

6

On the Proper Scales for Reaction Time

James T. Townsend
Indiana University

Reaction time or response time, as some prefer to call it, has been a key to the study of mental faculties, along with patterns of accuracy, virtually from the beginnings of experimental psychology (e.g., Boring, 1957; Luce, 1986; Townsend & Ashby, 1983; Welford, 1980). However, the question of its status within modern frames of measurement has remained somewhat opaque. I try to convince the reader that the question is important and that the scale properties of reaction time are strong, when used in appropriate ways. However, it is my belief that this cannot occur by way of mathematical proof. Rather, I argue that reaction time can, under the proper circumstances, be used in its physical sense. Alternatively, in other important cases, it can be used in ways that demand formal qualitative procedures to establish its scale properties. First it is necessary to limn in some background on measurement theory. Those readers who are already versed in measurement theory should jump to the next section. Obviously our treatment here cannot be technical, because of space limitations.

MODERN FRAMES OF MEASUREMENT

By "modern frames of measurement", I refer to the line of thought and research in psychology originating with S. S. Stevens (e.g., 1951) and undergoing rigorous treatment and much expansion by the foundational measurement group (e.g., Suppes & Zinnes, 1963; Krantz, Luce, Suppes, & Tversky, 1971; Roberts, 1979; Luce & Narens, 1987; Narens, 1985). We refer to this tradition as the "foundational approach".

That framework is similar in its logical format to Felix Klein's 19th-century

Erlangen program in which various geometries were classified in terms of their invariance properties under certain transformations. However, Stevens added a critical aspect for empirical science that viewed the starting point of measurement in terms of empirical operations on a set of "real world" objects. This supplementation greatly generalized the notion of additive empirical operations put forth by Campbell (1920, 1957) as necessary components of scientific measurement, although the latter is included in the foundational scheme at the strong level of extensive measurement.

Extensive measurement, as it is treated in this line of work, depends on "additive" or more formally, concatenative (literally "linked together") operations. Concatenation along with a set of axioms (see, e.g., Roberts, 1979, chapter 3) results in a representation theorem, which states the existence of a function from the objects into the set of numbers, obeying a homomorphic relation with the original empirical operations and real objects. "Homomorphic" implies that the important empirical relationships are preserved in the numerical system which represents the real world phenomena.

The second and final major stage is the uniqueness theorem. This theorem gives the ways in which the numbers assigned to the real world objects can be changed without damaging the truth of the numerical representation of these objects.

Perhaps the very first scale of measurement historically was also the strongest. *Absolute measurement* permits no alteration whatsoever of the numbers assigned to the real objects. The counting of things must obey this principle and surely emerged relatively early, although some primitive civilizations seem to have relied on the concepts of "one," "two," and "many" (see, e.g., Klein, 1972). The decibel scale of loudness provides a further example of an absolute scale. It depends on the ratio of a physical unit to a base (usually a psychological threshold) measure with the same unit, thereby canceling out the unit of measurement and delivering an absolute scale.

The next strongest scale is the *ratio scale*, associated in the foundational tradition with extensive measurement. Along with counting operations, it is certainly dominant if not omnipresent in physical measurement. It permits a change of unit but not of origin, that is, there is a natural unalterable zero point, such as zero mass or zero length. One may multiply the representing numbers by a fixed positive number, which thereby alters the unit of measurement. Thus, we change from mass in pounds to mass in kilograms by multiplication of the pounds measurement by .454 kg/lb because the latter is the weight in kilograms of a single pound. However, clearly it would be inappropriate to add or subtract a number from our measurements of mass, thus changing the meaning of "zero weight."

Below the ratio scale lies the *interval scale*, whose uniqueness properties allow change of the origin, that is addition of a constant, as well as change of the unit. Calendar time, for instance, Christian versus Chaldean versus Muslim and

so forth, permits the change of origin in going from one system to another, although many seem to be based on some notion of month and year. In 1989, the bicentennial year of the French revolution, it was curious to learn that one of the changes promulgated by the revolutionaries was a return to the ancient Greek week of 10 days! Fahrenheit and Celsius temperature are other examples of interval scales.

The next lowest type of scale in this hierarchy is *ordinal measurement*. As befits the name, the numbers preserve only the order of the original objects and consequently any function that preserves the order of the assigned numbers may legitimately be visited on them.

The final scale is the weakest: the *nominal*. Nominal measurement, to the extent it can be dignified by the title, depends only on being able to discriminate different objects and thereby give them unique names. A transformation that again assigns different indexes to distinct entities is acceptable. On the other hand, few meaningful mathematical operations, outside of certain set theoretical actions, can be performed on objects falling within this level of measurement.

THE POSSIBLE SCALES OF REACTION TIME

The central question here concerns the placement of reaction time within this scheme of modern measurement. It is a thesis of this paper that reaction time can be employed in a way that is nothing more nor less than the application of physical time to psychological science. Now, time in the physical sciences is usually considered on the same footing with mass and length and in fact, along with these, forms the three classical measurement dimensions of physics. (Obviously, length is employed in all three dimensions of space.) Much of physics would fall if time were suddenly relegated to a weaker level of measurement. Let us call this notion of time "physical time." The immediate inference is that reaction time would then lie on a ratio scale and merit all respect and privileges, as well as the constraints on alterations of representation, accorded physical time.

Another quite different usage of reaction time relies on it as an indirect measurement of some psychological quantity or quality. This mode requires establishment of its scale properties vis-à-vis the psychological attribute. We discuss both of these roles of time in what follows.

The question of the scale of reaction time has, for the most part, been ignored, even by reaction time "practitioners." Even the foundational measurement theorists have tended to neglect the issue, perhaps because they have been more concerned with the seemingly more egregious attribution of strong scale properties to psychological variables whose measurement status has not been formally established. Thus, Krantz and colleagues (1971, 1989, 1990), Roberts (1979), Pfanzagl (1971) and Narens (1985) omitted discussion of the scale of reaction

time, despite its importance for experimental psychologists. The situation appears little better in the reaction time literature itself. The more formal approaches as exemplified by Luce (1986) and Townsend and Ashby (1983) as well as the less formal, for example, Welford (1980) and Posner (1978) offer little or no discussion of this issue. The Townsend and Ashby book gives a brief commentary on pages 390–396.

If reaction time were supposed to always lie on an ordinal scale, which some psychologists apparently believe, then a great many of the conclusions drawn from reaction time studies regarding mental processes would have to be thrown out. Some of them may deserve this fate but perhaps not all.

UNCOVERING MENTAL ARCHITECTURE AND REACTION TIME MEASUREMENT

To understand this critical issue, we need to take a look (more like a peek) at how reaction time is typically used in psychology. This question itself deserves more space than we are able to devote to it here. Perhaps the least penetrating way in which it is used is to simply compare tasks, manual equipment, individual performance level, and so on, with reaction time as the measure. Of course, if only ordinal uses are made, then the scale strength question will not often arise. However, if it is claimed that, say, one individual performs twice as fast on a certain task as another, then the scale must be ratio. Averaging does not avoid this precept. Suppose reaction time were on an interval scale and we compared the average reaction times of two individuals. It is permissible to change the unit and the origin under the conditions for interval scales. But changing the origin implies adding or subtracting a constant that will not cancel out when a ratio of the two averages is formed, thus changing the ratio. As the ratio of the means is not invariant in this case, it does not constitute a meaningful statement.

Another, more profound role of reaction time, still in its physical sense, is to draw inferences about the make-up of mental processes and how they interact. Now it may turn out that the brain simply acts as a great bowl of smart Jello — the most extreme form of Lashley/Gabor/Pribram holographic functioning—as far as its cognitive status is concerned (see Lashley, 1933; Gabor, 1948, 1968, 1969; Pribram, Nuwer, & Baron, 1974). This seems unlikely except for certain processes and/or certain confined subregions of the brain. Even in that untoward event, or within such a homogeneous subsystem, reaction time may aid in describing the way in which it processes information. If there do exist modules (e.g., Fodor, 1983) or even less self-contained subsystems (e.g., Smith, 1968; Sternberg, 1969), then the question of the overall architecture and the dynamic functioning of the subset of these engaged in a task requires the employment of reaction time in this more abstract interpretation.

Cognitive theoretical and experimental methods based on reaction time, often

in conjunction with other measures, are limited only by the imagination of the researcher. However, there is one line of theory and associated methodology that seems especially apposite to the present discussion.

This strand of cognitive research goes back to the 19th century. It supposes that experimental factors, or variables, can be manipulated to shed light on mental architecture. Factorial experiments, where combinations of levels of several experimental variables are used in tasks to answer psychological questions, are ubiquitous in behavioral science, as they are in many other disciplines. However, the present topic is a special case of these where a factor does something to the underlying processes that permits conclusions about the structure or functioning of those processes. A good term for such methods might be *system-factorial methods*.

We cannot go into this topic in depth here as it is treated in considerable detail a number of other places such as: Ashby and Townsend (1980), Donders (1869/1969), Pachella (1974), Pieters (1983), Schweickert (1978), Schweickert and Townsend (1989), Sternberg (1969), Taylor (1976), Theios (1973), Townsend (1984), Townsend and Ashby (1983), and Townsend and Schweickert (1989).

Donders (1869/1969) believed that by manipulation of the complexity of the cognitive task, an investigator could determine the average amount of time required by a psychological process (e.g., perception, judgment, and the like). All the processes were assumed to be arranged in series. About 100 years later, Sternberg generalized the method to permit answering the question as to whether two processes were in series *and* separately affected by two experimental factors. Taylor (1976) pointed out the need for a more subtle methodology and offered ideas concerning series of processes that might overlap in time in their operations. Schweickert (1978) developed a theory and methodology for complex deterministic (processes take the same amount of time for their operations on each trial) architectures described by more complex networks. Townsend and Piotrowski (1981), Townsend and Ashby (1983), and Townsend (1984) developed factorial methods that could discriminate simple serial versus parallel probabilistic (i.e., stochastic rather than deterministic, so that processes could take different amounts of time on each trial) networks. Later papers by Schweickert and Townsend (Schweickert & Townsend, 1989; Townsend & Schweickert, 1985; Townsend & Schweickert, 1989) established general theorems for probabilistic latent mental networks. Important allied work on probabilistic mental networks was accomplished by Fisher and Goldstein (1983; Goldstein & Fisher, in press).

System factorial methods work within a general theory—*metatheory* might be a better term—of how cognitive processing takes place, leaving the details of the architecture to be determined by the factorial experiments. The best understood methodologies assume “selective influence,” that is, a particular experimental factor affects exactly one process in the overall system (e.g., Ashby & Towns-

end, 1980; Sternberg, 1969). The effect of changing an experimental factor is assumed to lengthen reaction time.

Suppose that two processes A and B are lined up in series with process A starting and finishing and B starting exactly when A is completed. It is readily seen that if two factors, call them (*a*) and (*b*), are used to lengthen the two process times, T_A and T_B , then (assuming selective influence) the effects of (*a*) and (*b*) will be additive (Sternberg, 1969; Townsend & Ashby, 1983, chapter 12). That is, the increment in overall mean reaction time caused by changing (*a*) will not depend on whether or not (*b*) was changed also. Now compare additivity with the case of independent parallel processing where both processes must be completed. In standard parallel processing, both processes begin at the same instant but either might finish first. Observe that in both serial and parallel system we assume that the processing times may be random, governed by some probability distributions. It has been shown that if selective influence is in force, then the effects of the factors (*a*) and (*b*) are subadditive (Townsend, 1984; Townsend & Ashby, 1983). This means that if, say, factor (*a*) has already incremented reaction time, then the effect of (*b*) in further incrementing reaction time will be less than if (*a*) had not been so employed. For instance, if a graph of mean reaction time is drawn as an increasing function of one factor on the x-axis and the other factor as parameter, then the two functions would appear to be converging. Superadditivity would hold if the same functions diverged as they increased.

If reaction time lies on a ratio scale, then these facts permit the experimental discrimination of serial versus parallel processing. The investigator simply picks two values of each factor, runs the appropriate experiment under the four combinations of factor values and plots the results; plus statistical tests if necessary. However, if reaction time lies on an ordinal scale, then foundational measurement theory allows any order-preserving function to be applied to the reaction times. The implication is that the parallel data could be transformed to look like the serial, that is, additive and therefore, the two diametrically opposed types of processing cannot be experimentally tested by this, and possibly any other, reaction time method. Although there exist other methods, based, for example, on accuracy, for testing parallel versus serial processing, those based on reaction time form a strategic part of the armada (e.g., Townsend, 1990).

It should be understood that as far as the methodology itself is concerned, it is irrelevant that the psychologist is dealing with organisms: All the concepts and inferences go over into situations where the "black box" is a machine. Note, too, that the present methodology makes no assumptions about the way in which the individual subsystems (processes, etc.) go about their business. As long as the precepts of the class of systems under consideration—in the preceding, parallel versus serial processing systems, and more generally, systems whose sub-processes complete their tasks in a discrete fashion before an immediately succeeding one can begin—are obeyed by the "black box," the state description does not matter. Of course, a state-space description within theories and related

methodologies could much strengthen system identification and theory testability, and psychology needs more of it.

It seems doubtful that more than a tiny fraction of the methods that employ reaction time to draw inferences about mental structure or functioning could withstand the damage caused by loss of the ratio scale property. Techniques that rely on the form of reaction time functions (not necessarily just probability distributions) are particularly susceptible to injury from ordinal scale changes. Such strategies that are robust must be impervious to any transformation on reaction times that preserves their order. I discussed such possibilities in a general context in a recent paper (Townsend, 1990b).

WHEN REACTION TIME SHOULD LIE ON A RATIO SCALE

Insofar as physics treats the dynamics of objects from an external point of view—which is certainly true of most Newtonian mechanics—the similarity of the use of reaction time to physical time is not so apparent. It is in the application of physics to identification of continuous time systems that the relationship may become more apparent.

Dynamic systems analysis is by definition intimately related to change in time and it should therefore be no surprise that an arbitrary monotonic transformation on the time axis would usually cause havoc in theoretical or practical measurement, prediction, or control interpretations. For purposes of illustration, we discuss only a few, extremely simple cases. Some recent excellent introductions, in my opinion, to systems theory are Luenberger (1979) for linear systems, Beltrami (1987) for nonlinear systems, and Devaney (1986) for chaos theory. My colleagues and I have found notions of systems theory to be of help in some of our work (e.g., MacCallum & Ashby, 1986; Townsend & Ashby, 1978; Townsend & Ashby, 1983, pp. 401–412; Townsend & Evans, 1983). Linear and nonlinear systems theory forms the backdrop for a good deal of the neuropsychological modeling of cognitive activity (e.g., Anderson, 1973; Grossberg, 1987a, 1987b; Kohonen, 1977).

Consider a so-called linear system. Such systems are called linear because if two signals are input instead of just one, the output will be the same as the sum of the outputs of the signals, each input alone. They can generally be described by sets of linear differential equations. The behavior of a linear system can be determined in exact form, from its differential equations. The theory of such systems is basically complete and they have been the workhorses of physics and the other sciences up to the present. Although nonlinear systems, including the exotic chaotic systems, are becoming increasingly useful in most disciplines, the assumption of linearity will undoubtedly continue to play an important role in scientific investigation.

One of the supreme theoretical results in linear systems theory states that any individual system can be completely described by its impulse response function. The impulse response function, often called $h(t)$, is a function of time that depicts the memory of the system for past signals. Thus, if $h(t) = \exp(-at)$, the interpretation is that as time grows larger, past inputs will have a decreasing effect on the current output.

The output, $x(t)$, of a linear system with input function $u(t)$, and assuming the state of the system at time $t = 0$ is itself 0, can be expressed as the integral

$$x(t) = \int_0^t [u(t') \cdot h(t-t')] dt'$$

Observe that the output, x , at time t is the integrated (summed) activity of present $[u(t) \cdot h(0)]$ as well as past inputs weighted by the memory of the impulse response function. For instance, the contribution from the beginning of processing is just $u(0) \cdot h(t)$; that is, the input at time 0 weighted by the memory of t time units in the past.

The impulse response function is called by that name because if an impulse is input to the system, the output will be the impulse response function. An impulse is intuitively just an instantaneous spike, so of course, the output of the system simply reveals the memory of that spike, which in turn completely describes the functioning of the system.

Suppose an engineer is interested in determining the nature of a linear system and for some reason cannot, or does not wish to, invade it with screwdriver and other tools. Our engineer can simply input an impulse and observe the output. The engineer has acted much like a behavioral psychologist who wants to understand the nature of a cognitive process:

1. Both approach a black box with largely unknown contents.
2. Both rely on input-output behavior in order to discover the functional workings of the system.
3. Both often need to assume time as a ratio scale variable for their assessment to work.

It is clear that even an order preserving function of time will usually greatly distort the impulse response function. For instance, take the impulse response function used earlier, $h(t) = \exp(-t)$. Now, let $t = \ln(s)$, that is, t is the natural logarithm of the substitute time variable. Then the new impulse response function, $h(t)$ becomes $h[\ln(s)] = 1/s$. This new impulse response function still decreases as a function of time, but its form is quite different from the original. In addition to simply giving a very different function, the vital nature of the system has been altered.

This can be appreciated through the assessment of the stability properties of the original system versus its counterpart, obtained by way of the time transformation. Basically, a system is stable if bounded inputs lead to bounded outputs.

A sufficient condition to make this happen is for the integral of $h(t)$ over all time, to be finite. This certainly occurs for the first impulse response function because that integral is just equal to 1. However, if our engineer began with the new function, say $h(s) = 1/s$, then the integral over all time turns out to be infinite and thus the system is determined to be unstable. For instance, input of the infinite square wave of height 1, starting, say, at time $t = 0$, leads to an unbounded output as time grows large. All would probably agree that the exact numerical description of the system has been irrevocably altered by the time change.¹

Many other qualitative aspects of system functioning and description will be disfigured by a time transformation. Thus, the so-called rise time and setting time depend immediately on the time scale. The frequency description and power spectrum of a signal will be changed, and so on.

The overriding point to be made is that the engineer is acting very much like a behavioral psychologist in her use of time and would think it quite silly indeed to permit any alteration of it other than a change in unit.²

RELEVANT DISCUSSIONS FROM THE PSYCHOLOGICAL LITERATURE

Several tomes on measurement, as well as a number of articles, offer very useful discussions about the correct and incorrect use of measurement in practical and scientific situations. Examples of books that emphasize some aspects of applications are Krantz, Luce, Suppes, and Tversky (1971, 1990) and Roberts (1979). For instance, Roberts took pains to provide applicators with practical tools to assess the meaningfulness of their everyday measurement operations (e.g., Roberts, 1985). However, most of the examples and provisions seem devoted to what the researcher should or should not do, given the specified level of measure-

¹It might be argued that there is another interpretation that saves the stability assessment. If one started with the original impulse response function, wrote down the expression for its integral and then performed the time change of variable according to standard techniques, the value of the integral would of course be unchanged (and therefore still finite). This topic may deserve more investigation than we can pursue here, but it might be asked if the scale (e.g., units, origin, etc.) of time is really unimportant, then why should one have to keep track of it to reach an understanding of the system? By analogy, if a geologist is working with the ordinal scale of hardness, then it is patently unimportant exactly what numbers are assigned to the various rock samples as long as the order is preserved.

²A difference between the example concerning impulse response functions and the typical reaction time methodologies is that the latter must always rely on data compilations over many trials whereas in principle, the engineers could determine the system characteristics with a single sample. However, in real life, engineers must ordinarily employ a number of samples due to system and measurement noise, and many advanced techniques rely by their nature on time or space averages. So, it seems that this difference is not really a substantive one.

ment scale with which one is working. In the present case, the latter is the primary question.

A judicious, nontechnical discussion of the general purposes, uses, pitfalls, and benefits of measurement is a paper by David Krantz in 1972. Although that was some time ago, the important messages for present purposes do not appear to be dated in any serious way. There are two particular aspects that are relevant to our discussion. The first is the major one for us and is perhaps one of the most valuable points in the paper overall.

Physical measurements obviously can be used as they are in, say, physics and chemistry to formulate new relationships, test predictions, and validate laws. If however, as noted earlier, they are employed to represent psychological qualities or quantities, then the scientist must follow the tenets and procedures of measurement theory in order to claim legitimate psychological measurement. It would seem that the use of time in establishing the Bunson–Roscoe (also known as Bloch's) Law of Temporal Integration of Energy in Vision, which says that up to about 100 ms, only the entire amount of energy in the visual signal is important for detection, is an instance of the first usage.

In contrast, the use of reaction time to *represent a psychological variable* or construct such as response strength, is categorically different. If the proper measurement steps are not taken, then this representation is little more than a circular definition. Krantz gave an example of this type of "definitional" measurement in color vision, whose dangers appear general (1972, p. 1434). An example of the use of time in establishing measurement laws relating several experimental variables to time considered as a measurement of response strength for animal learning is also given in that paper (pp. 1428–1429).

In my opinion, reaction time as implemented in the systems factorial methods and theories discussed earlier (e.g., Schweickert & Townsend, 1989; Townsend & Ashby, 1983; Townsend & Schweickert, 1989), should be viewed in terms of its *physical measurement properties*. Many other, though perhaps not all, applications of reaction time by sensory and cognitive psychologists also appear to fall within this framework. On the other hand, the interesting paper by Hans Micko (1969) refers to reaction time in the second sense, as a measure of some other psychological quantity. There, Micko argued that reaction time distributions stand for distributions on a psychological time scale that differ from physical time.

Another comment by Krantz (1972) points out how an investigator could assess a loosened notion of linearity in a generalized type of dynamic system, a system in which the input might only lie on a nominal scale and the output is measured ordinally. It turns out that such an assessment can be carried out using standard principles of foundational measurement. Now, this topic obviously pertains to the state description of the system, that is, the specification of the internal state of the system as a function of time, as well as that of the input and output. In Krantz's example, the state was taken as the output of the system, a

common practice in very simple systems. Further, it is germane to the earlier discussion of the engineer who is attempting to identify the impulse response function of a system. If the scales with which the engineer is working are not sufficiently strong, for example ratio, then her description of the impulse response function may not tell her what she thinks it will. Of course, she is assuming linearity, but if the scales are weak, such as in Krantz's example, then even her traditional ways of testing linearity may be flawed.

The intricacies of how psychological and neural state-space descriptions (not to mention input and output scale questions) relate to time and other dimensions remain to be explored; and this is an important problem. However, I assert that even if input is nominal, as long as the psychological systems function in a lawful fashion, that a punctilate measure of reaction time can be legitimately interpreted as a physical measure and fruitfully employed in identification of mental architecture and processing. By lawful, I mean that there exists a function taking the input to a unique output, on a single occasion. This notion permits the introduction of probability as the unique assignment could change from trial to trial. A generalized conception might permit one-many deterministic functions but we don't require this generality at present.

Suppose that the mental system under consideration is composed of a finite set of processes which each take a finite and (for now) deterministic amount of time to perform their operations. Let that set be designated as S , with individual members being s_i , $i = 1, N$. Now suppose that the reaction time is some function $G(t_1, t_2, \dots, t_1, \dots, t_N)$ of the times taken by the set of processes. Note that G need not be monotonic in the N variables, although it probably will be in some of them, in realistic cases. By the earlier precept, the mapping G carries a given set of the t_i to a unique point, although more than one set of the t_i might yield the same point. Methods like the systems-factorial methods work because each process is performing a task that could (if only we possessed a better window into the nervous system) be measured in terms of the physical time it consumes to complete that task. Certainly, it would violate the spirit of time measurement to change anything other than the units of such a measurement. Similarly, the overall reaction time is inherently related to the internal durations by $G(t_1, \dots, t_N) = RT$, and therefore one also is not sanctioned to perform arbitrary transformations on reaction time. This general argument is not influenced by an imposition of probability distributions on the internal times.

This discussion is pertinent to the earlier comments about state spaces in that a more microlevel description of the system components would pinpoint more detail about what subtask each process is performing. However, the actual nature of these state descriptions does not vitiate the argument about the physical nature of time as utilized in such methodologies. It also does not appear to be impugned by Krantz's point, mentioned earlier, about using measurement principles to prove a generalized system linearity. To see this, suppose that the input u , which might be on any type of scale, is subjected, as Krantz said, to an arbitrary input,

so that its effect may be only nominal, even if it were externally ratio. Call the internal input $v(t)$. Suppose for simplicity that only one processor is involved, so that the system state $x(t)$, is a function of v from $t = 0$ up until the present moment, and that a response is made when $x = x^*$. Now as long as our postulate about the uniqueness of a mapping or function is satisfied, $x(t) = H(v; 0, t)$, where the second and third terms in H designate the starting time and present moment, respectively, then the amount of time it requires for x to reach x^* is not distorted. That is, the time taken by the processor to complete its job cannot be arbitrarily transformed just because the internal version of the input is not immediately measurable on a strong scale. As a simple example, consider an electronic voice recognizer confronted with a person's name as input. Then the time consumed for recognition, correct or not, must be accepted as a physical measurement, whatever internal physical and information theoretic transformations are imposed on the input, and indeed, whatever use is made of previous memories or other information.

Suppose it be admitted that reaction time can be used as a physical scale in, say, the factorial methods espoused by Donders (1869/1969), Sternberg (1969), Taylor (1976), Schweickert (1978), Schweickert and Townsend (1989), and Townsend (1984), Townsend and Ashby (1983), Townsend and Schweickert (1989). Can something go wrong if the scales of two experimental variables (i.e., the factors are on weak scales such as nominal)? The answer is no, as long as we accept the relatively weak postulate that manipulation of the two factors leads to a legitimate probability distribution on reaction time with a finite mean.

Let us do our reasoning with ordinal scales and then make a comment about nominal scales. Thus, assume we have a probability distribution on reaction time as a function of the two factors and we calculate the mean in each of four cases. Note that the unit of the mean will change with a change in the unit of the individual times. Assume that the x and y factors are on ordinal scales and, for simplicity, assume that the mean can be written as a twice-differentiable function, with appropriate continuity conditions of them both. (The latter assumption is not necessary, as the same type of argument goes through with discrete operations.) Then as Townsend and Ashby (1983) showed, the qualitative contrast (i.e., additivity versus subadditivity versus superadditivity) can be represented by the sign of the second order mixed partial difference. That is,

$$C(x,y) = \frac{\partial^2 T(x,y)}{\partial x \partial y}$$

Now, let us perform admissible transformations on both x and y , that is, let $x' = r(x)$ and $y' = s(y)$, where both r and s are strictly increasing and appropriately differentiable functions of x and y . Suppose without loss of generality that $C < 0$ for all x and y in the experimental range. Now, $T(x,y) = T[r(x'), s(y')]$. By the chain rule, it follows that

$$C(x',y') = \frac{\partial^2 T}{\partial r \partial s} \cdot \frac{dr}{dt} \cdot \frac{ds}{dy} .$$

The signs located above the terms show that, overall, the signs of the contrast is invariant across the alterations of the factors. Therefore, admissible transformations in the scales of the experimental factors do not harm the qualitative conclusions, needed to assess the mental architecture. Of course, as noted earlier, the investigator should always be aware of the character of the scientific statements she is making, and not just the scale type per se. For instance, a ratio of differences of interval scales is scientifically and measurementally valid, even though a simple ratio of two interval measurements is not. What about change of scale of the dependent variable, that is, reaction time? As before, a usual change of unit is allowed but no looser transformation is permitted. If one could perform such a transformation, then the qualitative result could indeed be deformed.

If x and y are only nominal, what then? As long as the "levels" of x and y are chosen by the experimenter to have an effect then they can be reordered to be compatible with the experimental results. That is, we can order x_1 and x_2 so that the mean reaction time $T(x,y)$ obeys $T(x_1,y_i) < T(x_2,y_i)$, as long as this inequality holds for all y_i . If it does not, then the matter is more complicated and I have not worked out the consequences. But, if the inequality does hold, then all seems to proceed as before. This inequality is somewhat analogous to the fact that conjoint measurement (e.g., Krantz et al., 1971) can be built up from nominal scales, if the condition is interpreted as single-factor independence, of course, in conjunction with other requirements. However, the difference is that in conjoint measurement, the dependent variable scale may be transformed so that additivity results, but this is not permitted under the present circumstances.

CONCLUSION

This chapter attempts to examine reaction time within the framework of modern measurement theory and proposes that many of the theoretical and methodological uses of reaction time may be considered in the physical sense. Of course, as noted in the discussion, such use does not preclude its employment in the second sense, to represent another psychological variable. To some investigators, the idea that a dependent variable that possesses strong scale properties in physics can be used as a physical measure in the behavioral sciences may seem self-evident. However, others appear to have confused the two uses and cite such papers as Krantz (1972) as support for the contention that all uses of reaction time may necessitate the formal development of scale type; or even that reaction time simply "lies on an ordinal scale."

Arguments and examples were proffered to attempt to convince the more

skeptical reader of the existence of the two roles for reaction time, and of the separate regulations concerning their implementation.

REFERENCES

- Anderson, J. A. (1973). A theory for the recognition of items from short memorized lists. *Psychological Review*, *80*, 417-438.
- Ashby, F. G., & Townsend, J. T. (1980). Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology*, *21*, 93-123.
- Beltrami, E. (1987). *Mathematics for dynamic modeling*. San Diego, CA: Academic Press.
- Boring, E. G. (1957). *The history of experimental psychology*. New York: Appleton-Century-Crofts.
- Campbell, N. R. (1920). *Physics: The elements*. Cambridge, UK: Cambridge University Press.
- Campbell, N. R. (1957). *Foundations of science: Philosophy of theory and experiment*. New York: Dover.
- Donders, F. C. (1969). On the speed of mental processes. In W. G. Koster (Ed. & Trans.), *Attention and Performance, Vol. 2* (pp. 412-431). Amsterdam: North Holland Press. (Original work published 1869)
- Devaney, R. L. (1986). *Chaotic dynamical systems*. MenLo Park, CA: Benjamin/Cummings Publishing Co.
- Fisher, D. L., & Goldstein, W. M. (1983). Stochastic PERT networks as models of cognition: Derivation of the mean, variance, and distribution of reaction time using order-of-processing (OP) diagrams. *Journal of Mathematical Psychology*, *27*, 121-151.
- Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Gabor, D. (1948). A new microscopic principle. *Nature*, *161*, 777-778.
- Gabor, D. (1968). Holographic model of temporal recall. *Nature*, *217*, 584.
- Gabor, D. (1969). Associative holographic memories. *IBM Journal of Research & Development*, *13*, 156-159.
- Goldstein, W. M., & Fisher, D. L. (1991). Stochastic networks as models of cognition: Derivation of response time distributions using the order of processing method. *Journal of Mathematical Psychology*, *35*, 214-241.
- Grossberg, S. (1987a). The adaptive brain, I. *Advances in Psychology*, *42*. Amsterdam: North Holland.
- Grossberg, S. (1987b). The adaptive brain, II. *Advances in Psychology*, *43*. Amsterdam: North Holland.
- Klein, M. (1972). *Mathematical thought from ancient to modern times*. New York: Oxford.
- Kohonen, T. (1977). *Associative memory: A system-theoretic approach*. Berlin: Sperling-Verlag.
- Krantz, D. H. (1972). Measurement structures and psychological laws. *Science*, *175*, 1427-1435.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement, Vol. 1*. New York: Academic Press.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1989). *Foundations of measurement, Vol. 2*. New York: Academic Press.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1990). *Foundations of measurement, Vol. 3*. New York: Academic Press.
- Lashley, K. S. (1933). Integrative functions of the cerebral cortex. *Physiological Review*, *13*, 1-42.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- Luce, R. D., & Narens, L. (1987). The mathematics underlying measurement on the continuum. *Science*, *236*, 527-532.

- Luenberger, D. G. (1979). *Introduction to dynamic systems: Theory models and applications*. New York: John Wiley.
- MacCallum, R., & Ashby, F. G. (1986). Relationships between linear systems theory and covariance structure modeling. *Journal of Mathematical Psychology*, 30, 1-27.
- Micko, H. C. (1969). A psychological scale for reaction time measurement. *Acta Psychologica*, 30, 324-335.
- Narens, L. (1985). *Abstract measurement theory*. Cambridge, MA: MIT Press.
- Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. H. Kantowitz (Ed.), *Human Information Processing: Tutorials in Performance and Cognition* (pp. 41-82). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pfanzagl, J. (1971). *Theory of measurement* (2nd ed.). New York: Wiley.
- Pieters, J. P. M. (1983). Sternberg's additive factor method. Psychological processes: Some theoretical considerations. *Psychological Bulletin*, 93, 411-426.
- Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pribram, K. H., Nuwer, M., & Baron, R. J. (1974). The holographic hypothesis of memory structure in brain function and perception. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology*, Vol. 2 (pp. 416-457). San Francisco: Freeman.
- Roberts, F. S. (1979). *Measurement theory with applications to decision making utility and the social sciences*. Reading, MA: Addison-Wesley.
- Roberts, F. S. (1985). Applications of the theory of meaningfulness to psychology. *Journal of Mathematical Psychology*, 29, 311-332.
- Schweikert, R. (1978). A critical path generalization of the additive factor method: Analysis of a stroop task. *Journal of Mathematical Psychology*, 18, 105-139.
- Schweikert, R., & Townsend, J. T. (1989). A trichotomy method: Interactions of factors prolonging sequential and concurrent mental processes in stochastic PERT networks. *Journal of Mathematical Psychology*, 33, 328-347.
- Smith, E. E. (1968). Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, 69, 77-110.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and Performance*, Vol. 2 (pp. 276-315). Amsterdam: North Holland Press.
- Stevens, S. S. (1951). Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology* (pp. 1-49). New York: Wiley.
- Suppes, P., & Zinnes, J. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Vol. 1 (pp. 1-76). New York: Wiley.
- Taylor, D. A. (1976). Stage analysis of reaction time. *Psychological Bulletin*, 83, 161-191.
- Theios, J. (1973). Reaction time measurements in the study of memory processes: Theory and data. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation*, Vol. 7 (pp. 43-85). New York: Academic Press.
- Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology*, 28, 363-400.
- Townsend, J. T. (1990a). Serial vs. parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. *Psychological Science*, 1, 46-54.
- Townsend, J. T. (1990b). The truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, 108, 551-567.
- Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan and F. Rustle (Eds.), *Cognitive Theory*, Vol. 3 (pp. 199-240). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.

- Townsend, J. T., & Evans, R. (1983). A systems approach to parallel-serial testability and visual feature processing. In H. G. Geissler (Ed.), *Modern Issues in Perception* (pp. 166–191). Berlin: Veb Deutscher verlag der wissenschaften.
- Townsend, J. T., & Piotrowski, S. M. (1981). *Parallel systems and factorial additivity of mean reaction time*. Paper presented at Psychonomic Society Meeting, Philadelphia, PA.
- Townsend, J. T., & Schweikert, R. (1985). Interactive effects of factors prolonging processes in latent mental networks. In G. d'Ydewalle (Ed.), *Cognition Information Processing & Motivation, Vol. 3* (pp. 255–276). Amsterdam: North Holland Press.
- Townsend, J. T., & Schweikert, R. (1989). Toward the trichotomy method: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology, 33*, 309–327.
- Welford, A. T. (1980). *Reaction times*. New York: Academic Press.