes and reorganizes behavior in time in ways that are not described by theories that emphasize instead how reinforcement changes the amount of an already established behavior.

Cumulative Records

Cumulative records picture moment-to-moment behavioral variability and make it easy to see patterns, but they involve quantification in only a very minimal way in the sense that each response moves a pen a fixed distance across a piece of paper so that the total distance traveled before the pen resets reflects a specified number of responses, and the slope of the line estimates average response rate (although see Gallistel, Balsam, & Fairhust, 2004, for a potentially useful way to quantify a feature of cumulative records). All in all, however, picturing moment-to-moment behavior, as in Ferster and Skinner (1957), never led to serious quantification the way aggregating behavior subsequently did for molar analyses.

Another limitation of picturing behavioral variability in cumulative records is that it, in conjunction with the method of creating behavioral variability by hand shaping, may lead to inaccurate ideas about the organizing effect of reinforcement. For example, hand shaping is often complex and highly idiosyncratic, and it would be hard for a cumulative record to capture what a video shows about pigeon ping pong. And, at very short durations, shaping can establish patterns too short to be seen in conventional cumulative records.

Thus, shaping and cumulative records are two key methods for dealing with moment-to-moment behavior, but what they show about the effects of reinforcement and of behavioral variability, have not yet been unified. Unified analyses summarized below provide show how progress might be made toward this and other forms of unification.

It might seem as though about the last thing one would say about qualitative, useful, hand-shaped moment-to-moment behavioral variability, is that it is random in the sense it is temporally uncontrolled. As we will see next, however, assumptions intrinsic to molar theory, about response probability and aggregation, have often led free-operant behavior to be viewed in exactly those terms.

Molar Analyses of Behavior

Three components of molar analyses contribute to molar behavioral variability; 1) behaviors that occur at different times, possibly different places, and possibly produced by different individuals engaged in different moment-to-moment activities, are aggregated into one category of activities or responses that are ostensibly identical in the sense they have, or are assumed to have, the same reinforcing consequences, 2) the aggregated activities are assumed to be emitted randomly in time, i.e., are emitted probabilistically with the same probability over some period of time, and 3) a statistic is chosen and computed to describe the aggregate. All three components are implied by Skinner's assumption (1950) that the empirical average rate of a free operant estimates theoretical response probability. More recent time-allocation methods imply the same three components. Let us consider each of these components and describe how they affect what *molar behavioral variability* means and what it implies about what reinforcement does to behavior.

Aggregation

Why do molar analyses aggregate different occasions of presumably the same response or activity into a single category? I submit it is because Skinner argued that the average rate of a free operant estimates response probability and the classical relative frequency interpretation of probability (Feller, 1950) is in terms of relative frequencies of events over extended periods of time, technically, an infinite number of *trials* but in our case of probabilistic behavior, we assume an infinite duration of time. For example, suppose a response can occur once each second, a response occurs in 470 of 600 sec of observed behavior, so the relative frequency of responses per sec is 470/600, which is a rough empirical estimate of the theoretical probability of a response per sec. (A similar relative frequency can be computed in the case of time allocation, where the average rates of discrete switching responses estimate the likelihoods that an organism will be engaged in one activity or another.)

This kind of aggregation is so familiar that some of the assumptions required for it to be meaningful are sometimes forgotten. Indeed, it can be forgotten that there are any theoretical assumptions at all: aggregating responses is commonly viewed to be an empirical process. Each aggregated response is theoretically required, however, to be functionally identical to any other, except that it occurs at a different random time. If aggregated responses were otherwise different, the aggregate would consist of apples and oranges, as it were, and their average would fail to reflect meaningful relative frequencies of either apples or oranges. Skinner handled this problem by the concept of the generic nature of an operant response, i.e., responses in a specified category were assumed to be functionally identical (Catania, 2011) so that an experimenter did not need to know the idiosyncratic properties of individual responses. This definition is not without its ambiguity: it is *assumed* that aggregating behaviors that have the same function, i.e., the same reinforcing consequences, form a meaningful class or category. This paper would not have to be written, however, if it were always clear what *the same function* means. The most problematic case, described below, is when it is assumed in molar analyses of free operant behaviour that the shaping effect of reinforcement plays no role. Furthermore, this assumption may also be problematic when the presumed operant class involves complex shaped patterns extended in time, like playing pigeon ping pong. Modern literature on categorization might be useful in clarifying the nature of an operant, but that literature is seldom appealed to in the construction of operant aggregates: Molar aggregates are almost universally constructed without regard to what modern literature suggests makes an aggregate a meaningful category. Disregard of this literature may be a problem for any analysis that involves aggregation, including both molar analyses and the molar component of the unified behavior analysis I describe below.

Choice of a Statistic

Aggregating obtained behaviors generates what may be viewed as a sample of behaviors, and a sample affords an opportunity to compute a statistic, to begin quantifying behavior, and to use descriptive and inferential statistics.

Central tendency. The most commonly chosen statistics are those that describe an aggregate's central tendency. The behaviors in the sample, or the times allocated to them, are counted and then an average is computed, such as an average rate of responding over a period of time, or an average time allocated to the behaviors. Selecting and computing such a statistic to describe an aggregate of behaviors implies that doing so does not obscure important kinds of behavioral variability. From a molecular position, it is interesting that this implication is virtually never evaluated because a systematically overlooked molar assumption can tell us something about how molar behaviorists view their work. What might a molar researcher believe that might lead him to ignore an assumption? Consider the answer in the context of a common example where relative rates of responding are used to describe operant choices. One sees choices described this way most commonly without moment-to-moment results also described. From a molecular view, this is problematic because it assumes response-by-response sequences are irrelevant to understanding their molar aggregate; a molar perspective might lead someone to see the irrelevance not as an assumption, but as an established fact. A molar perspective can further lead to the belief that aggregated behaviors are *on a different level* from moment-to-moment behaviors. An adequate treatment of the *levels* idea deserves a more extended treatment than is possible here, although part of the issue is discussed below. Suffice it to say that from a molar perspective, molecular and molar analyses are generally independent (Baum, 2002a, b; Killeen & Hall, 2010; Rachlin & Laibson, 1997), and from a molecular perspective, they are not (Donahoe, 2013; Shimp, 1966, 2013; Tanno & Silberberg, 2012).

It is well known that aggregates may inaccurately represent the events they collect together. For example, cumulative records often show complex patterns which if aggregated and averaged might not represent much, or any, actual moment-to-moment behaving. An average of aggregated Fixed Interval (FI) scallops may not look like any individual FI scallop. This is, however, but a special case of the more general problem that an aggregate's average does not necessarily accurately describe the elements of the aggregate.

Variability. The idea that average response rate based on aggregated performance estimates response probability encourages analyses of steady-state performances; response rates averaged over changing performance do not provide a meaningful, stable, relative-frequency estimate of a constant response probability. The emphasis on steady-state performance seems to be weakening (Jozefowiez, McDowell, & Staddon, 2010; Marr, 2011), but it still exerts a dominant influence on how variability of an aggregate is interpreted. It seems clear that new theoretical methods will be needed to handle the meaning of variability in dynamic contexts (Marr, 2004). In the meantime, two important questions about variability over time of the average are whether response rate has reached a steady state, and whether variability over time or trials is random (Marr, 2012; Neuringer, 2002, 2012; Nevin, 1979). The former has received considerable attention and there are well known practical solutions to how to identify a steady state. The latter is our next topic.

The Assumption that Behavioral Variability Over Time is Random

The terms *random* and *probability* have appeared above without much discussion beyond the relative-frequency interpretation of probability, even though they have many different meanings (Kruger, Daston, & Heidelberger, 1990; Kruger, Gigerenzer, & Morgan, 1990). What they usually mean in behavior analysis has recently been discussed and critiqued in some detail (Barba, 2012; Marr, 2012; Neuringer, 2012) so there is no need here to provide more than one critical idea about the concept of probability and two analogies to physics that are favored in some justifications of molar analyses. For our purposes, the meaning of, and significance assigned to, random behavior begins with Skinner's (1950) assumption that free-operant responses are special in that they are emitted spontaneously and randomly instead of being elicited or controlled by momentary stimuli. In that sense, he saw free operants as distinctly different from elicited Pavlovian responses and from discrete-trials responding in general. He gave a tiny amount of empirical evidence that free-operant methods gave clearer results than discrete-trials methods, although vast amounts of modern quantitative discrete-trials data, as for example in visual search experiments (Blough, 2012) and in serial response time experiments (Shimp, Froehlich, and Herbranson (2007) appear to make this conclusion untenable.

It cannot be overemphasized that Skinner's suggestion that average rate of a free operant estimates a theoretical response probability *necessarily invokes assumptions* built into the idea of probability, including that responses are emitted spontaneously and are aggregated…spontaneously because otherwise they would be controlled and therefore not random, and aggregated because the concept of probability assumes responses are aggregated over very long durations, so that in practice, the longer the period of time over which responses are aggregated, the better the estimate of probability. As noted above, a molar view must also assume the response probability that average rate is deemed to estimate is constant. Thus, Skinner's suggestion imposes important constraints on the molar, aggregate concept of *behavior extended in time*. As Catania (2011) observed, "…aggregate measures are determined only over extended times" (p. 227), and I hope this paper clarifies how theoretical assumptions shape the molar meaning of *behaviour extended over time*. The molar meaning of the term demands that free-operant responding is random in time, therefore not predictable on a moment-to-moment basis, and intelligible only in terms of the average rate at which responses are emitted over an extended period of time.

Two simplistic analogies can be made to random phenomena in physics. (Full disclosure and apologies to the reader: I have forgotten what little I ever knew about physics, so these analogies are based on a layman's scant knowledge. Fortunately, the analogies have never been articulated in any detail so the following should suffice.) Both analogies are about how lawfulness can emerge in statistical analyses of aggregates of random events. According to early particle physics, lawfulness appeared not in terms of conventional deterministic causal explanations involving momentary events, such as in a Newtonian billiard-ball idea of linear chaining, but in terms of probabilities of events occurring randomly over extended periods of time, and estimated by their long-term relative frequencies of occurrence. Causal explanations involving momentary events were replaced by causal explanations involving events extended over time. For example, a single photon was viewed as spontaneously and randomly emitted; individual photons were said not to have momentary causal explanations. There was no momentary stimulus or event that elicited a photon or caused it to be emitted. (Modern quantum accounts are vastly more complicated than this, Nienhuis, 1988.) Instead, it was assumed that there is simply a probability that at any time a photon will be emitted, and the average rate of emission of photons over a long period of time estimated that probability. (For present purposes, there is no need to distinguish between discrete and continuous variables, Hogg & Craig, 1970).

The simplistic analogy to free-operant behavior is straightforward. Free operants are *free* to be spontaneously emitted very much as photons were assumed to be spontaneously emitted: they occur at different random times that are not explained in terms of momentary causal processes, and the average of the distribution of times between responses tells us nothing except the constant probability that a response will be spontaneously emitted. Average response rate can be experimentally manipulated, but momentary emissions cannot be. For present purposes it is vital to emphasize that from this molar view, average response rate is assumed to be unaffected by processes like shaping that reorganize how responses occur in time but do not necessarily affect overall average response rate or response strength. Molar free-operant results are assumed to involve random emission of free-operant responses that are independent of moment-to-moment behavioral processes, just as photon emission or particle decay is independent of moment-to-moment physical processes. All of this means that a search for causal explanations of momentary behavior is futile, and that causal explanations have to involve aggregates of behaviors that are emitted over extended periods of time.

These molar assumptions seem to me to be so unlikely to be correct that they merit repeating again. Skinner's suggestion that average response rate estimates response probability implies the necessity of aggregate analyses that hold that moment-to-moment emission of free-operant behavior is random, that sequential behavioral variability is random, and therefore that there is no shaping effect on sequential variability, just as in early particle physics, moment-to-moment emission of a particle was random, sequential particle emission was random, and no process affected the sequential emission of particles. Perhaps most importantly from the perspective of behavioral variability is that shaping effects of reinforcement are required to be absent from molar analyses. The purpose of shaping is, after all, to impose the very kind of controlled molecular, sequential organization incompatible with the presumed random nature of free-operant behavior.

As described so far, the analogy between free-operant behavior and photon emission or particle decay becomes awkward or entirely inapplicable to cases where it is patently obvious that responses at one time are more likely than responses at some other time. Average rates then do not estimate a constant response probability. In practice, local response rates in most reinforcement contingencies do of course vary, even when the average is steady, as cumulative records so vividly show, and molar arguments wisely seem to avoid trying to extend the physics analogy to these common cases. Stochastic theories that assume or predict that response probabilities change over time seldom if ever are based on Skinner's molar claim that mean rate of a free operant estimates response strength (e.g., Shimp, 1992; Shimp, Childers, & Hightower, 1990).

A second analogy between molar analyses and physics has a similar implication for the relation between moment-to-moment behavior and aggregated behavior. This analogy is to thermodynamics (Rachlin & Laibson, 1997). Random emission of individual operants is said to be analogous to random encounters among atoms or molecules, and randomness at the individual level can produce statistical laws on an aggregate level, as is described in the following;

> …the four laws of thermodynamics and the statistical relation, are very general statements which are completely *macro*scopic in content. They make *no* explicit references whatever to the atoms composing the systems under consideration. They are, therefore, completely independent of any detailed microscopic models which might be assumed about the atoms or molecules in the systems. These statements thus have the virtue of very great generality and can be used even in the absence of any knowledge about the atomic constitution of the systems of interest. (Reif, 1967, p. 285)

In short, it is not necessary to know the microscopic sequences of individual random events to discover macroscopic statistical laws. With *operants* replacing *random events*, and *molar* replacing *macroscopic* statistical laws, this sounds very much like Skinner's probabilistic formulation of the free operant, where it is not necessary to know individually emitted operants in order to understand molar, aggregate laws in terms of average rates.

These analogies between aggregated response rate, on the one hand, and phenomena in particle physics and thermodynamics, on the other hand, hugely simplify the development of quantitative theory IF they are correct. However, they are seldom explicitly articulated or systematically evaluated. Killeen and his colleagues, for example, in their study of possible molecular components of response rate, avoided examining the effects of reinforcement contingencies that involve explicit shaping of the organization of responses in time (Killeen & Hall, 2010; Killeen, Hall, Reilly, & Kettle, 2002). Jensen et al. (2012) appear to have adopted the analogy in their analysis of *truly random* responding, as did Nevin (1979) in his analysis of choice behavior. An original version of the matching law (Herrnstein, 1970) was virtually a defining exemplar of a molar analysis that adopts this analogy; it took aggregate response rate as fundamental and, like the classical version of particle physics, assumed times between response emissions, like times between individual photon emissions, were random and did not reflect causal mechanisms. Average response rate was not derived from un-aggregated, moment-to-moment processes, and instead it formed the quantitative basis for directly expressing basic laws. The principal effect of reinforcement was to strengthen behavior, and to understand the strengthening effect, it was not necessary to understand shaping. Shaping effects of reinforcement had to be excluded.

Skinner never showed how to unify or conceptually integrate shaping, or moment-to-moment analyses in general, on the one hand, and molar free-operant analyses of response strength, on the other hand. So far as I know, he never acknowledged the problem. Molar analyses still confront this dilemma of how to integrate moment-to-moment behaving controlled by shaping and aggregations of moment-to-moment behaving assumed to be random and uncontrolled. The problem facing molar analyses is to unify how shaping organizes and controls behavior in time and how strengthening changes how much of some activity there is averaged over some period of time while leaving temporal organization random and uncontrolled. The former is in terms of individual moment-to-moment behaving and the latter is in terms of aggregate behavior. Consider the following idea of the difference between moment-to-moment behavior and aggregated behavior: "Concurrent schedules of reinforcement induce highly variable allocations of response at the micro level while molar distributions can be described by a power-function relation between overall choices and obtained reinforcers" (Neuringer, 2012). This kind of statement is a recommendation of a molar aggregate analysis over a molecular analysis and sounds to me very much like the two analogies to physics. This contrast is, however, between two very different kinds of behavioral variability, variability in average response rate maintained by free-operant contingencies that assume behavior is meaningful only in the aggregate, and moment-to-moment variability in a behavior stream. In short, the quote is compatible with the molar idea that there is no causal explanation of moment-to-moment behavioral variability that is *truly random* (see Jensen et al., 2012), and that free-operant responding, being random from moment to moment, is generally free of the effects of shaping.

It is impressively bold to assume that free-operant behavior is randomly emitted, but it may be even bolder to assume further, as Skinner did, that free-operant behavior that is assumed to be emitted randomly from moment to moment can be said to be under the control of its consequences. This can mean only that random behavior when viewed as an aggregate is directly under the control of its consequences, because the claim that average rate estimates response probability implies there is no moment-to-moment shaping of individual patterns. To repeat, for aggregated free-operant behavior, such as average rate, to be directly under the control of its consequences, moment-to-moment behavior cannot be under the control of *its* consequences and has to be emitted independently of momentary stimuli, independently of times since previous behaviors and events, and has to have sequential properties that can described only in terms of random emission. Thus, if it is desired to develop a general theory of behaviour, a unified stochastic theory that assumes both shaping and strengthening effects of reinforcement operate continuously and simultaneously in time (e.g., Shimp, Childers, & Hightower, 1990), Skinner's molar assumption that mean rate estimates response probability is not useful or applicable.

The importance of these ideas and assumptions can scarcely be exaggerated. They seemingly make defining the Law of Effect a manageable quantitative enterprise that may not be so complex that it requires simulation, and their broader relevance extends far beyond the confines of scientific analyses of behavior because we often need to know what behavior is purposeful in order to determine what behavior an individual is responsible for. In plain English, we need to know when behavior is caused and when it is random. We need to know when it is environmentally, experimentally, or culturally controlled and when it is *free*. Given this general importance of the assumption of random behavior to molar analyses, several kinds of experiments have been conducted to test the assumption's validity.

One way to test this assumption has been to examine distributions of interresponse times (IRTs), likelihoods of response bursts, bouts of responding, pulses, or of successive choices and other sequential statistics as a function of changeover delay or intertrial interval to see if they correspond to a constant probability of random behavior. In general, they do not, and instead show what seem likely to be effects of shaping (Anger, 1954/1973, 1956; Blough, 1966; Jones & Moore, 1999; Nevin, 1979; Shimp, 1967; Shull, Grimes, & Bennett, 2004; Todorov, Souza, & Bori, 1993).

Can Random Behavior be Shaped?

A second way to test the molar view that free-operant behavior can be randomly emitted is to try to shape a random behavior stream corresponding to that produced by a random process with a constant probability of emitting a response. Can the sequential properties of random behavior be shaped? One might think at first glance that this is a very odd thing to try to do since the purpose of shaping is to control continuous behavioral variability and random behavior is assumed to be spontaneously emitted with no continuous control. Nevertheless, whether random behavior *can* be shaped might shed some light at least on whether it is ever possible for behavior to be random.

Many free-operant contingencies prevent determining whether behavior is emitted randomly. The contingencies most often called *basic* do not experimentally control the temporal patterns of behaviors preceding reinforcers and are not suitable for determining whether behavior is emitted randomly or reflects the effects of shaping (Anger, 1954/1973, 1956; Shimp, 1973). Moment-to-moment theories that maintain that shaping effects are pervasive and that free-operant behavior is therefore not random from moment to moment, have to modify conventional free-operant contingencies to discriminate between molecular shaping and molar strengthening effects (Anger, 1954/1973, 1956; Shimp, 1973; Tanno & Sakagami, 2008; Tanno & Silberberg, 2012). This can leave successful moment-to-moment demonstrations of shaping open to the criticism from a molar view that shaping might occur in unusual contingencies designed to permit discriminating between strengthening and shaping, but not in the basic, classic contingencies. Molar theories, conversely, are left open to the criticisms from a molecular view that they do not explain why shaping so often has an effect when it is explicitly arranged, why that fact does not imply that shaping has an implicit effect even when it is not explicitly arranged, and especially, why molar theory relies to such a large extent on contingencies where it is difficult to obtain conclusive findings one way or the other (Rachlin & Laibson, 1997), and why straight-line cumulative records that can obscure effects of shaping on short-term moment-to-moment behaving are so often used as the standard for defining a constant probability of a free operant response.

To my knowledge, no one has succeeded in shaping free-operant behavior that occurs with a constant probability over time, and although various efforts have come close in terms of some standards, perturbations attributable to seemingly minor aspects of contingencies seem difficult to avoid and to my knowledge none of these efforts have used extensive criteria for what constitutes *randomness* (Blough, 1966; Shimp, 1967). I believe it would be a serious mistake to accept that random free-operant behavior has been achieved or that the behavioral processes that have created approximations to randomness to date are those required to establish the kind of randomness involved in analogies to physics.

Thus, there is little evidence to support the assumption that free-operant behavior is randomly emitted from moment to moment, either in conditions where shaping is not explicitly arranged or when shaping is designed to create random behavior. If random responding defies efforts to establish it, then it is hard to see how Skinner's assumption that the average rate of a free operant estimates response probability could be generally correct. It is therefore also hard to see how average rate of a free operant could estimate response strength. In my opinion, the answer to the title of this section of this paper seems likely to be, no. From the perspective of unified analyses I discuss below, that turns out to be a good thing, not a bad thing, because it permits integrating moment-to-moment shaping with aggregated behavior.

What is Behavioral Variability in Molar Analyses?

So then, given this description of what a molar analysis is, what is molar behavioral variability and what does it tell us about the nature of reinforcement? From the perspective of molar theory and molar analogies to physics, in one important sense, there is no variability from a molar perspective! If a contingency produces a straight-line cumulative record and the contingency remains unchanged over a period of time, there is no moment-to-moment behavioral variability in the likelihood of emitting a free operant over time, the molar average rate based on an aggregate of responses remains constant, and hence there is no molar variability. Furthermore, individual responses are spontaneously emitted and there is no experimentally controlled sequential structure, and hence there is no molecular variability that is not random. In these ways, there is neither molecular nor molar variability! A Geiger Counter and a radioactive object provide a molar analogy. If the counter is held at a fixed distance from the object, particles are emitted with a constant probability without any controlled moment-to-moment variations in emission probability. A cumulative record of particles emitted would show an analogous straight line, there would be no non-random variability, and for as long as the distance between object and Geiger Counter were held constant, there would be no aggregate, molar variability, either. From the perspective of the analogy to physics, the variability molar analyses account for is the variability produced by moving a Geiger Counter closer to or farther away from a radioactive source. Once the distance is held constant, from a molar perspective, there is no longer any non-random, experimentally controlled, variability. Response strength becomes equivalent by this analogy to how temporally *close* an organism is to a reinforcing source. The *closer* an organism is to a reinforcing source, the more often on the average it randomly responds. This may be closely related to what Baum (1973) meant by the correlation based law of effect. Accounts of the effects of delays between responding and reinforcement also seem related to this analogy (Catania, 1971).

Furthermore, from the implications of Skinner's idea that average response rate estimates response probability and hence response strength, there is not even any behavioral variability within the aggregated category of responses itself because they are all assumed to belong to the same category of functionally equivalent responses, and in that sense are identical. Otherwise, the aggregation process would create a hodge podge of different responses. Even when one activity occurs for a longer time than another, if they are aggregated, they are assumed to be of the same kind, only differing in how long each is, so that they can all be collected together and their respective times added up to produce a meaningful accumulated aggregate, molar time.

I hope that the reader feels at this point that something is seriously amiss because on the one hand there is no molar variability to account for, but on the other hand molar analyses routinely conclude that they account for most of the obtained variability. How can that happen? In a nutshell, it is that what *behavioral variability means* depends on whether the context is molecular or molar; *behavioral variability* cannot be interpreted outside the context of the kind of analysis in terms of which it is defined and used. From a molecular perspective involving shaping, a molar analysis assumes that there is no non-random variability in the temporal structure of successive free-operant responses, yet from a molar perspective that molecular attribution of an assumption to molar analyses is incorrect because the attribution is about an assumption, and it is not an assumption, it is a proven fact. One can only hope that such a dispute would initiate a profitable discussion about molar variability; does it require the assumptions I have attributed to it? If it does, then molar claims that reinforcement has direct and general effects on aggregates that do not depend in any way on shaping effects need clearer justification.

I hope this discussion of molar variability clarifies what Skinner (1976) meant when he advocated for molecular analyses as being the most useful, because it is not clear how it helps a batter's performance if it is controlled by averages of aggregates when he faces a 105 mph fastball, or how it helps a pianist's performance of a difficult passage if it is controlled only by average allocation of time to practicing the piano over the last month, or how averages of aggregates helps a pigeon playing pigeon ping pong facing a ball rolling toward it. Molecular variability is the variability in the behavior streams of the batter, pianist, and pigeon. Molar variability might involve the batting average of players facing fastballs as a function of the speed of the fastballs, averaged over a season, or it might involve the average percentage of time spent playing the piano per day, over a month's time, but for the pigeon playing pigeon ping pong, I will leave it to the molar enthusiast to develop what a molar analysis might be, or to explain why such moment-to-moment social interactions between two organisms is not an important problem.

Lastly, from a molecular perspective, there is still one more kind of molar variability. Recall that the kinds of molar variability we have seen so far requires an aggregate that consists of members of an operant class that are elements of a behavior stream. These may be like key pecks or button presses, activities like standing on one side of a chamber, or any members of a reinforced class. This kind of aggregate has a clear relation to the moment-to-moment behavioral variability from which it was derived, as in Herrnstein (1970) and Baum and Rachlin (1969). This aggregate, by the molar analogy, is like the aggregate of particles emitted by a radioactive source over a period of time. This aggregate is based on the assumption that there is a moment-to-moment stream of emitted particles that generates the molar aggregate in the first place.

A second kind of aggregate is disconnected from any analogous stream of particles or from any stream of behavioural activities. In this case, there is no guarantee that the activities in the aggregate belong to the same operant class. In this second case it is not clear how to determine whether the behaviors aggregated together should actually be aggregated. This kind of aggregate is sufficiently different from the kind described previously that it deserves a different name. The term *extreme molar analysis* captures the extent to which it separates a molar aggregate from its molecular origins. Consider the following examples. Imagine archival research is used to collect data from highly disparate contexts, possibly including radically different contingencies that are known to produce different behavioral variability, and these diverse data are grouped together in the same operant class for statistical analysis, as has been the case with national crime data (Wilson & Herrnstein, 1985), and with economic data aggregated over diverse samples (Rachlin & Laibson, 1997). Behavioral methods to reduce or prevent *crime* behavior may be more successful if they consider how complex aggregates of different kinds of criminal behaviour derive from different kinds of behavior streams and different contingencies. Similarily, behavioral economics has great potential to impact public policy (Hackenberg, 2012), but it may facilitate progress to make clear what the connections are between public policy aggregates and the molecular origins of the elements of these aggregates.

I have just described potential problems with extreme molar analyses. Looked at from a molar perspective that sometimes equates moment-to-moment behaving with reflexology (Rachlin & Laibson, 1997), there is no need for concern such as mine that the connection between a behavior stream and a molar aggregate is unknown, or perhaps unknowable; it can in fact seem like a positive virtue. For example, consider the following: "The matching law predicts not … whether the pitcher will throw a fastball, curve, slider, or changeup at this moment but the proportion of each type of pitch he will throw to a given batter" (Rachlin & Laibson, 1997, p. 1). This position is entirely compatible with Herrnstein's definition of the difference between molar and molecular views: "…molecular processes, …processes that control the emission of single, or just a few, responses rather than hundreds or thousands of them." (Rachlin & Laibson, 1997, p. 68).

From a molecular perspective, this way of equating moment-to-moment behaving with either reflexology or with tiny sample sizes is a form of throwing out the baby (moment-to-moment behaving) with the bath water (reflexology or small samples). Reflexology and shaping have nothing in common, and Ferster and Skinner (1957) pictured many, many moment-to-moment patterns greatly extended in time as part of their molecular analysis of behaving. Theirs was not a study in reflexology. Similarly, small sample sizes of responses from an individual and moment-to-moment responding should not be confused with each other. They are generally unrelated.

Unified Analyses of Behavior

Molecular and molar analyses are still works in progress, and so is the notion of a unified analysis. I hope to show, however, that even in a preliminary version, unified analyses clarify what the other two kinds are in the sense that they clarify how molecular and molar analyses are usually mutually informing rather than competing. So, what is a unified analysis? A unified, or at least a partially unified, analysis involves three components; (a) automated shaping of quantitatively defined moment-to-moment temporal patterns, (b) aggregates of these shaped patterns, and (c) statistics to describe the aggregates and their relations to reinforcement contingencies and parameters. To anticipate, a unified analysis addresses, if not completely solves, the dilemma described above inherent in molar analyses to the effect that average rate is supposed to estimate response probability, implying moment-to-moment responses occur only with a constant probability over time, yet aggregates of such random behavior are supposed to be lawful. It does so by acknowledging a role for shaping in creating the aggregated responses: they resemble target responses having experimentally determined quantitative properties. We will find that several kinds of behavioral variability result, a quantitative version of moment-to-moment variability, and the kinds of molar behavioral variability that are derived from behavior streams. What I call extreme molar analyses that have unknown behavior-stream origins cannot be part of a unified analysis that relates moment-to-moment behaviors and aggregates of those behaviors.

Automated Shaping; Quantitative Precision and Naturalistic Realism

Automated shaping creates new, temporally-extended response patterns with a degree of quantitative precision impossible to achieve with hand shaping (e.g., Anger, 1956; Blough, 1966; Galbicka & Platt, 1986; Hawkes & Shimp, 1975, 1998; Machado, 1989; Platt, 1973; Shimp, 1973; Shimp, Fremouw, Ingebritsen, & Long, 1994; Terrace, 2012).

This precision has come, it must be said, at the expense of naturalistic realism. For example, IRTs have served admirably as one-dimensional, quantitatively-specified response patterns. Marr (2012) has reminded us, however, that realistic shaped patterns have a dimensional complexity and a continuous fluidity that IRTs do not display. Thus, the technology for automated shaping of IRTs does not readily generalize to shaping pigeon ping pong, training bears to ride motorcycles, or teaching a child to play the guitar.

There may be a surprising way to overcome these limits. The computer software and hardware now available for video games might permit shaping of complex behaviors in simulated naturalistic environments (Shimp, 2013). The software and instrumentation for driving simulators, pilot training simulators, teachable robots, and so on, combined with appropriate unified shaping contingencies, might suffice to lend automated shaping new power. We can already say that automated shaping can establish behavioral variability with specified quantitative properties, but only the future will clarify how automated shaping can create more complex and naturalistic forms of multidimensional behavioral variability.

Unified Aggregates

Once automated shaping has created new temporal response patterns having quantitatively specified features, it becomes possible to create conventional molar aggregates of those complex responses, in a manner described previously. Either counts of patterns or cumulated times allocated to them can easily be computed (Anger, 1954/1973;1956; Shimp, 1968, 1973, 1974; Staddon, 1968).

Unified Empirical Functions

Once aggregates of quantitatively shaped responses are formed and statistics are computed to describe the aggregates, it becomes possible to plot conventional molar functions that show how the statistics depend on response durations and reinforcement parameters and contingencies (Anger, 1956; Galbicka & Platt, 1986; Shimp, 1968, 1973, 1974, 2013; Shimp, Herbranson, & Fremouw, 2012; Staddon, 1968). The literature on these results has been reviewed elsewhere (Shimp, 2013), so it suffices here just to mention that it strongly suggests that molar function depends on molecular structure (Doughty & Lattal, 2001; Shimp et al., 1994). An emerging possibility of great possible significance is that these results may require new theoretical methods and concepts to handle the new quantitative control of complex and dynamic patterns of behavior extended in time (McDowell, 2013; Shimp, 1992, 2013; Tanno & Silberberg, 2012). I do not know of any theory that is not a simulation theory that handles results such as these.

What is Behavioral Variability in Unified Analyses?

Behavioral variability in unified analyses has elements of both molecular and molar variability. First, it involves moment-to-moment behavioral variability that resembles the variability within target patterns, as in different instances of pigeon ping pong, but now, unlike the case with pigeon ping pong, the variability has quantitative structure resembling quantitatively specified targets, like IRTs, interchangeover times, and similar patterns (Blough, 1966; Galbicka, 1994; Hawkes & Shimp, 1998; Platt, 1973; Shimp, 1967). Second, it also has molar behavioral variability in the form of statistics summarizing how aggregates of shaped patterns change as a function of contingencies and time, such as over sessions. Thus, it has two forms of behavior extended in time, the individual, shaped, moment-to-moment response patterns, and the aggregations of these patterns. Behavior varies over both. If an analysis neglects either one, it deals only with shaping or strengthening effects of reinforcement, not both. A unified analysis includes both.

An approximation to continuous behavioral variability, as within a game of pigeon ping pong, can be achieved in discrete-trials tasks involving individual sequential patterns (Barba, 2012; Blough, 2012; Fountain, Rowan, Muller, Kundey, Pickens, & Doyle, 2012; Jensen et al., 2012; Jones & Moore, 1999; Marr, 2012; Neuringer, 2012; Shimp, Froehlich, & Herbranson, 2007; Terrace, 2012). Stated otherwise, results of discrete-trials methods may involve both shaping and strengthening effects. This is a considerable advantage if one is trying to develop a general theory of behavior rather than separate theories of strengthening or shaping. On the basis of the breadth and precision of discrete-trials results now available, it would seem Skinner was premature to judge that discrete-trials methods produce data scientifically inferior to that produced by free-operant methods.

Conclusions and Implications

Behavior and *reinforcement* have different meanings in molecular and molar analyses. A molecular analysis is based on the temporal organization of moment-to-moment behaving and how it is controlled by shaping effects of reinforcement. A molar analysis is based on aggregations of behaviors and how they are controlled by the strengthening effect of reinforcement. *Behavioral variability* therefore has different meanings in different kinds of behavioral analyses. Each kind has its own advantages and disadvantages and the practical utility and theoretical success or failure of an analysis depends on an experimenter's purpose and the context in which it is used. Molecular analyses produce new, complex patterns but do not permit quantification and shed little light on the strengthening effect of reinforcement, molar analyses attempt to quantify strengthening effects but confound shaping and strengthening effects, and unified analyses show how shaping and strengthening interact but have not yet been generalized to many complex every day patterns of behavior. It would be satisfying for the reader to discover at the end of this paper that there is a kind of analysis that transcends the various kinds of limitations I have described, but such is unhappily not the case.

Kinds of Behavioral Variability and Practical Goals

Practical goals often determine the kinds of behavioral analyses that are suitable, which in turn determines what kinds of behavioral variability are generated. Sometimes this is obvious. Skinner's goal of establishing pigeon ping pong required him to use shaping, and all Skinner had to do in the form of data analysis for pigeon ping pong was to show pigeons playing the game. However, matching behavioral variability to practical goal is usually more complex. To exemplify this, imagine two different experiments, with each having the same goal but with each producing its own kind of behavioral variability. The goal is to improve healthful decision making by an individual at the grocery store.

Consider a molar analysis. We might observe how the individual allocates times separately to the candy aisle, to the produce section, to the meat counter, and so on, over several shopping excursions, and then we might aggregate these times and determine total times allocated to buying candy bars, to vegetables, and so on. A pie chart could show percentages of times allocated to different regions of the store. Suppose the pie chart showed an undesirably greater percentage of time allocated to the candy aisle than to the produce section. The experimenter might accordingly arrange contingencies that randomly delivered more reinforcers while the individual is standing in the produce section than while he is in the candy aisle. Sometimes these reinforcers might happen to be delivered while the individual had just entered the produce section, sometimes while holding an eggplant and looking at it, and sometimes while approaching the exit to the produce section, and so on. This method would ignore the actual moment-to-moment behaving while in the produce section; the sequential behaviors would be ignored and the aggregate behaviors would be paramount on the grounds that moment-to-moment behaviors are not causally determined and are only random, as I explained above. This experiment might improve the shopper's allocation of time to the produce section, or it might not. It would not show anything about any particular behavior stream, and an experimenter might not know the outcome until after the intervention was completed and data were analyzed, and of course, the variability in the data might obscure the effect to the point where it was not obvious. Null results might encourage the experimenter to collect data over more time or from more individuals to obtain a bigger aggregate in order to collect ever larger samples to better conform to Herrnstein's definition of molar behavior as involving large samples. The statistical analysis of variability might become more complex and the outcome could again be either the method worked, it did not work, or it was impossible to tell. The analysis might become ever more similar to the kinds of demands put on behavior analysts over half a century ago to investigate aggregate behaviors that encouraged them to initiate the *Journal of the Experimental Analysis of Behavior* aimed at the behaving of individuals. The reader can imagine further hypothetical molar experiments and their results. The point of this example is not to prove anything about the likely experimental outcome but to emphasize what a molar analysis might look like. The example illustrates a molar analysis because it involves aggregated behaviors and reinforcers delivered in a manner that does not require knowing the temporal organization of behaviors that precede reinforcers. There is no explicit shaping of moment-to-moment behavior. Any pie chart showing percentages of times allocated to different activities would confound shaping and strengthening effects.

Alternatively, consider a molecular analysis. This method requires more control of moment-to-moment behavior so let us imagine a simulated grocery shopping task. Real-time visual simulations like those for pilot or driver training could be adapted to show the interior of a grocery store as an individual moves about in it. The simulator would continuously show the grocery store from his viewpoint. A simulation might begin by placing the individual at the entrance to the grocery store, perhaps one that looks like the one he usually shops in. If he moves toward the produce section, he would be given a reinforcer. If he moves toward the candy aisle he would not be. If he actually enters the produce section, he would be given another reinforcer. He would continue to receive reinforcers to the degree to which his behavior resembled the target behavior of moving to areas with healthful food, picking them up, and putting them in his basket, moving to the checkout clerk, paying, and leaving the store. All this is conventional shaping using the method of successive approximations in a simulated environment. Again, this hypothetical example is not designed to prove which kind of analysis is *best*, just to illustrate a moment-to-moment method to shape desired behavior. These examples doubtlessly reveal my own background limited to abstract theory and laboratory work, rather than in improving practical outcomes and in translational research. I hope my mistakes might motivate better informed behavior analysts to improve the method.

In any case, if the method worked reasonably well the outcome would be immediately obvious; an experimenter would simply watch grocery shopping and see that it was wiser than before, just as Skinner simply watched pigeons playing ping pong. There would be no need for aggregating responses, choosing a statistic to describe the aggregate, or to perform any quantitative analysis (although many quantitative possibilities suggest themselves for this example, on behalf of such standards as *evidence based* or *data based* evaluation.) Naturally, there would be the difference between simulated and real shopping, but my example might be a beginning. And if the molar version confounds quantitative effects of shaping and strengthening, the molecular version may be more difficult to quantify at all. I do not advocate that one of the two methods is more likely to succeed than the other, just that they produce results with very different kinds of behavioral variability as a function of different conceptions of what behavior and reinforcement are.

In short, these two hypothetical examples are designed to clarify practical differences between molecular and molar analyses, not to show which is *right* or *better*. It is not too hard to see how each method might be more suitable, depending on the customer, the availability of technical resources, and other details. Perhaps most importantly, the possibility that each would work, in its own way, suggests that there is no reason for one method to claim universal generality at the expense of the other. Indeed, in a third unified version of the examples, both molecular and molar methods could be combined; the chief difference from the molar version described above would be that reinforcers would be administered not randomly in terms of moment-to-moment behavior but administered to shape quantitative patterns, of more or less specific durations of reading nutrition advice in the produce section, in the candy aisle, and so forth.

Kinds of Behavioral Variability and Theoretical Goals

There is literature the purpose of which seems to be to show that either moment-to-moment theories or aggregate theories are better (e.g., Baum, 2002b; Donahoe, 2013, in press; Nevin, 1979; Shimp, 1966, 1969, 1992; Rachlin & Laibson, 1997; Williams, 1988, 1990). Competitive comparisons of molecular and molar theories have sometimes involved comparing the degree to which one kind of theory describes its kind of behavioral variability to the degree to which the other kind of theory describes its kind of behavioral variability. This is an odd thing to have done. The result seems to have been that each side believed its side was better rather than simply believing that one or the other analysis is better, depending on the context. Consider a consumer choice example. Suppose a driver believes that a Toyota Prius is a better car than a Porsche sports car because it gets better gas mileage and is more socially responsible and egalitarian, while another driver says the Porsche is the better car because it has better brakes, is more nimble and can better avoid an accident and can better prevent the attendant pain and cost. In neither case does a driver explain why social responsibility in the form of gas mileage or active safety technology is the more important criterion, leaving a prospective car purchaser to continue to wonder which is the *better* car, and perhaps to wonder further if something is wrong with the idea of *better*. I am not suggesting that there are no criteria for evaluating theory, just that evaluating a theory by using the evaluative standards of a different kind of theory can be problematic.

Suppose that a theory predicts an individual's moment-to-moment or trial-to-trial behavioral variability and generates a behavior stream from which one can compute an individual's aggregate behavior over time. This type of theory usually uses computer simulation methodology (Catania, 2005; McDowell, 2013; Shimp, 1969, 1992; Shimp, Herbranson, & Fremouw, 2012; Tanno & Silberberg, 2012). Suppose another theory is a molar theory based on the idea that aggregates lead to more lawful results than moment-to-moment behavior (Baum, 2002b; Jensen et al., 2012; Neuringer, 2012; Rachlin & Laibson, 1997). It accordingly predicts average behavioral variability of aggregates and only aggregate variability. In each case, a theorist compares the theories in terms of the standards suitable for his own preferred theory, perhaps a theory he himself or his dissertation advisor developed. Consider the following example in a review of the previous draft of this paper. One reviewer, an extremely eminent researcher and scholar whom I greatly admire personally and professionally, wrote that "It seems to me you are ignoring that the molar versus molecular analyses of choices under concurrent schedules have one and the same goal: to predict choice allocations. The question is whether that is better done by, e.g., overall relative rates of reinforcement or by response-by-response, momentary relationships between choices and reinforcers." To me, this quote is a good practical definition of a molar view; the general goal is to predict *choice allocations*. (I assume from the reviewer's other comments that by *choice allocations* he meant aggregated allocations). From my more molecular view, I ask, why is that the goal? Why is an individual moment-to-moment behavior stream like Lincoln's Gettysburg Address not as important as the percentage of his time Lincoln allocated to giving speeches over a month's time? The reviewer sees it as an established fact that the goal is to understand an aggregate, like the daily times Lincoln allocated to giving speeches aggregated over a month. I personally am at least as interested in the one unique speech, in its phrasing, structure, meaning, political and cultural impact, than in how Lincoln on the average allocated his time to speeches. I can see why someone might be interested in the latter, but I am concerned about a belief that there is only one goal and it involves averages of aggregates.

Consider another kind of competiton between kinds of analyses, in which one kind is *pitted against* another, rather like how athletic teams are pitted against each other. In sports, there is supposed to be a common set of rules to determine the winner. But what is the common set of rules when moment-to-moment theories and aggregate theories generate different kinds of predictions? It is not as though there are point totals at the end of such an analysis like point totals at the end of a football game. It is more like one analysis gives a score in terms of a game of football and the other in terms of a game of golf. Let us suppose a molar theory is a very good molar theory and it describes 99.5 percent of *the data*. It is almost universally left unsaid that the theory also describes 0.0 percent of the moment-to-moment data, or rather makes no prediction at all, with no apologies offered because it is designed on purpose not to bother with generating a behavior stream viewed as having little diagnostic value. Let us suppose further that a moment-to-moment theory against which the molar theory is pitted describes only 90% of the molar variability in the same data, but does so by first generating a an appropriate behavior stream from which the aggregate results are derived. Which theory is better? All we have learned is that different kinds of theories describe different kinds of behavioral variability, and those different experimental arrangements, including all the molecular and molar contingencies, and the experimenters' choices of which empirical phenomena should be described, can advantage one kind of theory over another to one degree or another.

If pitting one theory against another is problematic, so too is the use of probes delivered in one context to determine response strengths acquired in another context. This is problematic because from a moment-to-moment perspective, response strength estimated in a probe context does not necessarily have anything to do with response strength in a training context, and vice versa. They may have nothing to do with each other because different moment-to-moment contingencies involve implicit shaping in different ways. How theories are evaluated changes over time, and much time has passed since it was believed that it was possible to subject a theory to a *critical test* (Benham & Shimp, 2004), and it is not clear in testing kinds of analyses how one could design a context so that neither a moment-to-moment analysis nor an aggregate analysis was advantaged over the other. In the grocery shopping example described above, for example, it surely would not prove either molecular or molar methods were generally better if the results seemed to favor one over the other. They are simply different ways of controlling behavioral variability defined either with respect to moment-to-moment shopping or aggregated shopping trips.

Perhaps a useful alternative approach would be to decide what kind of behavioral variability it is desired to describe or control, and then to proceed accordingly to develop a method and theory to do just that. We might use hand shaping to create and control complex individual behaviors extended in time and perhaps not even attempt to quantitatively describe the variability; we would just look at the experimental organism and see if its behavior sufficiently resembles the target we chose. Or, we might use either molar or unified theories for aggregate variability when the molar data derive from behavior streams, and unified theories for variability obtained when both moment-to-moment and molar contingencies are manipulated. From this point of view, it seems pointless to try to find an experimental context where any particular kind of theory is best. If one theory works better than another in an experimental setting, a useful task might be not to cheer if that theory is one's own, but to try to determine what it is about the empirical task that makes that particular kind of theory more suitable for that task.

Unification of, Versus Competition Among, Different Kinds of Behavioral Analyses

By now the reader will have correctly concluded that in my opinion it will be difficult to find contexts where comparisons among the three kinds of analyses are *fair* (Shimp, 2004); what does it mean to compare a theory that simulates behavior streams with a theory that describes averages of aggregates? Perhaps the unified category of analysis offers a beginning toward a way out of attempting futile tests. It integrates molecular-shaping and molar-strengthening effects of reinforcement into one analysis that, admittedly, has advantages and disadvantages of both molecular and molar analyses. It is not perfect. However, if we grant that if we want to construct a general theory we need to understand *both* sequential shaping and average strengthening effects, including how they interact, we could try to make progress by working to show *how they relate to each other* rather than by working toward what appears to be a doomed goal of establishing either alone as a general analysis. In other words, molecular and molar analyses can be seen as not necessarily in competition but instead as each informing our understanding of the other and as contributing to a unified analysis.

Molecular and unified analyses on the one hand are effective in organizing the moment-to-moment patterning of individual behavior, and in explaining aggregates that derive from such behavior, respectively, but they are not yet able to explain how national crime statistics change from year to year. Overcoming that limitation will surely require better understanding of the relation between moment-to-moment behavior and statistics that describe complex aggregates of diverse behaviors. An extreme molar analysis on the other hand might be able to handle the latter but not the former. That is, statistics that describe aggregated national economic or crime data seem unlikely to give clear guidance to a child holding a candy bar and trying to decide whether to buy it, or guidance to a wife faced with how to respond to a husband who called to ask for advice on how to guide a child who, he has learned, just stole a candy bar a minute ago. I repeat these and myriad other differences in effectiveness of kinds of analyses do not define a competition for which analysis is *best*. In my mind, they remind us of the inexhaustible diversity of kinds of behavioral variability and therefore the need not to focus prematurely on one kind of analysis. An anonymous reviewer of the first version of this paper remarked that "…molar analyses have the edge in terms of theory, but perhaps that is just because so much more effort has been devoted to theory at that level and because molar theories are regarded by many as somehow more acceptable." I hope this paper clarifies the sense in which molar theories of behavioral variability do and do not *have the edge*. I also hope it clarifies how molar analyses are *somehow more acceptable* or completely unacceptable, depending on whether one wants to understand and control aggregates of behavior, individual behavior in real time, or both.

References

Anger, D. (1954/1973). The effect upon simple animal behavior of different frequencies of reinforcement, Part II: separate control of reinforcement of different IRTs. Reprinted in *Journal of the Experimental Analysis of Behavior*, *20*, 301-312.

Anger, D. (1956). The dependence of interresponse times upon the relative reinforcement of different interresponse times. *Journal of Experimental Psychology*, *52*, 145-161.

Athens, E. S., Vollmer, T. R., & Pipkin, C. C. St. P. (2007). Shaping academic task engagement with percentile schedules. *Journal of Applied Behavior Analysis*, *40*, 475-488.

Barba L. S. (2012). Operant variability: A conceptual analysis, and associated commentary by Neuringer, Machado & Tourneau, Holt, & Marr. *The Behavior Analyst*. *35*, 213-227.

Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior, 20*, 137-153.

Baum, W. M. (2002a). The Harvard pigeon lab under Herrnstein. *Journal of the Experimental Analysis of Behavior*, *77*, 347-355.

Baum, W. M. (2002b). From molecular to molar: A paradigm shift in behavior analysis. *Journal of the Experimental Analysis of Behavior*, *78*, 95-116.

Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior, 12*, 861-874.

Benham, B., & Shimp, C. P. (2004). Falsification in social science method and theory. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 2, pp. 9-14). San Diego, CA: Academic Press.

Blough, D. S. (1966). The reinforcement of least-frequent interresponse times. *Journal of the Experimental Analysis of Behavior*, *5*, 581-591.

Blough, D. S. (2012). Reaction-time explorations of visual perception, attention, and decision in pigeons. In T. R. Zentall & E. A. Wasserman (Eds.), *The Oxford handbook of comparative cognition* (pp. 674-690). Oxford, England: Oxford University Press.

Breland, K., & Breland, M. (1966). *Animal Behavior*. The Macmillan Company.

Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. New York, NY: Wiley.

Bush, R. R., & Estes, W. K. (Eds.). (1959). *Studies in mathematical learning theory*. Stanford, CA: Stanford University Press.

Catania, A. C. (1971). Reinforcement schedules: The role of responses preceding the one that produces the reinforcer. *Journal of the Experimental Analysis of Behavior, 15,* 271-287.

Catania, A. C. (2002). The watershed years of 1958-1962 in the Harvard Pigeon Lab. *Journal of the Experimental Analysis of Behavior*, *77*, 327-345.

Catania, A. C. (2005). The operant reserve: A computer simulation in (accelerated) real time. *Behavioural Processes*, *69*, 257-278.

Catania, A. C. (2011). On Baum's public claim that he has no significant private events. *The Behavior Analyst*, *34*, 227-236.

Catania, A. C., & Reynolds, G. S. (1968). A quantitative analysis of the behavior maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *11*, 327-383.

Donahoe, J. W. (2013). Theory and behavior analysis. *The Behavior Analyst, 36*, 361-371.

Donahoe, J. W. (in press). Origins of the molar-molecular debate. *European Journal of Behavior Analysis*.

Doughty, A. H., & Lattal, K. A. (2001). Resistance to change of operant variation and repetition. *Journal of the Experimental Analysis of Behavior*, *76*, 195-215.

Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, *57*, 94-107.

Estes, W. K. (1959). The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A study of a science, Study 1. Conceptual and systematic* (Vol.2). *General systematic formulations, learning, and special processes*. (pp. 380-491). New York, NY: McGraw-Hill.

Feller, W. (1950). *An introduction to probability theory and its applications*. New York, NY: Wiley.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York, NY: Appleton-Century-Crofts.

Fountain, S. B., Rowan, J. D., Muller, M. D., Kundey, S. M. A., Pickens, L. R. G., & Doyle, K. E. (2012). In T. R. Zentall & E. A. Wasserman (Eds.), *The Oxford handbook of comparative cognition* (pp. 594-614). Oxford, England: Oxford University Press.

Galbicka, G. (1994). Shaping in the 21st century: Moving percentile schedules into applied settings. *Journal of Applied Behavior Analysis*, *27*, 739-760.

Galbicka, G., & Platt, J. R. (1986). Parametric manipulation of interresponse-time contingency independent of reinforcement rate. *Journal of Experimental Psychology: Animal Behavior Processes*, *12*, 371-380.

Galbicka, F., Kautz, M. A., & Jagers, T. (1993). Response acquisition under targeted percentile schedules: A continuing quandary for molar models of operant behavior. *Journal of the Experimental Analysis of Behavior*, *60*, 171-184.

Gallistel, C. R., Balsam, P. D., & Fairhurst, S. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences*, *101(36)*, 13124-13131.

Hackenberg, T. D. (2012). From demand curves to public policy: Introduction to the special issue on behavioral economics. *Journal of the Experimental Analysis of Behavior*, *99*, 1-2.

Hackenberg, T. D. (2013). What has happened to Skinner's empirical epistemology? *The Behavior Analyst, 36,* 277-281.

Hawkes, L., & Shimp, C. P. (1975). Reinforcement of behavioral patterns: Shaping a scallop. *Journal of the Experimental Analysis of Behavior*, *23*, 3-16.

Hawkes, L., & Shimp, C. P. (1998). Linear responses. *Behavioural Processes*, *44*, 19-43.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*, 243-266.

Hogg, R. V., & Craig, A. T. (1970). *Introduction to mathematical statistics* (3rd ed). London, England: Macmillan.

Jensen, G., Miller, C., & Neuringer, A. Truly random operant responding: results and reasons. In T. R. Zentall & E. A. Wasserman (Eds.), *The Oxford handbook of comparative cognition* (pp. 652- 673). Oxford, England: Oxford University Press.

Jones, J. R., & Moore, J. (1999). Some effects of intertrial-interval duration on discrete-trials choice. *Journal of the Experimental Analysis of Behavior*, *71*, 375-393.

Jozefowiez, J., McDowell, J. J., & Staddon, J. E. R. (2010). Editorial: Choice studies in transition. *Journal of the Experimental Analysis of Behavior*, *94*, 159-160.

Killeen, P. R., & Hall, S. S. (2010). The principal components of response strength. *Journal of the Experimental Analysis of Behavior*, *75*, 111-134.

Killeen, P. R., Hall, S. S., Reilly, M. P., & Kettle, L. C. (2002). Molecular analyses of the principal components of response strength. *Journal of the Experimental Analysis of Behavior*, *78*, 127-160.

Koch, S. (1959). *Psychology: A study of a science*. New York, NY: McGraw-Hill.

Kruger, L., Daston, L. J., & Heidelberger, M. (Eds.). (1990). *The probabilistic revolution: Vol. 1: Ideas in history*. Cambridge, MA: MIT Press.

Kruger, L., Gigerenzer, G., & Morgan, M. S. (1990). *The probabilistic revolution: Vol. 2: Ideas in the sciences*. Cambridge, MA: MIT Press.

Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, *52*, 155-166.

Marr, M. J. (2004). Dimension in action and the problem of behavioral units. In J. Burgos & E. Ribes (Eds.), *Theory, basic and applied research, and technological applications in behavioral science: conceptual and methodological issues* (pp. 151-177). Guadalajara, Mexico: Universidad de Guadalajara.

Marr, M. J. (2011). Has radical behaviorism lost its right to privacy? *The Behavior Analyst*, *34*, 213-219.

Marr, M. J. (2012). Operant variability: Some random thoughts. *The Behavior Analyst*, *35*, 237-241.

McDowell, J. (2013). Representations of complexity: How nature appears in our theories. *The Behavior Analyst, 36*, 345-359.

Morse, W. H. (1966). Intermittent reinforcement. In Honig, W. K. (Ed.). *Operant behavior: Areas of research and application* (pp. 52-108). New York, NY: Appleton-Century-Crofts.

Neuringer, A. (2002). Operant variability: evidence, functions, and theory. *Psychonomics Bulletin & Review*, *9*, 672-705.

Neuringer, A. (2012). Reinforcement and induction of variability. *The Behavior Analyst*, *35*, 229-235.

Nevin, J. A. (1979). Overall matching versus momentary maximizing: Nevin (1969) revisited. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*, 300-305.

Nienhuis, G. (1988). Photon emission as a random event. *Journal of Statistical Physics*, *53*, 417-438.

Palmer, D. C. (2013). Some implications of a behavioral analysis of verbal behavior for logic and mathematics. *The Behavior Analyst, 36*, 267-276.

Peterson, G. B. (2004). A day of great illumination: B. F. Skinner's discovery of shaping. *Journal of the Experimental Analysis of Behavior*, *82*, 317-328.

Platt, J. R. (1973). Percentile reinforcement: Paradigms for experimental analysis of response shaping. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 7, pp. 271-296). New York, NY: Academic Press.

Rachlin, H., & Laibson, D. I. (Eds). (1997). Herrnstein, *The matching law: Papers in psychology and economics*. Cambridge, MA: Harvard University Press.

Reif, F. (1967). *Statistical physics*; *Berkeley physics course* (Vol. 5). New York, NY: McGraw-Hill.

Schlinger, H. D., Jr. (2013). A functional analysis of psychological terms redux. *The Behavior Analyst, 36*, 255-266.

Schoenfeld, W. N, & Farmer, J. (1970). Reinforcement schedules and the "Behavior Stream." In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 215-245). New York, NY: Appleton-Century-Crofts.

Shimp, C. P. (1966). Probabilistically reinforced choice behavior in pigeons. *Journal of the Experimental Analysis of Behavior*, *9*, 443-455.

Shimp, C. P. (1967). Reinforcement of least-frequent sequences of choices. *Journal of the Experimental Analysis of Behavior*, *10*, 57-65.

Shimp, C. P. (1968). Magnitude and frequency of reinforcement and frequencies of interresponse times. *Journal of the Experimental Analysis of Behavior, 11,* 525-535.

Shimp, C. P. (1969). Optimal behavior in free-operant experiments. *Psychological Review*, *76*, 97-112.

Shimp, C. P. (1973). Synthetic variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *19*, 311-330.

Shimp, C. P. (1974). Time allocation and response rate. *Journal of the Experimental Analysis of Behavior*, *21,* 491-499.

Shimp, C. P. (1992). Computational behavior dynamics: An alternative description of Nevin (1969). *Journal of the Experimental Analysis of Behavior*, *57*, 289-299.

Shimp, C. P. (2001). Behavior as a social construction. *Behavioural Processes*. (Special Issue on The Longer View: 20th Century Quantitative Analyses of Behavior), *54*, 11-32.

Shimp, C. P. (2004). Scientific peer review: A case study from local and global analyses. *Journal of the Experimental Analysis of Behavior, 82*, 103-116.

Shimp, C. P. (2013). Toward the unification of molecular and molar analyses. *The Behavior Analyst, 36,* 295-312.

Shimp, C. P., Childers, L. J, & Hightower, F. A. (1990). Local patterns in human operant behavior and a behaving model to interrelate animal and human performances. *Journal of Experimental Psychology: Animal Behavior Processes*, *16*, 200-212.

Shimp, C. P., Fremouw, T., Ingebritsen, L. M., & Long, K. A. (1994). Molar function depends on molecular structure of behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*, 96-107.

Shimp, C. P., Froehlich, A. L., & Herbranson, W. T. (2007). Information-processing in pigeons: Incentive as information. *Journal of Comparative Psychology*, *73*, 73-81.

Shimp, C. P., Herbranson, W. T., & Fremouw, T. (2012). From momentary maximizing to serial response times and artificial grammar learning. In T. R. Zentall & E. A. Wasserman (Eds.), *The Oxford handbook of comparative cognition* (pp. 674-690). Oxford, England: Oxford University Press.

Shull, R. L., Grimes, J. A., & Bennett, J. A. (2004). Bouts of responding: The relation between bout rate and the rate of variable-interval reinforcement. *Journal of the Experimental Analysis of Behavior*, *81*, 65-83.

Skinner, B. F. (1948). *Walden Two*. New York: Macmillan.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, *57*, 193-216.

Skinner, B. F. (1959). A case history in scientific method. In S. Koch (Ed.), *Psychology: A study of a science, Study 1. Conceptual and systematic, Vol.2. General systematic formulations, learning, and special processes* (pp. 359-379). New York, NY: McGraw-Hill.

Skinner, B. F. (1976). Farewell, My LOVELY! *Journal of the Experimental Analysis of Behavior*, *25*, 218.

Skinner, B. F. (1979). *Shaping of a behaviorist*. New York, NY: Knopf.

Staddon, J. E. R. (1968). Spaced responding and choice: a preliminary analysis. *Journal of the Experimental Analysis of Behavior*, *11*, 669-682.

Tanno, T., & Sakagami, T. (2008). On the primacy of molecular processes in determining response rates under variable-ratio and variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, *89*, 5-14.

Tanno, T., & Silberberg, A. (2012). The copyist model of response emission. *Psychonomic Bulletin & Review*, *19*, 159-178.

Terrace, H. S. (2012). The comparative psychology of ordinal knowledge. In T. R. Zentall & E. A. Wasserman (Eds.), *The Oxford handbook of comparative cognition* (pp. 615-651). Oxford, England: Oxford University Press.

Thorndike, E. L. (1911). *Animal intelligence*. New York, NY: Macmillan.

Todorov, J. C., Souza, D. G., & Bori, C. M. (1993). Momentary maximizing in concurrent schedules with a minimum interchangeover interval. *Journal of the Experimental Analysis of Behavior*, *60*, 315-435.

Williams, B. A. (1988). Reinforcement, choice, and response strength. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology: Vol. 2. Learning and cognition* (pp. 167-244). New York, NY: Wiley.

Williams, B. A. (1990). Enduring problems for molecular accounts of operant behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, *16,* 213-216.

Wilson, J. Q., & Herrnstein, R. J. (1985). *Crime & Human Nature: the definitive study of the causes of crime*. New York, NY: Simon & Schuster.

Zeiler, M. D. (2006). An architect of the golden years. *Journal of the Experimental Analysis of Behavior*, *86*, 385-391.

Financial Support: No financial support was declared
Conflict of Interest: The author of this paper declared no conflict of interest.

Submitted: *September 1st, 2013*
Resubmitted: *November 27th, 2013*
Accepted: *December 19th , 2013*