Citation: Neufeld, R.W.J. (in press). Composition and uses of formal clinical cognitive science In B. Shuart, W. Spaulding & J. Poland (Eds.), *Modeling Complex Systems: Nebraska Symposium on Motivation, 52.* Lincoln, Nebraska: University of Nebraska Press.

Composition and Uses of Formal Clinical Cognitive Science

Richard W. J. Neufeld

University of Western Ontario

"In any particular theory, there is as much real science as there is mathematics". So wrote Immanuel Kant about the status of natural philosophy, as he saw it at the time. This declamation seemingly reflected a certain frustration not unlike that possibly experienced some two centuries earlier by Francis Bacon, who decried "... the distemper of learning [that occurs] when men study words and not matter." Contemplating these quotes, the reader senses they may be viewed as a commentary on much of the activity done in the name of science by the authors' contemporaries– activity possibly characterized by considerable redundancy of effort to resolve the day's purportedly key issues, contentious debate devoid of clear resolution, a lack of precise place keeping regarding the status of important segments of "the body of knowledge", and absence of salient direction for potentially profitable future thrusts of investigation.

More contemporary, and emanating from within our discipline, Paul Meehl (1978) bemoaned the slow progress of "soft psychology" (which presumably included psychological clinical science). He indicted the dominant research paradigm (Kuhn, 1962) of which a prominent modus operandi was, and for that matter remains, Fisherian- and Pearsonian-based statistical methods. Theoretical predictions were derided as somewhat anemic, typically embracing simply the presence of non-zero relations among ad-hoc measured variables ; as nature is said to abhor an absence of association, so-called support rested on statistical power, whereby the ever-popular ANOVA and related tables of results, in effect were summary statements of such power (cf. Cohen, 1988, pp. 16, 17; Steiger & Fouladi, 1997). He noted that "theories", at least in fields of soft psychology, typically fell by the wayside, not because they were toppled by uncorroborated risky predictions, but because the discipline became inured to them, and their proponents grew weary of marketing them. It might be added nevertheless that their proponents need not worry, because it is well known that old theoretical ideas frequently re-emerge, in a different guise, and with modernized labels (Staddon, 1991). (Most readers will have their favorite accounts of such instances, from their particular fields of expertise.)

I would contend that the foregoing lamentation retains considerable currency when it comes to contemporary psychological clinical science. This observation provides much in the way of motivation for the developments presented here. It is maintained that the implementation of formal theory in psychological clinical science can go far to redress current quagmires, and accelerate progress. Emphasis of course is on the field with which I am most familiar, cognitive clinical science. However, merits of formal theory disclosed in that domain of investigation arguably are general, providing additional incentive for the particular sphere of application.

To begin, a clear delineation of what does an does not constitute formal theory is presented. Benefits and selected (practical) liabilities are described. Following on is an

exemplary case in point from the domain of clinical cognitive science. Novel developments include proposals for monitoring cognitive aspects of an individual's response to treatment over time, and likewise dynamical assessment of treatment-program efficacy with respect to cognitive functioning. Some surprising noteworthy spinoffs from excursions into formal modeling, that impinge directly on prominent issues in this area of study, are expounded upon. One such by-product emanates from a personal "labor of love", entailing the appropriation of stochastic modeling to the study of cognitive efficiency as it relates to schizophrenia, to psychological stress, and to their combination. It comprises a new form of construct validity, using mathematical heuristics. Another bears on an issue that I have engaged more or less since its appearance in the literature; namely the so-called psychometric artifact problem in the study of differential cognitive deficit. The problem is translated into simple but comprehensive measure-theoretic terms, which reveal the shortcomings of typical efforts at redress. From there, formal theory is shown to re-cast the entire issue onto its own terms, and in so doing renders it as more or less obsolete.

Demarcations of Formal Theory

Described here are features of formal theoretical systems that distinguish them for all practical purposes from non-formal and quasi-formal systems. Drawn upon is the rigorous, still-cogent taxonomy of logical-deductive systems in science, provided by R.B. Braithwaite (1968).

The most prominent type of theoretical system in psychological clinical science, Braithwaite's mixed deductive system, can profitably be used as a point of departure in explicating the nature and role of formal theoretical systems. Mixed deductive systems consist of formal and non-formal subsystems. A survey of contemporary issues of APA clinical-science journals (e.g., *Journal of Abnormal Psychology*; *Journal of Consulting and Clinical Psychology*), and any number of Psychiatry research journals, reveals the following predominant structure. Within the non-formal subsystem, predictions of empirical observations are dictated by syllogistic verbal reasoning, possibly accompanied by visual aids, and most certainly by antedating empirical findings. There is, however, no particular governance by a set of rules constraining the format by which the verbal arguments are posed, or the way in which one statement follows form another (e.g., mathematical derivation, computer language, or symbolic logic). The sequence of qualitative arguments culminates upon transition to a prediction(s) of results, or at least guiding question(s), credited with motivating the investigation at hand (further elaboration of informal and other systems, in selected domains of applied psychological science, is presented in Neufeld, 1989). Summarily speaking, the study then is executed, and results obtained.

It is at this point where the formal subsystem makes its entry. Measures of dependent variables, guided, if not "prescribed" as such, by the non-formal theoretical apparatus (see, e.g., McFall & Townsend, 1998; Meehl, 1978), provide the data array submitted to statistical analysis, expedited by associated software. Statistical treatment is part and parcel of a formal subsystem, because its constituent operations rest on closed-form formulae that by and large emerge from theorem-proof continuity (nowadays increasingly accompanied by numerical

simulation). Note that so-called qualitative research methodology replaces this formal data-analytic arm of the logical-deductive system with non-formal, verbal operations. In this way, the entire logical-deductive enterprise is rendered informal.

Conversely, formal-theoretical approaches replace non-formal reasoning in the above mixed deductive design with a formal logical-deductive subsystem, thus creating an entirely formal logical-deductive structure. Whereas the typical statistical apparatus (often elegant in its own right, and with a virtual industry of software support) is substantively generic-- its application transcending theoretical-content domains-- formal theoretical models are content-specific. For example, Doob (1953) has described one such type of theoretical formulation, a stochastic model, as a "... mathematical abstraction of an empirical process, whose development is governed by probabilistic laws p.v". Whereas statistical computations comprise mathematical methods for assessing arrays of values rendered by empirical investigations, formal theoretical models are concerned with processes that make the data what it is (although, as will be demonstrated below, formal theories are far from mute about ways of analyzing the data whose genesis they address). In other words, in the explanatory chain of events, such models weigh in with quantitative constraints *ab initio*.

The nature of authentically formal theory can be thrown into sharper relief by considering procedures that may appear to qualify as the former, pending closer scrutiny. Integrative and evaluative reviews, for example, potentially useful in their own right to be sure, are a case in point. Apart from a rare application of synthesizing quantitative models of possible agents of surveyed outcomes, verbal schemata as a rule are restatements of what is already known (e.g., McFall, Townsend & Viken, 1995).

Formal theory also can be distinguished from prominent higher-level statistical methods, including "structural-equation causal modeling" (SEM; e.g., Kline, 1988; Tomarken & Baker, 2003; c.f., MacCallum & Ashby, 1986), taxometric procedures (Meehl & Golden, 1982; Waller & Meehl, 1988), and Hierarchical Linear Modeling (Bryk & Raudenbush, 1992; see Piasecki, Jorenby, Smith, Fiore & Baker, 2003). With respect to SEM, the subject of prediction comprises its parameters[1], to all and intents and purposes meaning the path coefficients of conjectured relations among the posited latent variables. (It is surmised that criterion values for contemporary indexes of overall statistical fit, a sort of overarching prediction, have been met.) Note that ordinarily, the hypothesized structure of inter-latent-variable linkage, and the relative magnitudes and/or signs of its associated path coefficients themselves are posited non-formally.

In a similar vein, confirmatory factor analysis forecasts factor loadings and coefficients expressing inter-factor relations, including those involving higher-order factors. Concern is with an empirical covariance structure, but whose features are defined according to antedating empirical observations, or non-formal argumentation, rather than derivations spawned by specific axioms, definitions and assumptions.

The chief subject of prediction for taxometric methods, in turn, are certain configurations of data summaries (calculated via methods prominently known as "Maxcov", "Mambac",

"Maxslope", : "Maxeig", and "L-mode"), indicative of dichotomous versus continuous latent distributions of target variables (e.g., specific symptoms, or syndromes). Selected parameters, such as base rates of a taxon (a discrete group with the symptom, genetic endowment, or other target characteristic) can be estimated. The data configuration corresponding to the conjectured distribution, however, once more is not in and of itself the product of formal-system reasoning. By these characteristics, taxometric methods and SEM fall into the category of statistical analysis, and in that way constitute the formal subsystem of a mixed deductive system (elaborated upon in Neufeld, 1989). Suffice it to say that similar observations attend typical applications of Hierarchical Linear Modeling.

Formal theoretical systems as well entail parameter estimation and testing for goodness of fit between predictions and observations. Their parameters, however, are substantively significant, as imbued by their explanatory roles in the theoretical model in which they participate (Braithwaite, 1968; exemplified below). Moreover, their variation generates selective changes in empirical data *sui generis* to the localized problem (e.g., Link, 1982; Wenger & Townsend, 2000; Townsend & Wenger, in submission). This treatment contrasts an essentially generic interpretation of parameter values, or other model properties, across even vastly divergent content areas, and a focus on highly similar aspects of data (e.g., empirical covariance structures).

We return momentarily to the non-formal system widely known as "qualitative research methodology", because of its growing popularity, especially in the social sciences. Qualitative research increasingly has been touted as a legitimate, even superior alternative to systems of science implanted with formal methods.

Non-formal strategies give the appearance of flexibility in dealing with concepts and measures whose implementation may strain formal systems. The former, however, accommodate such challenges handily, not because they somehow transcend the explanatory merits of formal systems, but because they simply sidestep the constraints on precision of expression demanded by formal systems or subsystems (Staddon, 1984). As such, despite their purported checks and balances (e.g., Miles & Huberman, 1994), they are vulnerable to the onslaught of frailties in reasoning, and inefficiencies accompanying dependence on essentially verbal contentions, defects that formal procedures are designed to confront (Hintzman, 1991; Kline, 1985). Indeed, the history of science makes salient that systematic retreat from decidedly formal inferential methods inevitably detracts from progress (e.g., Braithwaite, 1968; Kline, 1985).

Assets of Formal Theory

Explanatory Insights.

A major advantage bequeathed by formal developments is the unique insight afforded into the studied phenomena. Interplay of critical variables is disclosed by inspecting the structure of derived predictions, and the steps to obtaining them. Examples in the hard sciences of course abound. Consider, for instance, a notable case in point, Newton's universal law of gravity: $\gamma\, m_1 m_2 / r^2$. The law defines the force of attraction between two particles, with masses m_1 and m_2,

separated by distance *r*. The constant γ varies only with the units of measurement. The actual physical mechanism of gravitational pull may be enigmatic, but the mathematical definition is not. Moreover, to be taken seriously, any proposed physical mechanism would have to comply with this mathematical statement. The function of each term, and of the law itself, in turn are decipherable from the nature of the formal assertion, and from the way in which the law serves in broader theoretical contexts.

Similarly, something of the nature and function of Einstein's limiting constant, the speed of light, can be appreciated contextually, by considering its operation in a formal deductive system. Such exposition, for example, discloses that any meaningfulness of a larger value rests on a solution to the imaginary number $(-1)^{½}$. .

In psychological clinical science, elucidation of obtained data configurations can be bolstered by stipulating relations between observed and inferred variables in terms analogous to the above. As opposed to ad-hoc interpretation, statistical interactions, for one, can be understood from formal depiction of the psychological processes under study, and associated scales of measurement (e..g., Busemeyer, 1980; Staddon, 1984).

An illustrative instance entails the stochastic dynamical modeling (see e.g., McFall, Townsend & Viken, 1995) of effects of stress (ambient noise) on cognitive efficiency (Weinstein, 1974). One such model expresses the expected (mean) latency for proof- reading lines of prose (e.g., Glass & Singer, 1972). The expression is $mr/(k\text{-}1)$ (Neufeld, 1996; 2002). In this formula, *r* conveys stress-related impairment of performance speed (*r* goes up with increasing stress, elevating the mean, reflecting diminished speed), *k* is an increasing function of task-relevant competence (associated here with practice), and *m* indicates task load, quantified as the mean number of processed stimulus elements (chunks of letters scanned per line of prose).

The expected latency for a given level of stress, and amount of practice, for the summary task load *m*, thus can be predicted as a function of substantively-significant parameters. The expression also discloses the fabric of interplay between stress and practice, making for their "interaction": effects of stress can be mitigated by practice, practice becomes more important as stress goes up, and effects of this interplay on performance latency are "scaled" according to the prevailing task load *m*. In fact, the form of the stress-practice interaction is one of "superadditivity" of the interaction's second-order difference ({[mean latency under higher stress, and lower practice] - [mean latency under lower stress, and lower practice]} - {[mean latency under higher stress, higher practice] - [mean latency under lower stress, and higher practice]}) > 0; Townsend, 1984). Note that each of the above parameters also is entrenched in its own quantitative infrastructure, endowing it accordingly with further analytical and substantive meaning (Evans, Hastings & Peacock 2000; elaborated upon below).

Beyond their revelatory function, closed-form, analytical expressions also can serve the practical cause of computational economy. It is enticing to depend on numerical simulations as surrogates for analytical derivations, given contemporary computational technology. Analytical derivations, however, not only enhance understanding, according to their detailed structures,

but can reduce computational steps, and ease computer memory demands when simulations are inevitable. Partly for this reason, theorists in longer-established sciences will opt to take analytical solutions to the limit, before invoking numerical algorithms[2].

Self-Diagnostic, Self-Indicting Properties.

Desiderata of useful theory are Popperian bold conjecture, and falsifiability. In Bayesian terms, bold conjecture is tantamount to a predicted observation being negligible, apart form viability of the proposed theory. In like fashion, distinct falsifiability is identified with an essentially zero probability of the tendered theory being defensible, upon failure of its specified prediction. It is readily shown that in the first instance, the probability of the theory being tenable, given that the predicted evidence occurs, approaches 1.0; in the second case, the probability of the evidence, given the theory also approaches 1.0 (Neufeld, Vollick, Carter, Boksman & Jetté, 2002).

Predictions emanating from strong theory stand to be formidably precise. A weak meteorological forecast, for example, predicts rain in the Northwest, during November. At the opposite extreme, a strong prediction stipulates the precise amount to fall within a given time period, in a specific location (Meehl, 1978). Formal theory arguably makes for desirably strong prediction, including that located in psychological clinical science (Neufeld, et al, 2002). This is why formal theories die, while non-formal theories seemingly defy all laws of natural selection (Staddon, 1991)[3] .

Formal theories provide specific predictions, that can compete directly in terms of differential conformity to observed values. Predictions set in quantitative terms, in turn, expedite quantitative expressions of their relative superiority (e.g., tests for "goodness of fit", and "comparative goodness of fit"). In addition, however, models can prescribe their own qualitative signatures of validity. The latter entail the contour of data across gradations of independent variables, and their combinations (e.g., superadditivity of means, above; Neufeld & Williamson, 1996). Such patterns, moreover, can be mutually exclusive, in that even with very weak assumptions, competing models cannot produce each other's configurations (e.g., Wenger & Townsend, 2004, in press; cf. "empirical equivalence", e.g. Townsend, 1990).

If formal models mandate their own diagnostics, they necessarily ease the problem of empirical measurement. In fact, formal deductive systems stand the selection of measures on its head. As with longer-established sciences (e.g., physics examples, above), theory is potentially strong enough to mandate certain empirical ramifications. Displaced, therefore, is the often ad-hoc process of selecting off-the-shelf measures, conjecturally connected to dependent variables (McFall, Townsend & Viken, 1995; Meehl, 1978). Owing to precision of the formal deductive system, theory and measurement become intertwined. Assessment of the generated measures, for the usual statistical properties of fallible estimates (e.g., bias, maximum-likelihood status), to be sure remain applicable (e.g., Riefer & Batchelder, 1988; Townsend & Wenger, in submission). Nevertheless, they apply to indexes substantively tied to the theory at hand.

Especially befitting clinical science and assessment, individual differences in model expression can be accommodated within the context of model diagnostics. Included are variation in the structure of cognitive transactions (e.g., serial processing, variations on parallel processing, and selected combinations of these structures; Townsend & Fific, 2004). Variation in model properties, notably parameter values, across individuals can be profitably accommodated also through Bayesian techniques (Batchelder, 1998; Chechile, 1998; Neufeld, et al, 2002).

Not least among the endowments of formal theory, is its inherent challenge to preferred interpretations of addressed phenomena. Explicitness of expression helps to make salient current boundaries of knowledge-- flaws and limitations in extant formulations. In particular, empirical equivalence, on current measures, of nonetheless opposing positions are disclosed. However, so too are alternatives that can break the encountered experimental-mimicking logjams (Casti, 1989; Townsend, 1990).

Absorption of Theory-Exogenous Variance.

With formal logical-deductive systems, random empirical variance can be explained by the theory itself. Quantitative theory, in other words, encompasses apparent observational indeterminacy, versus relegating it to "error" or "noise variance" (cf. ANOVA linear model, e.g., Kirk, 1994; "local independence", Bollen, 2003).

An additional classification of theoretical systems, cogent apropos of this avowed asset, is that of Busemeyer and Townsend (1993; depicted in Table 1). Systems are categorized as dynamic versus static, and stochastic as opposed to deterministic. First, static stochastic models (lower left quadrant) predict frequencies of response values (e.g., categories, ratings, and so on), if not the times for producing them. An example is Tversky's Elimination-By-Aspects Model of preference and choice. (Tversky, 1972a,b; see Batsell, Polking, Cramer & Miller, 2003). Comparative frequencies of choosing items from those available are considered to follow a multinomial distribution (e.g., Evans, et al, 2000). Accordingly, probabilities of selection are the main focus. The probabilities are governed by parameters of the Elimination-By-Aspects Model, which are subjective utilities of the item attributes (e.g., likelihood of winning a lottery, versus its payoff; Rappaport, 1983; cf. Kukde & Neufeld, 1993; Morrison, Neufeld & Lefebvre, 1988, for applications in stress-and-coping research). Note that an earmark of stochastic models is their prediction of probabilities, or distributional frequencies of events, versus predictions of an all-or-none nature.

Turning to the upper-right quadrant, deterministic dynamic theory is exemplified most notably by nonlinear-dynamical systems models, popularly knows as "chaos theory" (e.g., Gregson & Pressing, 2000; Koçak 1989). To illustrate, variables such as psychological stress, coping and related variables can be depicted as being in constant flux, reciprocally influencing, and being influenced by one another. The system is represented by differential equations, each indicating the momentary change in a variable. The changes themselves are defined in these equations essentially by the variable's own extant state, and those of the other variables making up the network (Neufeld, 1999a).

Stochastic dynamical theories are exemplified by time-related extensions of decision-and-choice models. The above Elimination-By-Aspects model, for example, has been reformulated by Marley and colleagues (e.g., Marley, 1989; Marley & Colonius, 1992), so as to add temporal predictions about selection responses, to the original model's predicted frequencies of selection from among the competing items.

Busemeyer and Townsend (1993; also Roe, Busemeyer & Townsend, 2001) have presented an empirically compelling model of decision and choice, providing for speed of choice, type of choice, and their interconnections. Variability of expressed preferences, and their latencies, once again are accommodated, in that the crux of prediction involves frequency distributions.

Stochastic dynamical models represent possibly the most widely used version of formal theory in psychology. Generally speaking, they depict probability distributions of events, as a function of time (see, e.g., Johnson, Kotz & Balakrishnan, 1994; 1995; Luce, 1986; Townsend & Ashby, 1983). Part of the model fabric, then, is variability of event latencies across occasions of observation (e.g., experimental trials). The provision for variability is graphically illustrated by considering the model-prescribed variance of time taken to complete a relatively simple cognitive process.

The process comprises k' stages, each stage having a specified frequency distribution, as a function of time. Each stage is executed at the rate of v stages per unit of time (whereby the mean stage-wise completion is $1/v$ time units). Such a process may correspond to the encoding of k' stimulus features (e.g., curves, lines and intersections, of an alphanumeric item) into a cognitive format facilitating further processing, such as selected manipulations involving the item in short-term memory (e.g., Sternberg, 1975). By the present model, known as the Erlang Distribution (Evans, et al, 2000), the mean time across trials, for process completion is k'/v. The across-trial variance of completion, in turn, is k'/v^2. Note that this statement of dispersion is squarely embedded in the substantively-significant parameters of the theoretical model of process completion.

Further provision for stochastic properties are available by allowing for individual differences in the values of the parameters k' and v. This extension (a mixture model) "acknowledges" inter-participant variation in the above variance, and related model expressions (e.g., Neufeld, et al, 2002). Latency variance for a given value of k' and v maps onto the familiar within-participant variance in (M)ANOVA, and variation across participants in k' and v maps onto between-participant variance (see, e.g , Parzen, 1962, p. 200; also, *A Stochastic-Modeling Translation of the Statistical-Property Issue*, below).

Other contemporary exemplars of dynamical stochastic models include selected extensions of multidimensional scaling (see, e.g., Schiffman, Reynolds & Young, 1981). These extensions quantitatively relate categorizations, or ratings of multidimensional stimuli to their cognitive-process underpinnings, where the variable of time plays a central role (e.g., Carter & Neufeld, 1999; in submission; Nosofsky, 1992; Takane & Sergent, 1983).

Before leaving this subsection, it should be noted that nonlinear-systems theory in principle can be augmented to incorporate a stochastic component. A specifically random element can be added to the otherwise deterministic differential equations describing inter-variable influences across time (e.g., Brown & Holmes, 2001; Huber, Braun & Krieg, 1999). This interlaced stochastic thread perturbs the momentary change in each variable, and the latter's consequent influence on the coupled network of variables. Interestingly, the temporal trajectories of such a "stochasticized" system can deviate dramatically from those of its strictly deterministic counterpart.

Suffice it to say that overall, unaccounted-for variance in observed values may never be eliminated (cf. Gilden, 2001). However, such "model-exogenous variance" stands to be substantially diminished with increasing comprehensiveness of theory construction.

Aesthetic Appeal.

It was once declared by the Physics Nobel Laureate, P.A.M. Dirac, "I would rather have my equations be beautiful, than that they fit the experiment." Bolstered by their historical success, apparent failure of beautiful equations was held to be a temporary aberration. Perhaps, in their presented form, the equations were an inadequate approximation of empirical events. Or possibly, the experimental observations themselves were flawed. Either of these shortcomings could be redressed. Beautiful equations were intrinsically self-vindicating, and the rest was detail. Formal theory thus potentiates the scientific merit of equational elegance. There is no inherent reason why psychological clinical science should be barred from this asset, which arguably is a proprietorship of formal developments.

What, then makes an equation aesthetically appealing? One criterion is succinctness. The simplicity of beautiful formal expressions belies their profundity, which becomes apparent when their copious and far-reaching implications are unveiled (Polkinghorne, 2003). One example, emanating from the field of nonlinear dynamics, is Mandelbrot's formal representation of fractals. Although parsimonious, a feature of its constituent variables is that they take on a multitude of intriguing trajectories. The proliferation of exotic patterns has been widely disseminated, thanks to modern computer-graphic technology (e.g., Gleick, 1988; Townsend, 1994).

A more modest example is available from the Erlang Distribution of cognitive-process completion, described above. Here, the probability of completing the stated process, by a given time interval t is

$$\frac{(k'-1)!-\Gamma(k',vt)}{(k'-1)!} \tag{1}$$

Note that one can safely bypass the definition of each term in the equation and its equivalents, below, and still appreciate the pithiness of this expression. The somewhat involved operations, summarily contained above, include

$$1 - \int_{vt}^{\infty} x^{k'-1}\ e^{-x} dx/(k'-1)! = \sum_{j=k'}^{\infty} (vt)^j/j!\ e^{-vt}. \quad (2)$$

These formulae, in turn, can be used to create visual aids that facilitate insight into the interplay of model parameters (see Clark & Paivio, 1989, for a treatise on theory-related intuitions afforded by visual imagery). Three-dimensional response surfaces convey variation in the present probability across alternate combinations of *k*', *v* and the time window *t*. Two examples are provided in Figure 1, for values of *k*' ranging from 1 to 10, and values of *v* ranging from 0.01 to 3.0. The lower response surface is constructed for a value of *t* equal to 4, and the upper one is for a value of 7. This figure presents but two instances of a virtually infinite number of possible graphical portrayals of the formally-stated process. This is why, if a picture is worth a thousand words, a formula is worth a thousand pictures.

Liberating Qualities.

Formal logical-deductive systems are constrained by observationally- and quantitatively-principled initial assumptions, definitions, and axioms.[4] Building on this foundation, rigorous derivations render deductions about events of principle interest. In this way, precisely-framed empirical implications in effect can be evaluated for their potential importance, in advance of actual empirical testing.

Fruits of these labors, then, include the resulting insights, but also a potential increase in the overall economy of inquiry. Important questions can be posed before launching out empirically. Will the forecasted yield of information justify the associated investment; is that which may be found actually worth finding? Does the theory posit results which, if corroborated, could significantly impact on explanation, measurement and indeed future predictions?

The point to be made is that the theorist is emancipated to safely wade into the staked out substantive territory, with the above and related issues in hand, because her deliberations are constrained mathematically. These constraints represent an antidote to the hazards of armchair philosophy, and Psychology's ensuing reaction against exploration free of constant empirical vigilance (cf. instrumentalist versus realist distinctions in scientific theorizing; e.g., Casti, 1989).

As for the empirical enterprise, formal theory can increase the latitude of investigation in clinical science. It can do so by incorporating abnormalities associated with psychopathology into hypothesis formation, and testing. Such possibilities are exemplified in the cognitive-science application, below. In the enterprise of clinical science, certain hypotheses regarding pathognomonic disturbances may defy oft-used methodological options, and formal theory can supply essential alternatives.

To elaborate, one tack to testing hypotheses about cognitive debilities, for example, involves attempting to duplicate performance deviations in normals using experimental manipulations that are designed to mimic their etiology in patients. Formal theory can complement this strategy, and moreover pick up where it leaves off. The approach is to modify

quantitative models of performance in ways implied by the hypothesized agent of deviation, and then to evaluate accuracy of the resulting predictions (elaborated upon, below). Such methodological endowments come into play especially where tendered sources of disturbance are beyond the reach of direct manipulations. Selected agents of dysfunction, although model-hospitable, may be fundamentally alien to normal functioning, or demand unethical extremes of experimental induction.

Barriers: Attitudes, Aptitudes, and Investigator Role.

Ideally, formal theory at minimum should be a ready-to-hand arrow in an investigator's hypothesis-development quiver. It is recognized, however, that there are certain practical barriers to this ideal. One such barrier quite simply involves prevailing attitudes toward formal theory. Mathematical models are regarded with suspicion in some circles unless their implications are accompanied by readily-appreciated parallels. In clinical science, such parallels often take the form of behavioral neurological mechanisms that clearly align with the mathematical developments.[5] Interestingly, this position is counter to that of longer-established disciplines, where accepted mathematical necessity pre-dates identification of a corresponding physical mechanism (e.g., Newton's universal law of gravity, above). One price paid for suspended belief in quantitative formulations is the forfeiture of potential unification of co-existing explanatory systems, and new angles on synthesizing current evidence (e.g., Herrnstein, 1979).

Another practical impediment concerns formal theory's demands for hands-on tasking, by one closely involved with the subject matter. This requirement contrasts practices where assignments can be delegated to an associate or research assistant (as might be done, for example, in data entry, or say statistical analysis). Formal modeling furthermore, being customized to the immediate theoretical problem, is not a matter of output from routinely-applied software. This is not to say of course that advances in computational power, and computer-algebra programs, such as MAPLE and MATHEMATICA, do not substantially aid the theorist's job.

On balance, it can be said that such hurdles to formal theorizing may be daunting at first blush. Current practices, however, need not be abandoned while cultivating the formal alternative as an increasingly friendly ally. Eliminating the present hurdles can reap the substantial dividends on outlay in time and labor, enumerated above.

Application to Clinical Cognitive Science: The Case of Stimulus-Encoding Dysfunction in Schizophrenia

This section leads off with a presentation of the philosophical blueprint guiding the present application. Its exposition culminates in the appropriation of group-level inferences about cognitive functioning to the individual participant. This appropriation is mediated by the use of Bayesian-based procedures. Their promised value launches an exposition of the parent Bayesian framework dominating the current development. Positioning of stochastic models of cognitive performance within this framework is pinpointed.

Aided by the present instantiation, entailing selected stimulus-encoding dysfunction in schizophrenia, potential payoffs of the approach are expounded. They include: (a) precision in estimating performance-model parameters for the individual; (b), customization of the process-model's performance-latency distribution to the individual; (c), monitoring changes in cognitive functioning with the progression of treatment, at both the individual and group levels; (d) importing the informational infrastructure endowed by the model's cognitive-science roots; (e), complementing neuroanatomical regions of interest, with times of interest, in contributing to contemporary technology for measuring evoked neurophysiological responses; and, (f), facilitating links of cognitive-performance indexes with clinically-significant quantitative formulations on stress and coping.

Philosophical Orientation.

Associations among cognitive science, clinical cognitive science, and applications to individuals for clinical assessment, or other purposes, are depicted in Figure 2. Its arrows denote a smooth if not seamless transition among these investigative domains.

In the first instance, a model of cognitive performance addressed to the targeted territory of mentation is chosen. Deliberations include viability of the experimental paradigm, and relevance of its response parameters. Addressed cognitive faculties, for example, may be those identified with visual, memory and collateral processes (e.g., stimulus encoding, and response mechanisms); the paradigm may take the form of a conventional visual- or memory-search task (Sternberg, 1975; Townsend & Ashby, 1983); and response parameters may include performance latency and/or accuracy.[6] Several candidate models may present themselves. Selection from among contenders is based on congruence between the model features (e.g., its parameters,[1] or other properties) and concepts indigenous to the substantive research question (e.g., cognitive capacity, processing architecture-- serial, parallel, or hybrid--, efficiency of capacity allocation, and so on). Extant models may require elaboration or modification, befitting the currently driving issues. It may even be necessary to develop a model *ab initio* (see *Barriers ...* , above), depending on points of contact between the available armamentarium, and the clinical-science queries that are pressing.

The formal account of cognitive-processing is then adjusted so as to accommodate performance deviations of disordered individuals. Alterations can be directed toward the architectural structure of the modeled processing system, and/or toward model parameters. Architectural structure refers to the arrangement of constituent processes involved in task transaction. To illustrate, in ascertaining and registering the existence of a presented alphanumeric item in a memorized set of items, component processes composed of encoding, memory-search, and response operations, may be conducted in serial, in parallel, or in some combination of the two. Structure refers as well to the arrangement of constituent stages (subprocesses) of the said processes (e.g., in the case of encoding subprocesses, parallel versus serial input of the physical item features for template matching to those of the memorized item set; Townsend, 1984; Townsend & Ashby, 1983). Parameters refer to variables characterizing the processes, such as their size-- operationalized as the number of subprocesses making up the process-- and capacity,--operationalized as the speed with which these subprocesses are

transacted (parameters denoted above as *k*' and *v*, respectively).

For the sake of parsimony, the normal model is minimally perturbed, pending realization of the target configuration. Necessitated adjustments purportedly indicate disorder-affected functions, and those remaining intact signify functions that are spared. In general, processing architecture tends be preserved, however the same cannot be said for the values of model-parameters (Neufeld, et al, 2002).

It is apparent from the foregoing that models most closely aligned with the present purposes are those that are stochastic and dynamic (see *Absorption of Theory-Exogenous Variance*, above). By certain extensions of their stochastic element, beyond that attending response parameters *per se*, these models become poised to accommodate unique aspects of individual performance. Specifically, mixture models are fashioned to incorporate unequal values of performance-model parameters across units of observation-- in this case, members of diagnostic groups (Batchelder, 1998; Berger, 1985; cf. Busemeyer & Stout, 2002). The distribution of individual parameter-values in effect becomes a "prior distribution", in Bayesian statistical terms. Through the implementation of Bayes' Theorem (below), performance-model parameter estimates, as well as performance-latency distributions, stand to be individualized with considerable precision. Analogous to a medical-laboratory scenario, where a modest blood specimen is referred to the broader body of hematological knowledge, a modest cognitive-performance sample now is embellished with information conferred by the prior distribution of parameter values emanating from the participant's group.

Note that with the introduction of Bayesian-prior distributions of parameter values, deviation-accommodating adjustments no longer take place directly at the level of the task-performance process model. In the case of mixture models, there is instead movement of the deviant group's Bayesian-prior distribution of parameter values. In this way, effects of disorder are registered as features of the prior, and inasmuch as what has been altered is a *distribution*, provision for individual differences in expression of these effects is retained.

Arrows in Figure 2 are unidirectional, conveying the thrust of the current argument. There is nevertheless reciprocity among the above investigative domains. Successful capturing of deviant performance among clinical samples speaks to model robustness; those that readily accommodate performance extremes are preferred to those that are strained, or fail in this regard (Neufeld, 1982; Neufeld & Mothersill, 1980). This widening of application represents an unique opportunity for generalization testing (Busemeyer & Wang, 2000), and thus potentially reciprocates benefits back to basic cognitive science.

Nor are the layers of investigation mutually exclusive (see left-most column of Figure 2). Successful titration of stochastic models can lead to parallel manipulations at the connectionist level of analysis, whose own success or failure adjudges theory robustness (Carter & Neufeld, in submission; Roe, Busemeyer & Townsend, 2001).

Neurophysiological levels of investigation also stand to benefit. As described above, performance-model parameters (e.g., *k*', *v*) can be customized to individuals, by amalgamating each one's performance sample with the parameter-distribution prior, ascertained for the associated diagnostic group. With the individual estimates in hand, parametrically-homogeneous subgroups can be formed. Resultant specification of cognitive-process latency distributions (e.g., that of cognitively encoding a presented item into a task-facilitative format) in principle can be used to establish the likely epoch of occurrence of the process for the constructed group (e.g., Neufeld, et al, 2002, elaborated upon below). When it comes to measuring evoked neurophysiological responses (as in fMRI, and ERP), estimated "time windows of interest" thus may be prescribed, complementing neuroanatomical regions of interest, in calibrating the space-time co-ordinates of measurement.

Group and Individual Differences in the Encoding Process.

Bayes' Theorem in the current context. Recall that Bayes' theorem states

$$Pr\,(A|B) = [Pr\,(A)\,Pr(B|A)\,] \,/\, Pr(B)\,, \qquad (3)$$

where *A* and *B* are events; *Pr* (*A*) is the prior probability of *A*, or the probability of *A* preceding consideration of *B*; *Pr*(*B*|*A*) is the conditional probability of *B*, given *A*; and *Pr*(*B*) is the unconditional probability of *B*. As in the present context, *A* often is an hypothesized event, and *B* refers to experimental observation. This statement appears simple enough, until its wide-reaching ramifications are unraveled. The parsimony of expression that nevertheless belies its profundity obviously qualifies this theorem as one of statistics' aesthetically exquisite equations.

Perhaps a more familiar general expression of Bayes' Theorem is

$$Pr(A_{i'}\,|B) = [Pr(A_{i'}\,)\,Pr(B|A_{i'})] \,/\, [\textstyle\sum_i Pr(A_i)\,Pr(B|A_i)], \qquad (4)$$

where $A_{i'}$ is one of a set of hypothesized events A_i. In its present implementation, $A_{i'}$ becomes a parameter value of a test-performance process model, and *B* becomes a participant's performance sample comprising *N* empirical latencies $\{t_1, t_2, ..., t_N\}$[7], or {*}, for short. The latencies refer to stimulus encoding because cognitive translation of presented items into a format that facilitates collateral functions (e.g., search for a matching memory-held item) tends to be elongated in schizophrenia (Neufeld, et al, 2002).

The following formula adapts these equations to the present setting of application, and in so doing provides a comprehensive backdrop for explicating the potential assets enumerated at the beginning of this section.

$$w(\theta'|\{*\}) = \sum_{g=1}^{G} \left(\{\Pr\,(g|\{*\})\} \left[w(\theta'|\{*\}, g \right] \right). \qquad (5)$$

In this equation, $w(\theta'|\{*\})$ is the probability, or probability density (in the case of discrete and continuous parameters, respectively) of the parameter value θ', given the set of empirical latencies $\{*\}$; $Pr(g|\{*\})$ is the probability that the participant belongs to diagnostic group g, given $\{*\}$; and $w(\theta'|\{*\}, g)$ is the probability (density) of θ', given $\{*\}$ as well as the prior distribution of θ for group g.

As stated above, it is through variation in the prior distributions of θ that group-related performance deviations become registered in Equation (5). It is the presence of cognition-relevant parameters θ in this equation moreover that takes equations (3) and (4) beyond a theoretically-barren actuarial status to expressions imbued with substantive significance. The precise role played by the group-specific priors in Equation (5) is pinpointed in its expansion (presented in Appendix A). Note that the equation is partitioned by elongated braces $\{\}$ and elongated square brackets $[\,]$ according to the respective terms bearing on the main developments, below.

Experimental paradigm, and process of interest. What then corresponds to θ in the present application? The answer requires consideration of the cognitive task on which the model parameters bear, and whose performance supplies the latency sample $\{t_1, t_2, \ldots, t_N\}$. A variant of a memory-search paradigm (Sternberg, 1975) eliciting performance deviations among schizophrenic patients is described (Highgate-Maynard & Neufeld, 1986). The latter are representative of deviations occurring with other memory-search and related paradigms. Encoding of presented items is identified as the source of observed abnormalities in performance latencies, which compose the subject of prediction. Pertinent developments in quantitative cognitive science are exerted to discern responsible parameters. The theoretical design of the encoding process is delineated. Its model parameters are deemed to vary randomly across participants, with their distribution being perturbed according to group membership. The several potential benefits of this Bayesian layout, enumerated above, are then detailed.

The study by Highgate-Maynard & Neufeld (1986) extended typical memory-search methodology by incorporating methods emanating from Paivio's dual-coding theory (1971; 1986). Individuals with schizophrenia and controls indicated "as quickly and accurately as possible" whether of not the real-life size-- "overall volume"-- of a presented item (probe item comprising either an object or animal) was similar to that of a member of a previously-memorized set of items (memory set; "fixed-set procedure"). Memory-set size, ranging from 1 through 4, varied randomly across trials. Presence versus absence of the probe's size properties in a member of the memory set (positive versus negative trials), were similarly scrambled. Half the participants in each group were presented with similar-sized drawings of probe items, whereas the other half were presented with names. Encoding demands accordingly were deemed to be greater in the latter instance, because the initial verbal presentation of the probe would require referral to the imagery system for access to the demanded spatial-size properties (Paivio, 1986; see George & Neufeld, 1984 for examination of dual-coding and related predictions among schizophrenia participants).

Scanning memory-held items for the presence of the probe's size properties was considered to be comparatively taxing. It required comparison between the probe and memory-held items' overall volume, as set against subjective criteria of similarity in this property (Hockley & Murdock, 1987; Wright, 1977). Size attributes of items in the memory set were spaced such that the means of their normative-size ratings (Paivio, 1975) were at least two standard deviations of their Thurstonian discriminal-difference size dispersions apart (practically, at least two standard deviations of the distribution of difference scores, between their normative size ratings; Highgate & Neufeld, 1986). Responses were designated as correct or incorrect, based on whether or not they conformed to the following criteria for positive and negative trials. Negative trials, whose correct response was a "no" button press, meant that the probe item's size did not resemble that of a memory-set item. Here, the probe-item's mean real-life size rating was at least one standard deviation of the Turstonian discriminal-difference dispersion from the normative mean rating of each item in the memory set. Positive trials, whose correct response was a "yes" button press, were those where the mean of the probe item, and that of a memory-set member, to all intents and purposes were identical. Practice trials, ensuring familiarity with task requirements and the nature of correct responding, were similar in number for each group.

Further specifics, including those surrounding provision for potentially confounding clinical and demographic variables, and ascertainment of the viability of the paradigm for each diagnostic group of participants (e.g., applicability of normed item properties to each one), are detailed in the original article (Highgate & Neufeld, 1986), and summarized in Neufeld, et al, (2002). Considering the current emphasis on latencies, namely estimated encoding durations for correct responses, it should be emphasized that error rates were comparable across groups, and did not contribute to latencies in any confounding way. Finally, trimming from the examined latency data, those times estimated for processes residual to the one of current focus-- namely encoding-- is detailed in Neufeld, et al (2002; see also, Neufeld, Vollick & Highgate, 1993).

Specific processes thought to be tapped by this task, then, include cognitively translating the probe item into a format abetting comparisons with the memory set, or the memory-set's relevant properties; second, the comparisons proper; and third, response processes. Results from Highgate and Neufeld's (1986) study represent a case in point of delayed completion of the encoding process among schizophrenia participants, amidst integrity of remaining processes.

Individuals with schizophrenia taking part in this study included paranoid and nonparanoid subgroups. Comparison groups involved non-schizophrenic psychiatric controls, as well as non-patient controls. The present developments are expedited by focusing on the paranoid-schizophrenia-participants, and the non-patient controls. The clinical significance of this delineation has been described in Nicholson & Neufeld (1993); paranoid subgroups also have tended to be more elongated in probe-item encoding than nonparanoids.

Convergent support for this form of deficit, evinced on the present and related paradigms, has been presented in other sources (Neufeld, 1991; Neufeld et al, 2002; Neufeld & Williamson, 1996). Convergent support for the modeled composition of the deficit (elaborated upon below)

arises from investigations of divergent paradigms and forms of analysis. They include studies analyzing expressly the architectures and parameters of item encoding (Vollick, 1994; Vollick & Neufeld, in preparation); studies combining memory-trace theory and multidimensional scaling in assessing deviant similarity judgments (Carter & Neufeld, 1999); and neuroconnectionist assays of temporal and substantive aspects of item categorization (Carter & Neufeld, in submission). Finally, potential significance as a source of symptomatology, notably thought-form disorder (delusions and thematic hallucinations), has been discussed in other sources (Neufeld, et al, 1993; Neufeld & Williamson, 1996).

Parameter Estimation at the Diagnostic-Group Level of Analysis

Equation (5), above, is relatively complex (see expansion in Appendix A), but it comprehensively allows for the possibility that we may not know the group of origin for the individual supplying the current cognitive-performance sample. Thus, we first evaluate the probabilities of group membership, given the performance sample $\{ Pr(g|\{*\}) \}$, and then cipher the group-wise posterior probability (density) of θ', conditional on group membership $\lfloor w(\theta' | \{*\}, g) \rfloor$. This provision is brought into play, below, when we wish to use performance samples taken over time to profile an individual's progress, in terms of proximity to variously symptomatic groups (cf. Nicholson & Neufeld, 1993). It comes into play as well, when attempting to evaluate dynamical aspects of treatment efficacy, in terms of estimated temporal changes in group density.

Meanwhile, we may know perfectly well the group to which the present individual belongs. Our concern therefore is strictly with the expression $\lfloor w(\theta' | \{*\}, g) \rfloor$. We may isolate on the latter by setting $Pr(g) = 1.0$, where g is the relevant group. This constraint reduces the parent equation (5) to the target expression, and spotlights its associated computations (consider remnants left by cancellation of terms, in the expanded version of equation (5) presented in Appendix A, when $Pr(g)$ is set equal to 1.0)).

Process-latency model. The process-latency model has already been alluded to in discussing the nature of stochastic dynamical models (e.g., section on *Absorption of Model-Exogenous Noise*). It consists of the *Erlang distribution*, which is invoked to portray the stimulus-encoding process. The encoding process is highlighted because of the several processes putatively involved in task transaction *in toto*, it is one that is diagnostic of group-performance differences. Rationale for adoption of the Erlang and other distributions used here, and qualifying considerations surrounding the present applications, have been detailed in Neufeld, et al (2002).

The model postulates k' stages of encoding (e.g., constituent operations of covertly accessing size properties of the probe item). The stages proceed at the rate v per unit of time (seconds, in this instantiation). In the deterministic case, the process would invariably terminate after an interval of k'/v. Ambient trial-to-trial influences naturally infiltrate processing time, and perturb this pristine expression. These influences are summarized according to the probability density of process completion at time t:

$$f(t) = (vt)^{k'-1}/(k'-1)!\ ve^{-vt}, \qquad (6)$$

which renders a mean latency of k'/v, and a variance across trials of k'/v^2.

The values of parameters k' and v are not invariant, but are posited to vary across group members. Thus, the present mixture model provides for individual differences according to the group's Bayesian-prior distribution of parameter values, that randomly enter into the above process-latency model. Figure 3 displays the $f(t)$ for four combinations of parameter values, for four hypothetical inidividuals.

Parameter-value mixing distributions. It is now apparent that two parameters k' and v are involved in the process-latency model. Equation (5) may be readily extended to accommodate the present θ' as representing this pair of parameter values (Neufeld, et al, 2002). We consider the distribution of each, in turn. The parameter v is held to be gamma distributed, its gamma distribution having parameters r and k. Its probability density function, proportional to the relative frequency of its occurrence, is

$$(rv)^{k-1}/\Gamma(k)\ re^{-rv}. \qquad (7).$$

Note that Γ is the gamma function, which is a continuous-variable analogue to the factorial; where k is an integer $\Gamma(k)$ is $(k-1)!$.

The parameter r tenably expresses stress-induced impairment in processing speed. In contrast, k arguably conveys the speed-augmenting effects of task-related competence or performance skill (Neufeld, 1991a; 1994; 1996). As v formally conveys "processing capacity" (e.g., Townsend & Ashby, 1978; cf. Wenger & Townsend, 2000), its distribution is identified with the task-wise capacity pool, and r and k correspond to influences on this pool of stress and performer competence. Construct validity for the mixing-distribution parameter interpretations, particularly that of k, is examined under *Construct Validity of Distribution Parameters*: *Analytical Properties*, below. Introduced is a new format of construct validity endowed by the parameters' formal moorings (cf. Haynes, 2001).

Sufficient for the purposes at hand is the observation that the distribution of v parsimoniously is regarded as common across diagnostic groups and conditions of encoding load (supporting analyses are enumerated in Neufeld & Williamson, 1996; Vollick & Neufeld, in preparation). In the case of schizophrenia, therefore, processing capacity is spared; its distribution parameters remain unaltered. The estimates of r and k, obtained essentially through fitting theoretical to empirical moments (e.g., Townsend, 1984), were 0.03735 and 2.5044, respectively. Note that in order to circumvent certain interpretative hurdles apropos of the present purposes, incorrect responses were excluded, as were positive trials (Neufeld & Gardner, 1990; Neufeld, et al, 1993). The common distribution of v is presented in Figure 4.

Unlike the distribution of v, that of k' varies with diagnostic group and encoding-load condition. Specifically, this distribution is shifted upward in a similar fashion, with elevation the

encoding load, and with paranoid-schizophrenia diagnostic status. In this sense, these factors represent exogenous and endogenous sources of encoding-load increase, respectively.

The distribution of *k*' is considered to be Poisson (Evans, et al, 2000). The single parameter of this distribution is designated *m*. Its mean-- the modeled average value of *k*'-- is moved to the right by a constant amount for each source, as is its variance. This perturbation of the mixing distribution for *k*' therefore signifies an additive increase in encoding subprocesses attributable to each factor of the diagnostic-groups by encoding-load layout. The Poisson "intensity parameter" *m* is taken to indicate tendency toward encoding-subprocess recruitment upon engagement of the probe item.

The mean and variance of the Poisson distribution, then, both are *m*. Its base value for the non-patient controls, under the low-encoding load condition, is denoted *m*'. It is increased by the amount *h* for each group under the higher-encoding load, verbal-probe format, and by an amount *g* with paranoid-schizophrenia diagnostic status, under both higher and lower encoding demands.

Reference to the binomial-distribution's counterpart to *m* allows a more concrete understanding of this parameter. This counterpart, denoted *np*, approximates *m*, as follows. The quantity *n* indicates the pool of encoding subprocesses *potentially* transacting the encoding process for any participant. The term *p* in turn portrays the probability of a latent subprocess being marshaled to the encoding operation. The value of *n* is $m'+h+g+x$, where *x* is merely some very large number. That of *p* approximates m/n, where *m*'s base value, for controls under low-encoding demands is *m*', to which is added *h* and/or *g*, depending if exogenous (encoding-condition) and/or endogenous (paranoid-schizophrenia) related increases in propensity toward greater subprocesses are present.

The term *n* is tantamount to a dormant "subprocess pool", and *p* may be viewed as a "disinhibition" term. Clearly *m* increases with each term. The present interpretation designates *p* as the source of variation in *m* (Neufeld et al., 2002). In keeping with this stance, as *n* approaches the limiting value of infinity, and *np* approaches the limiting value of *m*, the Poisson and binomial distributions converge. It is the value $m \approx np$ that governs the probability of a specific value of *k*'. This probability is

$$Pr(k';m) = m^{k}/k'!\, e^{-m}. \qquad (8)$$

The estimated value of *m*' was .00001, that of *h* was 19.7390, and the estimated value of *g* was 70.0. The value of *m*' suggests that the controls required hardly any processing to encode the rudimentary limned drawings of the probe items. Plots (smoothed) of the $Pr(k';m)$ for the present values of *m* are plotted in Figure 5.

The general structure of the mixture-distribution model is summarized in Figure 6. The gamma and Poisson mixing distributions of *v* and *k*', respectively, correspond to those displayed in Figures 4 and 5. The Erlang distribution of *t,* on which *v* and *k*' bear in the individual case

(exemplified in Figure 3), is the mixture's "base distribution". With this structure in hand, estimation of the base distribution's parameters for an individual in principle can be substantially sharpened; all that is needed is a modest sample of the participant's own process-latencies {*}.

Individualized Parameter Estimation.

With appropriate application of $\lfloor w(\theta')|\ \{*\},g \rceil$ to the case of {*v*, *k*'}, estimation of these parameters stands to become substantially more precise (computational formulae are presented in Neufeld, et al, 2002). To illustrate the point, representative performance samples from each diagnostic group, under each condition of encoding load, are brought into play. In each instance, a sample was extracted from one of the 10 participants in each of the four factorial combinations (2 diagnostic groups by 2 encoding-load conditions). Each sample size was 4, one for each memory-set size, under negative trials. Adjusting these latencies to highlight the encoding process has been described in Neufeld, et al, (2002). The sets of trimmed values are presented near the top of Table 2.

A graphic example of tailoring a parameter distribution to the individual is exemplified for a tentative case. Specifically, consider the performance sample supplied by the Control participant under the low-encoding condition {1.037, ...}. Although unlikely according to Table 3, this sample still could have been produced by a member of the paranoid-schizophrenia group, under the low-encoding load. If so, the customized distribution of *k*' would have moved markedly to the left, relative to the prior distribution aligned with the paranoid-schizophrenia–low-encoding combination. This relocation reflects the joint probability of comparatively lower latencies making up this encoding-process performance sample. The Bayesian posterior distribution of *k*' (smoothed) is presented in Figure 7, alongside the prior distribution corresponding to $m = 70.00001$. Note as well the increased precision of estimation, indicated by the diminished width of the posterior distribution. In fact, the standard deviation of this posterior distribution of *k*' is 1.8110, whereas that of the prior is 8.36. Similar observations attend the individualized estimation of *v*.

Individualized Latency Distributions.

These slenderized posterior parameter-mixing distributions can be used to create posterior latency distributions $f(t|\{*\}; g)$. The customizing effects of a performance sample can be appreciated by contrasting $f(t|\{*\}; g)$ with those for the alternate mixture distributions $f(t|g)$, sans reference to the performance sample {*}. The selected performance sample is that for the paranoid-schizophrenia low-encoding condition {2.177, ... }. Figures 8 and 9 show the posterior latency distribution constrained by {*}, and Figures 10 and 11 give their unconstrained mixture-distribution counterparts. Observe that in Figure 10, the distribution for $m = 0.00001$ essentially is not visible, because the stochastic mechanisms involved result in its hugging the *y* axis.

Posterior Probability of Group Membership.

We now consider the case where a performance sample has been obtained from an individual, and we wish to use it to evaluate the person's "proximity" to variously symptomatic groups. This scenario may apply to treatment evaluation over time: "Is the individual being edged closer to normal functioning, as assessed by methods rooted in contemporary cognitive

science?"

Estimates are available in terms of the posterior probability of a representative of the target group– a group with less or no symptom severity– generating the sample, in contrast to a member of a symptomatic or more severely symptomatic group. In effect, what is being asked is, "In Bayesian terms, what are the compatibilities of the obtained process-latency specimen with each group's pair of prior parameter-distributions?" The expression of equation (5) enclosed in elongated braces $\{\!\{$ now is brought forth.

Allowing that the performance is sampled at selected points in time, it is assumed that the representativeness of each symptom group along the way is known. In other words, transitions in prevalence of the candidate groups ostensibly are monitored, so that base-rates for a testee's group membership $Pr(g)$, $g = 1,2, \ldots G,$ are available. It is assumed, as well, that the mixing distributions for these groups, established at the outset, remain valid; or if not, that new priors have been formulated for the ensuing points of performance sampling. In the present application, the assumption of initial-prior stability is invoked.

Finally, the present computations imply that the G classes are mutually exclusive and exhaustive. The classificatory system is therefore regarded as comprehensive. This feature underwrites the intended description of the presenting individual, which in fact is a finite, discrete posterior-probability mixture. The adopted approach entails an "odds-ratio evaluation strategy" -- an odds-ratio evaluation strategy for assessing the appropriateness of the G respective priors. It is deemed to facilitate the flow of the present developments; other Bayesian strategies, with unique statistical merits, nevertheless are available (see, e.g., Karabatsos, 2006).

In contrast to the scenario painted above, it is possible that diagnostic-group base rates indeed are not available for the designated time points. Instead of monitoring an individual's progress, by profiling his/her likelihoods of belonging to symptom-groups of known base rates, the primary issue may entail the base rates themselves. Here, the focal problem may be one of evaluating the collective effectiveness of treatment, by ascertaining movement in size of symptom-group membership. A question of this nature once more can be broached through the elongated-braced expression $\{\!\{$ of equation (5). This second case is taken up in its turn, below.

To expedite the exposition, the present classifications formed by factorially combining diagnostic status and encoding-load condition are appropriated. These combinations again provide the Bayesian mixing-distribution priors of k' and v, as well as the representative performance samples listed in Table 2. We conveniently allow for a progressive decrease in symptomatology with movement across groups formed from combinations of paranoid-schizophrenia status/ high-encoding demands, to paranoid-schizophrenia/ low demands, through non-patient status/high demands, finally to non-patient status/low demands.

Apropos of the encoding manipulation as a surrogate for symptom severity, to be sure responding to increased encoding demands with an increased number of encoding stages hardly is indicative of disorder. It nevertheless conveys an escalation in recruited subprocesses, as a

stand-in for the elevated number that may well go along with increased symptomatology (Nicholson & Neufeld, 1993). In all events, this pretense for expedience is inconsequential to the ensuing line of argument, and in no way detracts from its analytical points.

We posit a scenario where encoding-latency samples have been obtained from 4 participants, as arrayed in the first row of data in Table 2. The values have been obtained say at a time of interest specified by the treatment regimen. We wish to estimate probabilities of symptom-group membership for each participant, given the obtained performance data. These group-wise posterior probabilities are endowed by the prior parameter-mixing distributions aligned with the candidate groups.

To proceed, it is necessary to know the values of the unconditional joint probability-density of each latency sample with respect to each group $ucd(\{*\}|g)$. The $ucd(\{*\}|g)$ is the joint density of the obtained sample of latencies, given group g, all values of v and k' considered. The numbers assumed by $ucd(\{*\}|g)$ will change with g because the parameter mixing-distributions are group specific. The values of $ujd(\{*\}|g)$ for the current performance samples are presented in Table 2. Base rates assigned to the respective groups $Pr(g)$, proceeding from left to right of Table 2 (Paranoid-schizophrenia/High-encoding load to non-patient Control/Low-encoding load) are 0.05, 0.20, 0.5, and 0.25.

The posterior probabilities of group membership are entered in Table 3. Each individual is profiled in terms of distance from the respective groups, enumerated as the Bayesian posterior probability of group membership $Pr(g|\{*\})$. Not surprisingly, the highest values tend to be aligned with the groups from which the samples originated. The sole exception occurs for the sample from the Paranoid-schizophrenia/ High-encoding load combination. This result is attributable to both $ucd(\{*\}|g)$ and $Pr(g)$ being highest in connection with the Control/High-encoding load amalgam.

The point of the exercise, however, mandates that the clinical scientist/practitioner is blind as to the actual origins of the latency samples. Table entries then can provide pivotal information regarding participant status, as conveyed by the Bayesian expansion of the present mixture model. Allowing that all individuals were members of the same group at the outset, clearly the far-right column presents the most encouraging profile, and that second from the left, the least so.

As intimated above, such multivariate profiles can be charted across successive times of interest. The resulting individual configurations of treatment (non)response, at least as quantified by focal cognitive functioning, in turn may be statistically mediated to psychometric and other predictors.

Maximum-Likelihood Estimation of Symptom-Group Base Rates.

In the first case, above, base-rates for the variously symptomatic groups were known, and the desired information consisted of the alternate likelihoods of the individual's membership, given his/her specimen of process latencies. In the present case, we entertain a sample of

individuals who have been allocated to one or another group on the basis of presenting symptomatology at the selected time of interest. However, unlike the first case, the base-rates now are unknown, and it is this information that we seek. These densities of enclaves with differing symptomatology may be useful in assessing treatment efficacy across the designated times of interest.

It may be argued that the desired course of action is simply to directly examine the array of diagnostic assignments that have been made at the selected times. It may be prohibitive, however, to diagnostically interview (or assess by some other means, such as psychometric) a large enough sample to ensure reasonable fidelity of estimates (see, e.g., Batchelder, 1998; Luke & Homan, 1998). If so, it seems desirable to bring to bear our available bank of information on cognitive functioning. In importing this information, we may settle not only for a modest sample of individuals whose classifications are updated , but also a modest performance sample from each. The required sampling may feasibly be tacked onto sessions hosting symptom assessment at the successive times of interest.

The strategy for estimation is straightforward enough. It consists of constructing a likelihood function for the updated participant assignments, and then maximizing the function with respect to $Pr(g)$, $g = 1,2, \ldots G$. The likelihood function is made up of the joint probability of the updated participant categorizations. This joint probability comes down to the product of the posterior probabilities of those categorizations, given the latency samples {*} linked to the categorized individuals. The procedure comprises a somewhat unusual but valid combination of maximum-likelihood estimation of a meta-Bayesian parameter (see Appendix A). The likelihood function, and related computational details of maximum-likelihood estimation of $Pr(g)$, $g = 1,2, \ldots G$, are presented in technical Appendix B.

The instantiation of this application once again makes use of entries in Table 2. For reasons of mathematical tractability, and in the interests of illustrative neatness, only 3 of the 4 data sets listed in the this Table are considered. The set corresponding to the non-patient Controls/High-encoding load is excluded, deleting the third row and third column of $ujd(\{*\}|g)$ values in the Table. However, in practice, such a restriction is not necessary, as described momentarily.

The maximum-likelihood estimates of $Pr(g)$ were obtained by solving for each value using closed-form solutions to two simultaneous equations (two, rather than three, as the base-rate probabilities sum to 1.0; see Appendix B). It was found that increasing G beyond 3 became infeasible, even with the assistance of a computer-algebra program. In larger-scale applications, however, where $G > 3$ and/or the number of diagnostically-updated participants is somewhat larger, one may resort to a numerical-search algorithm (e.g., MATLAB Optimization Toolbox), rather than pristine closed-form solutions, to estimate the respective values of $Pr(g)$.

For the present purposes, the 3 sets of performance samples in the provisionally reduced Table 2, now represent those putatively obtained at the period of updated individual classification. It is sufficient moreover to allow each individual to have been assigned to the

group adjacent to the displayed performance sample. In other words, three individuals have been classified, one to each group. The classified individuals furthermore have putatively supplied the very process-latency samples atop the columns of their respective groups.

With these arrangements in place, base-rate estimates of the three groups, proceeding from left to right, become 0.6748, 0.0852, and 0.2399. The most symptomatic category thus remain in the majority at this time of evaluation, and those who are moderately symptomatic are sparsest. The least symptomatic category nevertheless is reasonably well represented, with an estimated relative frequency of nearly 0.25.

It is tempting to insert these base rates into the expression enclosed by $\{\}$ in equation (5), and in turn to evaluate the posterior probabilities of membership in the respective groups, given the associated performance samples {*}– as done in the first case, above. Note, however, this exercise is nothing more than instructive, regarding the involvement of $ujd(\{*\}|g)$ and $Pr(g)$ in arriving at the values of $Pr(g|\{*\})$. The latter, again proceeding from left to right in Table 2, are 0.119786, 0.119736, and 1.00. The first amount resembles the second, because the comparatively low $ujd(\{*\}|g)$ connected with the Paranoid-schizophrenia/High-encoding group for the individual so assigned, is offset by that group's having the highest value for $Pr(g)$. The value of 1.00 for the non-patient Control/Low-encoding load assignment, evinces the fact that $ujd(\{*\}|g)$ associated with this specific group exceeds the next highest value by a factor of $.2766325061(10^{21})$.

Qualifying assumptions. The above estimates of $Pr(g)$ assume that the parameter-mixing distributions initially prescribed for the symptom-based classifications extend to the time of estimation. This assumption is intrinsically bound up with stability of conditions under which the current sample of task performance is obtained. The assumption also is yoked to continuation of the initial classificatory criteria leading to participant allocation. These requirements come down to retention of their earlier values, by parameters of the mixing distributions r, k, and m; continuing tenability of the distributions themselves; and transference of the previous meaning of g, in the presently estimated $Pr(g)$.

Dislodgement of the established priors from their host symptom groups may occur because of a shift in the nature of task performance, owing for instance, to practice effects. Potential redress may require the substitution of different task items, analogous to using of parallel forms of psychometric tests. To guard steadfastness of the G categories, symptom criteria for designations by and large need to be consistent.

It is additionally assumed that differential dropout across the G groups, or other sources of distortion in their representativeness are prevented. Finally, the computations once again imply that the G classifications are mutually exclusive and exhaustive. In other words, transitions in symptomatology eventuate in one or the other of the G classes addressed at the outset.

It is evident from these developments that tenability of the proposed analytical strategies makes for new and potentially useful avenues of formulating clinically significant information. Available are inferences about the status of specific patients, as conveyed by cognitive-functioning aspects of symptomatology. Available as well are broader inferences about treatment efficacy. Furthermore, assumptional constraints are relatively salient. Lastly, the developments remain entrenched in a formal cognitive-process model.

When events are processes: Potential contributions to functional MRI.

Recent years have seen an explosion of advances in fMRI technology. (e.g., Frackowiak, Friston, Frith, Dolan, Price, Zeki, Ashburner & Penny, 2004; Huettel, Song & McCarthy, 2004; Reiman, Lane, Van Petten & Bendettini, 2000). Clinical cognitive scientists have capitalized on these successes, to map deviations in neuronal-activation patterns corresponding to symptom-significant mentation (e.g., Boksman, Théberge, Williamson, Drost, Malla, Densmore, Takhar, Pavlovsky, Menon,& Neufeld, 2005; Lanius , Williamson, Densmore, Boksman, Neufeld, Gati, & Menon, 2004). Note that the ultimate subject of interest, in documenting neuronal-activation concomitants, is not that of external-stimulus events per se. Rather, it comprises the cognitive processes that are instigated by these events, and that parallel the activation profiles that are thus set in motion (cf. Friston, Harrison & Penny, 2003).

Considerable dynamical modeling of the blood-oxygen-level-dependent (BOLD) response, the kernel of MRI measurement, has been put forth (e.g., Aguirre, Zarahn & D'Esposito, 1998; Boynton, Engel, Glover & Heeger, 1996). Similarly impressive advances in statistical treatment of fMRI signals have taken place. Quantitative cognitive science meanwhile is poised to complement these developments. Implicated are stochastic dynamical models, as follows.

Consider the stochastic dynamical trajectories presented in Figure 9. The areas under these curves are proportional to relative frequencies of process completion. They can be integrated into posterior cumulative probability functions $F(t|\{*\})$ whose values for the respective curves in Figure 9 are shown in Figure 12.

The amounts displayed in Figure 12 in principle may be consulted to synchronize times of interest in BOLD-response recording, on the one hand, with probability of target-process occurrence, on the other. In moving from right to left in Figure 12, for example, probability of process expiry decreases, however, the time window for measurement does as well. The interplay of such considerations now has a formal backdrop, against which to estimate times of interest, complementing regions of interest, in calibrating of MRI measurement. Moreover, imputation of cognitive functions now stands to be anchored in a verisimilitudonous performance model, avoiding the oft-encountered circular reference to the very neurocircuitry whose functional significance is under study.

The trajectories in these figures have been constructed for an individual performance-sample--prior-distribution combination. They could, however, be extended to accommodate subsets of individuals whose sets of performance samples are similar, and to whom the same

priors apply. The resulting amalgams should render subsets with relatively homogeneous parameter profiles (see Figure 7), along with associated latency trajectories. Aggregation across participants often is necessary to attenuate noise of MRI signals, or for selected statistical treatment. Amalgamation of parametrically homogeneous participants guards against an amalgam of systematic individual differences in the expression of cognitive functions under study.

Multiple processes. Typically, even the simplest cognitive task recruits multiple stochastic processes (Smith, 1995; Townsend, 1984). Separability of the respective processes then comes into play. The processes may occur in parallel, meaning they commence simultaneously. Their finishing times nevertheless are staggered, according to their respective stochastic distributions (Townsend & Nozawa, 1995; Wenger & Townsend, 2000).

Let the probability-distribution function for one of a pair of processes be $F_1(t)$, and that for the second be $F_2(t)$. It is stated here without elaboration that the greatest likelihood of the second process surviving, with the first having elapsed, is defined by $max\big([1 - F_2(t)] - [1 - F_1(t)]\big) = max\big(F_1(t) - F_2(t)\big)$. Measurement epochs addressed to the second process would be synchronized accordingly.

The processes instead may proceed in serial. Let the probability-density function of the initial one be $f_1(t)$. Then the greatest likelihood that it has been accomplished, while the second has not yet elapsed, occurs at the modal value of $f_1(t)$, or $max\big(f_1(t)\big)$.

Cognitive Debility and Stress Susceptibility.

Adaptive significance of cognitive functions may be appreciated by deciphering how processing operations interface with environmental exigencies. Consideration of cognitive psychopathology invites the following extension: examination of quantitatively dissected cognitive debilities in relation to likewise quantified environmental pressures on which they bear.

As applied to stress negotiation, the above approach implies, among other analyses, delineating information-processing requirements for selecting advantageous, threat-minimizing options when stressor situations are multifaceted. This endeavor is potentiated through the application of mathematical combinatorics, and auxiliary procedures, to environmental prototypes. Products of the analysis include the specification of, (1), cognitive costs, in the form of information-processing demands; (2), stress-reducing benefits, in the form of minimization of threat (e.g., physical danger/discomfort; social sanction); and, (3), sources of vulnerability stemming from information-processing frailties. The synthesis should ultimately provide a more precise picture of compromised negotiation of environmental demands, harboring implications for prediction and intervention.

A prominent form of coping with stress, alluded to above, is labeled *decisional control* (Averill, 1972; cf. Lees & Neufeld, 1999). This type of coping basically takes the form of situating oneself in a multifaceted stressing context, so as to engage the situation's most

innocuous option. Dissection of decisional control, through the application of mathematical combinatorics and related methods to the essential structure of this form of coping, discloses it to be a prototypical expression of selection and choice (Morrison, Neufeld & Lefebvre, 1988; Neufeld, 1999a). Unveiled, in turn, are its demands on cognition, including option-cue encoding, visual and memory search, and response processes. Unveiled alongside is the undermining potential of debilities in the required cognitive functions, including those comprising retarded completion of encoding processes (Nicholson & Neufeld, 1992; Neufeld, 1991; 1999b).

A factor potentially exacerbating the above risk, is the toll taken on encoding speed by stress itself. Apropos of the formal depiction of encoding, this effect translates into a general reduction in the parameter *v* of the base distribution of latencies (Figure 3). Recall that a stress-induced movement of the distribution of *v* to the left is identified with an increase in its mixing-distribution parameter *r* (Figure 4; Neufeld, 1994; Neufeld & Carter 2000).

Construct Validity of Distribution Parameters: Analytical Considerations

Construct validity is a form of psychometric validity entailing a corpus of evidence brought to bear on the espoused interpretation of a measure. Haynes and O'Brien(2000) summarily have stated that it "Comprises the evidence and rationales indicating the degree to which data from an assessment instrument measures the targeted construct; includes all evidence bearing on the measure, and includes all types of validity (Haynes, p.75)." In line with this definition, construct validity for the purported psychological meaning of scores from a test, inventory or other mode of assessment emphasizes the measurement tool's portfolio of empirical correlates (Campbell & Fiske, 1959; Cronbach & Meehl, 1955; Embretson, 1983; for a mathematical critique of selected data-analytic strategies for adducing psychometric construct validity, see Neufeld & Gardner, 1990).

An overriding principle governing evidence for a measure's construct validity is that the measure act in accordance with theory (Wiggins, 1973, p. 406). In the present treatment, "measures" consist of model parameters. Construct validity for their imputed meaning, in turn, emanates from their mathematical properties, which are part an parcel of the stochastic dynamical models in which they participate. The light thrown on the nature of parameters by the mathematical properties they possess arguably embodies a new form of construct validity (cf. Cronbach & Meehl, 1955).

The parameters of the appropriated process-latency model included *v* and *k*'. Their interpretations stemmed from the assumed structure of the encoding process, as translated into a corresponding distribution of latencies. Distributions of the above parameters in turn have their own parameters, *r* and *k* in the case of a gamma-distributed *v*,[8] and *m* in the case of a Poisson-distributed *k*'. Because they pertain to the mixing distributions for *v* and *k*', the parameters *r*, *k*, and *m* are a step removed from the empirical touchstones of an individual's latency data. Accordingly, called for is support for the constructs they purportedly express, arguably from a commensurate level of analysis.

Recall that the parameter *m* was aligned with "encoding load", and conveyed exogenous (task demands) and endogenous (paranoid-schizophrenia) sources of increase. Its binomial-based dissection cast a certain light onto the potential composition of this parameter (cf. "construct representation" Embretson, 1983). The parameter *r* was identified with stress effects on processing speed. Its increase moves the stage-dispatching capacity parameter *v* to the left. Finally, the parameter *k* ostensibly corresponds to process-transaction competence, and its increase shifts the distribution of *v* to the right.

The parameter *k* is endowed with additional mathematical properties that effectively illustrate the type of construct validity imbued by analytical derivations. The behavioral significance of these properties hinges on the concept of maintenance of cognitive performance, over and against its breakdown, as follows. Within the present analytical context, integrity, versus collapsing of performance requires that the value of *k* exceed a certain threshold, paralleling a critical level of performer skill. In contrast, increasing *r* and *m* prolong performance times, but unlike the case for *k*, actual failure of the processing system is not precipitated when a finite range of values is violated

The current formal expression of performance breakdown takes the form of certain latency-distribution moments acquiring infinite values: more succinctly, $E(T^n) = \infty$, where *n* is the order of the computed moment. Behaviorally, the result signifies that the population of trials occurring under the prevailing value of *k* includes a critical mass that are extremely long, or essentially outstanding, non-completions. The integral

$$\int_0^\infty f(t)t^n dt = E(T^n)$$

consequently does not converge (see also Appendix C.a for further technical exposition).

The first four moments of a stochastic distribution ($n = 1, 2, 3, 4$) are somewhat familiar, comprising or being involved in its mean, variance, skewness and kurtosis. In model evaluation, empirical moments beyond the second seldom are sampled, because of instability. The present theoretical deliberations, however, operate in the realm of population properties, bypassing issues of sampling and measurement weaknesses. The order of the moment *n* therefore can be driven to whatever amount is prescribed by the governing theoretical analysis, which in this case concerns consequences of variation in the processing system's constituent parameters. At the same time, such theoretical explorations retain a bona fide bearing on the interpretation of the parameter values, as fallibly estimated from empirical data. Observe that this tack to discerning the nature of a theoretical system's parameters has much precedent in fields ranging from mathematical ecology, and economics, to astrophysics.

The above integral indicates how values of *t* and the order of the moment *n* act in concert to affect $E(T^n)$. As *n* increases, its exponentiation of *t* correspondingly magnifies the consequences of elongated latencies for the computed moment, including its transition to infinity (Appendix C). Apropos of the model parameter *k*, strengthening process-performance skill

should influence the distribution of t downward, resulting in $E(T^n)$ to an ever greater extent being bounded by finite values. That is to say, it should enable $E(T^n)$ to withstand heightened values of n, before transmogrifying to infinity.

In effect, k acts to diminish densities $f(t)$ for higher latencies, including extremes, through its shifting the distribution of the process-completion rates v upward. Indeed, the threshold value of k issuing in bounded moments is defined precisely in terms of n (below, and Appendix C). Insufficient k implies that appreciable densities of v, $f(v)$, are perilously close to $v = 0$, generating the critical magnitudes of $f(t)$ at acutely high ranges.

The above support for the construct validity of k as a capacity-affecting competence parameter is robust, as assayed from multiple analytical standpoints. It transcends process-latency base distributions with diverse characteristics; its interpretation is enhanced by pertinent Bayesian extensions; its effects are modified in expected ways by formally defined efficiency of capacity-application by the operative base distribution; its effects align with those of related parameters from similar mixing distributions; and it is evinced in familiar base-distribution statistical properties. Finally, its effects are distinguishable from those of parameters to which a different meaning is ascribed, notably r, the avowed index of stress effects on processing capacity. These variations on analytical support are dealt with in succession. Their largely verbal descriptions are presented below, with technical justifications allocated to the several subsections of Appendix C.

Critical Values of k: Convergence from Multiple Base Distributions.

Finite moments of mixture-model latencies require that k exceed the order of the moment n for diverse process-latency base distributions. The current selection of base distributions are prominent in the stochastic-modeling literature, and express alternate structures of a processing system. Their descriptions are available in various sources (e.g., Evans, et al, 2000; Johnston, Kotz & Balakrishnan, 1994; 1995; Ross, 1996); presentation here is restricted to theoretical analysis of the espoused substantive significance of k. Note that the present treatment, although theoretical, is malleable with respect to empirical application. Almost any empirical process-latency distribution can be well approximated by creating a finite probability mixture of one or more base distributions (defensibly expressing the relative percentage of trials best characterized by the respective base distributions; e.g., Townsend & Fific, 2004). Importantly, results of the present analyses regarding moment finiteness are preserved with such a mixture.

The above relation between k and n occurs for the Erlang distribution (Figure 3), with a fixed stage parameter k'. A specific case is the single-stage exponential distribution, where $k' = 1.0$. It includes, as well the extension to k' as a Poisson-distributed random parameter, otherwise known as the compound Poisson distribution (e.g., Feller, 1966; Ross, 1996). Together with the mixture on v, the latter distribution makes up the dual-parameter mixture model depicted in Figure 6. In the present context of base distributions, whose parameter v is gamma-distributed, the compound Poisson distribution is considered simply to be a base distribution whose Erlang stage parameter k' is Poisson pre-mixed.

The rate parameter v is constant for each stage in both the Erlang and compound Poisson distributions. The above relation between k and n, however, transcends base distributions with a static rate parameter. It extends to those with rate transitions across stages, known as the general gamma distribution (McGill & Gibbon, 1965). A notable version of the general gamma is a "pure death system with a linear death rate", which is a parametric version of an independent parallel-processing model, with unlimited capacity (see Townsend, 1990; Wenger & Townsend, 2000). This model has figured prominently in selected theorizing about the architecture of human information-processing mechanisms (see, e.g., Townsend & Ashby, 1983; pp. 80-95). Another well-known and used distribution to which the present relation applies is one with a single stage of process completion, but whose rate continuously changes with time (Weibull distribution).

The selection of base distributions, and their distinguishing properties, to which the relation between k and n applies is summarized in Table 4 (further details are briefly presented in Appendix C.b, additional elaboration being available in Neufeld, 1994). Also listed are the moments for $k > n$ and $k = n$. Apropos of the latter, note that if the nth-order moment is unavailable (infinite), as in the case of $n = k$, so are moments of order exceeding n (see, e.g., Harris, 1966, p. 102 for k as an integer; where k is not an integer, and $k < n$, the function $\Gamma(k - n)$ appearing in the moment's expression takes on a succession of negative, infinite and positive values as k-n decreases from 0, whereby moments about the origin become meaningless).

At the heart of the matter is the attestation to k as a process-transacting competence parameter, according to invariance of its infinity-avoiding values across base distributions of heterogeneous makeup. Robustness of this nature bears a certain resemblance to empirical convergent validity, which entails agreement among divergent methods of measuring the same trait (Haynes, 2001).

Bayesian Extension.

Additional light can be thrown on the proffered interpretation of k by dissecting its properties within a Bayesian context. Insight is afforded according to k's give-and-take with empirical performance samples, and with collateral parameters. A poignant exemplar involves the $\text{Erlang}_{r,k}$ distribution whose rate parameter v is gamma mixed. The current Bayesian

extension once more incorporates a set of process latencies $\{t_1, t_2, ..., t_N) = \{*\}$, and the nature of latency-distribution moments again occupies center stage.

Fixed-parameter status of k' streamlines the present exploration without curtailing generality of available inferences. Findings tenably apply to classes of individuals sufficiently homogeneous regarding k' that this simplification is a reasonable approximation of the quantitative goings on. The required moment expression, so conditionalized on k', can be immediately extracted from derivation (A.10) of Neufeld, et al (2002). It appears as follows

$$E(T^n|\{*\},k') = \frac{[r+\sum_{i=1}^{N} t_i]^n \ \Gamma(Nk'+k-n)(k'+n-1)!}{\Gamma(Nk'+k)(k'-1)!} \qquad (9)$$

This expression affords a certain perspective on the posited interpretation of k from the standpoint of rudimentary classical psychometric theory surrounding performance-sample size, or test length. The excursion into the psychometric analogy is followed once more by observations on viability of the proposed interpretation of k.

Isolating first on the term $\Gamma(Nk' + k - n)$, an infinite value of $E(T^n|\{*\};k')$ is avoided if $Nk' + k$ exceeds n. In this case, it is possible to have $k \leq n$ without $E(T^n|\{*\};k) = \infty$ if, and only if $Nk' > n-k$, allowing of course that the present $t_i < \infty$, $i = 1,2, ... N$. A value of $k \leq n$ signifies that a finite $E(T^n|\{*\};k')$ does not characterize all members of the class of individuals to whom the current value of k applies. The specific member at hand nevertheless is characterized by a latency moment bounded by finite values.

The psychometric analogy seemingly throws light on this conclusion, as follows. Within this analogy, the stage parameter k' corresponds to the number of parallel items composing a psychometric test. Second, N corresponds to the number of independent test administrations. The collective item sample therefore is Nk'. With a sufficiently large empirical sample of finite latencies, combined with the extant value of k, a finite value of $E(T^n|\{*\};k')$ is assured for the testee at hand.

The argument can be carried further as follows. Consider that the Bayesian posterior moment $E(T^n|\{*\};k')$ in principle can be partitioned according to a probability mixture of two partial expectations, $\pi(E(T^n|\{*\};k') = \infty)$ and $(1-\pi)(E(T^n|\{*\};k') < \infty)$, where π is the mixing parameter, $0 =< \pi =< 1.0$. A sufficiently large performance sample of finite process-completion latencies, in concert with k, in effect renders π equal to 0. Conversely, a value of Nk', such that $Nk' + k \leq n$, fails to nullify the viability of the partial expectation $E(T^n|\{*\};k) = \infty$; $0 < \pi \leq 1.0$, and an infinite moment ensues for the present individual.

Bringing the above formulation to bear on convergent support for k as a process-transaction competence parameter, the following inference from Equation (9) is apparent. A performance sample size that is inadequate to ensure stability of finite latency values for the

person at hand, can be offset by a sufficiently large value of k for the class of individuals to which the present performer belongs, specifically $k>n-Nk'$.

Inspection of equation (9) reveals further interplay between k and N, and between k and k', both informative as to the nature of k. Allowing $Nk' + k>n$, consider the following ratio, extracted from equation (9):

$$\frac{\Gamma(Nk'+k-n)}{\Gamma(Nk'+k)} = \frac{1}{(Nk'+k-1)(Nk'+k-2) \ \ldots \ (Nk'+k-n)} .$$

$E(T^n|\{*\};k')$ evidently decreases as N and k increase, other things remaining constant. This result stems from the participant's performance-based maximum-likelihood estimate of v, MLE(v) being

$$\frac{Nk'}{\sum_{i=1}^{N} t_i}.$$

Clearly, MLE(v) increases with N, $\sum_{i=1}^{N} t_i$, along with k' remaining constant. Likewise, the estimated process-completion capacity of the current participant is raised in a parallel fashion with elevation in k (equation (A.6) of Neufeld, et al, 2002), as should be the case, considering the imputed meaning of k as a capacity-enhancing competence parameter.

Finally, equation (9) discloses that the interplay between k' and k is comparatively complex. Computer computations indicate that changes in $E(T^n|\{*\};k')$ with elevation in k' can be non-monotone, decreasing and then increasing. The decrease is understandable, in light of the influence of k' on MLE(v), above. Note, however, that as k' goes up, so does the process magnitude, with respect to the number of stages involved. Evidently effects of the latter ultimately surpass those on MLE(v), eventuating in a movement upward of the Bayesian posterior moment. Apropos of the current thrust regarding k, interestingly such eventual elevation can be replaced with continuing decline, if k is made increasingly large. This result again should be observed if k represents a competence parameter, whose influence can override that of a progressive increase in task-load k'on $E(T^n|\{*\};k')$.

Interplay with Capacity Utilization.

Unlike the Erlang, or general gamma distributions, but similar to the Weibull distribution, above, the non-homogeneous Poisson distribution specifies a continuous change in processing rate with time. This dynamical rate is $vt^{\beta-1}$, $\beta > 0$. If $\beta > 1$, the rate continuously increases; if $\beta < 1$, there is a continuous decline in rate, reflecting say a fatigue effect. The parameter β therefore is designated a capacity-application efficiency parameter. Metaphorically, it signifies the exploitation of capacity dealt to the processing system by the "capacity-resource pool"(gamma$_{r,k}$

distributed v).

It is apparent that the Weibull distribution can be obtained from the non-homogeneous Poisson by setting $k' = 1$, and taking v to the power of β. Releasing these constraints in the study of k, however, affords certain insights stemming from the relation of its effects to those of β.

As the rate parameter of the Erlang distribution now is replaced with $vt^{\beta-1}$, $f(t|v) =$

$$\frac{\beta((vt^{\beta-1})t)^{k'-1}}{(k'-1)!}\ vt^{\beta-1}\ e^{-(vt^{\beta-1})t}\ .$$

The coefficient β provides for

$$\int_0^\infty f(t|v)dt=1.0.$$

Consequently,

$$E(T^n) = \int_0^\infty \int_0^\infty \frac{(rv)^{k-1}}{\Gamma(k)}\ re^{-rv}\ \frac{\beta(vt^{\beta})^{k'-1}\ vt^{\beta-1}e^{-vt}t^{n}}{(k'-1)!}\ dt\ dv$$

$$= \frac{r^{n/\beta}\ \Gamma(k-n/\beta)\ \Gamma(k'+n/\beta)}{(k'-1)!\Gamma(k)}.$$

Upon inspecting the structure of this moment, it is apparent that the burden placed on k for retention of finite values now is alleviated, or indeed intensified– for $\beta>1.0$, and $\beta<1.0$, respectively– by a factor of $1/\beta$. Observe in passing that stress effects on moment magnitude also are modified according to β in the term $r^{n/\beta}$.

Similar to the escape of $E(T^n)$ from finite values if $k\leq n/\beta$, $E(T^n)$ breaches finite bounds if $\beta\leq n/k$. In this way, performer competence as conveyed by k, and capacity-application efficiency as conveyed by β, reciprocate with respect to moment retention of finite values.

A complementary angle on the relation of k to capacity utilization β is afforded by a rigorous and robust index of "process-completion potential," as articulated in contemporary cognitive science (Townsend & Ashby, 1983; Wenger & Townsend, 2000). This index is aligned with $E(T^n)$; it consists of $-\ln\ (\bar{F}(t))$, where $\bar{F}(t)$ is the process-latency distribution's survivor function, or the complement of the cumulative probability distribution function, $1\text{-}F(t)$ (see Appendix C.a; also Figure 12 and related text). The current interpretation of k implies effects on $-\ln\ \bar{F}(t)$ that should closely resemble or mimic those of β. The parameters k and β also should play off one another in maintaining specific levels of $-\ln\ (\bar{F}(t))$. These associations are apparent in Figure 13, which plots the above index of process-completion potential as a function of k and β for selected values of r, k', and t (formulae underling Figure 13 are presented in Appendix

C.c).

The parameter k, then, as a capacity-supply competence index behaves as it should, as set against β as a capacity-application efficiency index. It does so both with respect to $E(T^n)$, and also the allied expression of process-completion potential $-\ln(\bar{F}((t))$.

Convergence with Allied Mixing-Distribution Parameters.

The above developments surrounding construct validity for k specify that performer competence is a key contributor to capacity resources bearing on process transaction, and that this shape parameter of gamma-distributed v is an index of such performer skill. There is nothing about this formulation that excludes other capacity mixing-distribution parameters from assuming a role like that of k. The shape parameter of related mixing distributions tenably could express performer competence, with properties like those of k. In other words, k need not have a corner on conveying performer ability; indeed parameters that take on a similar role in other distributions, and affect performance indicators in a similar fashion to k, indicate that the derived analytical properties are not idiosyncratic to a gamma mixing distribution. The concept of capacity-infusing competence should transcend the shape parameter k of the gamma distribution as a particular quantitative instantiation. This version of analytical construct validity again is analogous to empirical convergent validity, where a trait evinces parallel values across divergent methods of measurement (e.g., Haynes, 2001).

A notable example whose distributed variate, like gamma's, ranges from 0 to ∞, with an intensity parameter v and shape parameter β, is the Weibull distribution, described above. As the Weibull now serves as a mixing distribution, these parameters are designated V and B, respectively. Appendix C.b shows that whether they are mixing or base distributions, the gamma and Weibull are nested in the generalized gamma distribution (additionally allowing k' of the generalized gamma to assume the status of k, as a continuous variable).

An exhaustive analysis of the two distributions' effects seems unnecessary. Sufficient to make the point are definitions of moment finiteness, in relation to k and B. Selected process-latency base distributions now include the Erlang and Weibull itself. In the latter case, the intensity (capacity) parameter v alone, and then v exponentiated by β, v^{β}, comprise the gamma- or Weibull- mixed random variate.

The respective moments are presented in Table 5. Isolating on the far right-hand gamma function $\Gamma(x)$ in each case, maintaining moments within finite bounds evidently requires identical values of k and B from gamma and Weibull mixing distributions, respectively. Convergently, then, these equivalent thresholds of k and B add to the constellation of analytical properties endorsing k's construct validity.

Finally, similar observations attend the discrete-variable analogue of the mixture involving the Erlang base whose parameter v is $\text{gamma}_{r,k}$ distributed. The Pascal distribution defines the probability of completing a process comprising k' stages on the jth trial, $j = k', k'+1, \ldots$. Its parameters are k' and p, where p denotes the probability of successful stage completion on

a trial (Patil & Joshi, 1968). The parameter p thus is the discrete-variable analogue of stage-completion capacity, and in that way maps onto v of the Erlang distribution.

The Pascal-gamma mixture defines p as e^{-u}, where u is gamma distributed, with parameters r and k. Clearly the directions in which r and k affect the distribution of the base processes' capacity parameter now are reversed, suggesting that moment finiteness may depend on critical values of r. Accordingly, the mean and variance of process-completion trials, for example, escape finite bounds if $r = 1$, and the variance does so if $r = 2$.

Summary Statistics of Prototypical Performance Samples.

The final aspect of analytical construct validity considered here pertains to statistical summaries of latency distributions for representative individuals, as related to k for the class to which these participants belong (see, e.g., Carter, et al, 1998, and Neufeld, & McCarty, 1994, for discussions of empirical estimates of class-composing homogeneity). The analysis affords another glimpse on k, as indicative of performance-ability level required to contain the statistical summaries of performance within finite bounds. Each formally-prescribed hypothetical participant is representative, inasmuch as the amount appropriated by the capacity parameter v of her/his Erlang distribution, generates the mean, or expected value of the summary statistic for the represented class.

The first summary statistic is the mean process latency, which is k'/v. With v being $\text{gamma}_{r,k}$ distributed, the grand mean taken across all participants is $k'r/(k\text{-}1)$. The result shows that the value of v for an individual whose own mean latency equals that of the grand mean, necessarily is $(k\text{-}1)/r$, which is the modal value of the $\text{gamma}_{r,k}$ distribution. Similar observations apply to the individual whose standard deviation in latency equals the expected standard deviation, or $k'^{1/2}r/(k\text{-}1)$. Finite values of these summary statistics demand that $k > 1.0$.

In like fashion, v for an individual whose variance in process latencies is equal to that of the expected value -- the average variance -- is $((k\text{-}1)(k\text{-}2))^{½}/r$. The result follows because the latency variance, given v is k'/v^2, and the expected or mean variance is $k'r^2/((k\text{-}1)(k\text{-}2))$. (Note that the theoretically-prescribed value of v for the expected standard deviation, and that for the expected variance, are not identical, as $E((s.d.(T|v))^2) = E(Var(T|v)) \neq (E(s.d.(T|v)))^2$. Here, finite values of the summary statistic requires that $k>2$.

Complementing Empirical Support.

The above exposition of the analytical side of construct validity should dovetail with empirical support in its varied forms (enumerated in Haynes & O'Brien, 2000). A tactic of choice, when it comes to formal-model parameters, involves selective sensitivity to experimental manipulations, or to variation in pertinent organismic, participant variables. Concrete illustrations from the literature include those addressed to the understanding of parameters symbolizing determinants of retro-active interference with traces of memorized items (Chechile, 1987). They address, as well, parameters linked to the detection and retention of informational sources (Riefer, Knapp, Batchelder, Bamber & Manifold, 2002).

As for the present parameters, *k* has been released as a free parameter to express variations in cognitive (visual scanning) performance across groups differing in stress proneness (Hamilton, 1980). Model predictions of empirical performance configurations failed to compete with those where *r* instead was freed to vary with stress susceptibility (Neufeld & Carter, 2000). In other instances (Neufeld, 2002), variation in *k* effectively expressed effects of practice, unlike a competing parameter considered to indicate the number of stimulus features to be processed during task trials (Weinstein, 1974).

In Summary.

The method of illustration was used to motivate and exposit engagement with analytical aspects of construct validity, as it related to interpretation of model parameters. Focus centered on support for espoused meaning of a parameter bearing on the capacity-resource pool of a cognitive process (e.g., stimulus encoding). The capacity-resource pool was operationalized as a stochastic distribution of a process-model parameter conveying speed of process transaction. Construct validity of the imputed substantive significance of the capacity- affecting parameter, *k,* in turn took the form of its mathematically-derived analytical properties.

The stochastic distribution of capacity values was "gamma", with parameters *r* and *k*. Increasing values of these parameters shifted the capacity distribution downward, and upward, respectively. The shape parameter *k* was tendered as an indicator of process-completion skill or competence.

Evidence constellated for the posited interpretation of *k* included substantively-meaningful properties transcending specific process-latency (base) distributions. Contrasting structures of base-distribution models were employed, tenably standing for diverse types of cognitive-task operations. Bayesian extensions lent additional support to educed inferences.

Consequences of varying *k* interfaced with the status of other processing-system parameters, in ways to be expected with the interpretation accorded *k*. Convergent support was endowed also by alternate-distribution parameters, whose effects on task performance should resemble those of *k*, given that they occupy a similar mathematically-prescribed role.

Next, performance of hypothetical participants, whose distributions of process latencies typified those of their embedding class, were inspected for workings of *k*. Performance properties (summary-statistic functions of *k*) of these idealized participants again conformed to the claimed substantive significance of this parameter.

Finally, compatibility of analytical results with empirical support for construct validity was described.

Stochastic-Modeling and the Psychometric-Artifact Issue in the Study of Differential Cognitive Deficit.

Stochastic dynamical modeling of cognitive performance can speak to certain longstanding measure-theoretic predicaments accompanying the study of differential cognitive deficit in schizophrenia, and other disorders (Carter & Neufeld, 1999; Neufeld,, et al, 2002). At the forefront is the "psychometric-artifact problem", which states that the apparent separation between pathological and control groups on a given measure conflates differences on the addressed cognitive faculty with the measure's psychometric properties. A measure therefore can spuriously indicate greater deficit on the faculty it putatively taps because it possesses greater psychometric precision than measures directed to other faculties (Chapman & Chapman, 1973; 2000). Advocated redress of the problem has entailed pre-matching measures' psychometric properties using a standardization group composed of non-patients who vary widely on the faculties being examined. Matching purportedly should incorporate variance in the measures' observed scores, and in their respective reliabilities. In addition, simplicity of factorial structure is recommended (Strauss, 2001; cf. Neufeld, 1984; Neufeld & Broga, 1981).

The issue, and its endorsed solution continues to elicit considerable adherence (e.g., Brenner, Wilt, Lysaker, Koyfman & O'Donnel, 2003). The proposed solution of psychometric-property matching also has received substantial challenge, and alternatives have been articulated (Carbotte, 1978; Knight & Silverstein, 2001; Neufeld, & Broga, 1977; 1981; Neufeld, 1984; 1989; Neufeld, et al, 2002). The issue is dissected here by translating it into simple but arguably sufficient measure-theoretic terms, and then evaluating effects of typical psychometric-adjustment practices.

Boiled down to essentials, the argument for some three decades has been more or less as follows. More psychometrically-precise measures will have more statistical power for detecting control-pathological group separation, and thus be differentially favored over psychometrically inferior measures in indicating statistically significant deficit, or disproportionate deficit (expressed as an interaction with corresponding group-wise simple main effects). Comparative effect-size estimates of course should conform to the above inter-measure differences in statistical properties, other values (notably sample size) being equal.

Note that for two groups, statistical power for detecting a difference on a selected measure is an increasing function of the squared non-centrality parameter $n(1/4d^2)$, where d refers to Cohen's (1988) d, or the standardized population group difference in means $(\mu_1 - \mu_2)/\sigma$. In this expression, σ is the common within-groups standard deviation, and n is the common group sample size. The expression $1/4(\mu_1 - \mu_2)^2$ conveys between-group effects, and tenably is composed of some amalgam, first, of those that are unique to group membership, and constant among its members (corresponding to Meehl-Golden (1982) taxonic variables), denoted here as β; and, second, those that represent between-group inequalities with respect to latent variables that also generate true-score variance across participants, within groups τ. The net contribution to the numerator of $1/4d^2$ from the latter source is denoted $c\tau$, where the scalar c increases as sources of between-group separation on the measure overlap with those of within-group, inter-

participant variation; $0=<c=<1.0$.

The expression $1/4(\mu_1-\mu_2)^2$ thus can be replaced with $c\tau + \beta$. Finally, duly replacing σ^2 with $\tau + e$, where the term e represents measurement-error variance, makes $1/4d^2$ itself become

$$(c\tau + \beta)/(\tau + e). \qquad (10)$$

Although not essential, for mathematical tractability, and that of illustration, it is assumed that none of c, τ, or e are functions of one another.

Turning to the issue of psychometric adjustment using standardization-group data, it is apparent from equation (10) that the nub of between-group inequality, c and β are bypassed; within-standardization-group measure reliability is equal to $\tau/(\tau + e)$, and observed-score variance is equal to $\tau + e$.

Interestingly, by equation (10), where all between-group variance is conveyed by β, $\beta > 0$, $c = 0$, $1/4d^2$ evidently can be increased with reduction in τ. That is to say, the measure becomes less group-discriminating as its standardization-group psychometric precision goes up. In fact, the effects of reducing τ are unveiled by differentiating (10) with respect to τ, the result being $(ce - \beta)/(\tau + e)^2$. Clearly, increasing τ-- and along the way, increasing measure reliability $\tau/(\tau + e)$-- has adverse effects on power if $ce < \beta$, and the opposite if $ce > \beta$. Thus the consequences for d^2 of varying τ are complex, and in fact unknowable from standardization-group variance, because it is mute about the structure of group differences composed of c and β. Of course, decreasing e invariably increases reliability as well as d^2, as would be expected, and as is obvious from equation (10). But c and β remain unknown, leaving psychometric matching to grope blindly when it comes to amounts by which e should be altered (say through increased accuracy of instrumentation, or pre-aggregating measurements; see, e.g., Neufeld, & Gardner, 1990). The upshot is that equalizing a pair of measures on τ and e has no problem whatsoever in leaving them grossly unequal in their group-discriminating power.[9]

Empirical examples instantiating the theoretical dissociation between control-pathology group discriminability, and classical psychometric properties, have been amply documented elsewhere (Knight & Silverstein, 2001; Neufeld, 1984; 1989; Neufeld & Broga, 1981). These instantiations need not be recapitulated here. More productively, it is shown that equation (10) can be superposed over and against a stochastic-modeling formulation. The conversion metamorphoses the equation and essentially leaves behind as vestigial the statistical-property issue.

A Stochastic-Modeling Translation of the Statistical-Property Issue.

The terms of equation (10) can be mapped onto substantive stochastic-modeling expressions. Those of the dual-parameter mixture model are invoked in the present instance, but the resulting observations transcend the immediate case. Each translated term potentially becomes a model prediction of correspondingly partitioned data. Statistical power considerations now pertain to rejection of a non-trivial competing model (cf. Cohen, 1988, chap. 7; pp. 16,17).

More generally, ushered in is the substantial methodology of model selection. Included are procedures for evaluating empirical fit, and comparative fit, of model predictions, that go beyond conventional goodness-of fit statistical-significance tests (e.g., Myung, Forster & Browne, 2000). Along with accessing advances in model testing and refinement, the translation of equation (10) breathes content into its otherwise substantively desiccated terms.

The term *e* in equation (10) signifies "psychometric measurement error," taking the form of random variation in participants' empirical values across observational trials. The current mixture model, in turn quantifies participant-wise inter-trial variance in latencies as the expected or mean conditional variance, given *k*' and *v*,

$$E(Var(T|k' \cap v)_{m_{1,2},r,k}).$$

Here, $m_{1,2}$ signifies that the expected variance is taken across both control and pathological groups, which for simplicity, and in keeping with the design depicted in Figure 6, differ only with respect to *m*. The actual amount prescribed by the mixture model is

$$(r^2 m_{1,2})/((k-1)(k-2)). \tag{11}$$

The model counterpart of τ is the variance in the conditional mean or expected latency, given *k*' and *v*,

$$Var(E(T|k' \cap v)_{m_{1,2},r,k}) = \frac{m_{1,2}\ r^2(k-1+m_{1,2})}{(k-1)^2(k-2)}. \tag{12}$$

Both the expected conditional variance, and variance in the conditional expectations, above, provide for random variation in *v* and *k*' among participants, within groups.

Model analogues of $c\tau$ and β cast *k*' as a Poisson$_m$-mixed, and then fixed parameter, respectively. In the first case, *k*' is randomly distributed across participants within groups. The distributions themselves are separated, according to the values of $m_{1,2}$. For example, m_1 may be *m*' of Figure 6, and m_2 may be *m*' + *g*. The differences in m_1 and m_2 thus are affiliated with *c* of equation (10). The scaled value of τ, $c\tau$ therefore has a model-representation as the following function of the between-groups difference in means:

$$1/4 \left(\frac{(m_1 r)}{(k-1)} - \frac{m_2 r}{(k-1)} \right)^2. \tag{13a}$$

The analogous value for *k*' as fixed within groups is

$$1/4 \left(\frac{(k_1^{/}r)}{(k-1)} - \frac{k_2^{/}r}{(k-1)} \right)^2 . \qquad (13b)$$

Equations (11) through (13, a or b) together embody twelve model predictions of empirical terms. The data of each diagnostic group, under each experimental condition (encoding load, of Figure 6), can be assembled into the expected conditional variance, variance in conditional expectations, and group mean. Parameter estimates, obtained for example through moment-fitting procedures, are five in number according to the prevailing model design. It is apparent that each parameter is called upon simultaneously to participate in empirical values whose sources range from within-participant variation to inter-group separation. Among other considerations, severity of model testing increases with the ratio of empirical observations to parameter-set-size, and heterogeneity of conditions to which the model applies.

Inferential validity regarding proposed dysfunction, then, ultimately rests on rigor of the model diagnostics that are exerted. Predictions desirably include qualitative properties of empirical configurations. For example, the present model prescribes additive effects on means, of diagnostic groups and experimental conditions, which is tantamount to a predicted second-order difference (above) being equal to 0 (e.g., Neufeld, et al, 1993; Neufeld & Williamson; 1996). Augmenting significance tests for goodness-of-fit- involving predicted distribution moments (crafted from equations (11) through (13)) also can be constructed (Neufeld, Vollick, Carter, Boksman, Levy, George & Jetté, 2007).

Further tests are available from distribution properties, which in the case of dynamical stochastic modeling, are functions of time *t* (García Pérez, 1994; 2000; Van Zandt, 2000). These properties flow from the very design spawning the above distributional moments (e.g., density functions, illustrated in Figures 10 and 11). Bayesian extensions moreover stand to further proliferate model predictions. Posterior density and distribution functions, given an individual sample of latencies (e.g., Figures 8, 9, 12), once more are engineered by the stochastically-modeled system giving rise to Bayesian-posterior distributional moments themselves (illustrated in Neufeld et al., 2002).

Concluding Comments

Placing clinical cognitive science on a decidedly formal platform can unearth productive routes of study, and discern otherwise hidden substantive nuances. Brokering the clinical ramifications of contemporary cognitive science makes for a potentially seamless transition between these two domains of investigation. Over-arching model designs can be constructed so as to mediate group-level findings to individual differences in their expression amongst constituent members. Developments self disclose implications for evaluation of the cognitive side of treatment-program efficacy, and progress of an individual's own response to treatment, with respect to cognitive functioning. Also disclosed are assumptive underpinnings of the adumbrated assessment technology, along with lingering ambiguities awaiting resolution.

Formal clinical cognitive science uniquely convenes a quantitative infrastructure for assessing construct validity of meaning assigned to its parameters and potentially other properties. It also engages measurement and statistical challenges to inferences about cognitive dysfunction. The current offering at minimum represents a beachhead from which a stronghold of formal clinical cognitive science and assessment methodology hopefully may be consolidated.

Acknowledgments

Sources of support for this research include operating grants from the Social Sciences and Humanities Research Council of Canada, The Workplace Safety and Insurance Board, and operating grants and a New Emerging Teams grant from the Canadian Institutes of Health Research. Appreciation is expressed to the University College of the Fraser Valley, the Department of Psychology, University of British Columbia, Kenneth D. Craig, sabbaticant host, and to Ruth and Douglas Barr, private citizens, for providing facilities during the writing of this manuscript.

References

Aguirre, G.K., Zarahn, E., & D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *Neuroimage, 8,* 360-369.

Averill, J. R. (1973). Personal control over aversive stimuli and its relationship to stress. *Psychological Bulletin, 80,* 286-303.

Bandura, A. (1984). Representing personal determinants in causal structures. *Psychological Review, 91,* 508-511.

Batchelder, W. H. (1998). Multinomial processing tree models and psychological assessment. *Psychological Assessment, 10,* 331-344.

Batsell, R.R., Poling, J.C., Cramer, R.D.,& Miller. C.M.. (2003). Useful Mathematical Relationships Embedded in Tversky's Elimination by Aspects Model. *Journal of Mathematical Psychology, 47,* 538-544.

Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis* (2nd edition). New York: Springer.

Boksman, K., Théberge, J., Williamson, P., Drost, D., Malla, A., Densmore, M., Takhar, J., Pavlovsky, W., Menon, R. & Neufeld, R.W.J. (2005). A 4.0 Tesla fMRI study of brain connectivity during word fluency in first episode schizophrenia. *Schizophrenia Research, 75,* 247-263.

Bollen, K.A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology, 53,* 605-634.

Borowski, E. J. and Borwein, J. M. (1989). *Dictionary Of Mathematics.* HarperCollins.

Boynton, G.M. Engel, S.A., Glover, G.H. & Heeger, D.H. (1996). Linear systems analysis of fMRI in human V1. *Journal of Neuroscience, 16,* 4207-4221.

Braithwaite, R. B. (1968). *Scientific explanation.* London: Cambridge University Press.

Brenner, C.A., Wilt, A., Lysaker, P.H., Koyfman, A., & O'Donnell, B.F. (2003). Psychometrically Matched Visual-Processing Tasks in Schizophrenia Spectrum Disorders. *Journal of Abnormal Psychology, 112,* 28-37.

Brown, E. & Holmes, P. (2001). Modeling a simple choice task: Stochastic dynamics of mutually inhibitory neural groups. *Stochastics and Dynamics, 1,* 1-33.

Bryk, A., & Raudenbush, S. W. (1992). *Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods.* Newbury Park, CA: Sage.

Busemeyer, J. R. (1980). Importance of measurement theory, error theory, and experimental design for testing the significance of interactions. *Psychological Bulletin, 88,* 237-244.

Busemeyer, J. R. & Stout, J. C. (2002) A Contribution of Cognitive Decision Models to Clinical Assessment: Decomposing Performance on the Bechara Gambling Task. *Psychological Assessment, 14,* 253-262

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432-459.

Busmeyer, J. R., & Wang, Y. (2000). Model comparisons and model selections based on generalization test methodology. *Journal of Mathematical Psychology, 44,* 171-189.

Campbell, D T & Fiske, D W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Carbotte, R.M. (1978). Converging operations or matched control tasks? *Journal of Psychiatric Research, 14,* 313-316.

Carter, J.R., & Neufeld, R.W.J. (1999). Cognitive processing of multidimensional stimuli in schizophrenia: Formal modeling of judgment speed and content. *Journal of Abnormal Psychology, 108,* 633-654.

Carter, J.R., & Neufeld, R.W.J. (in submission). Cognitive processing of facial effect: Neuro-connectionist modeling of deviations in schizophrenia.

Carter, J.R., Neufeld, R.W.J., & Benn, K.D. (1998). Application of process models in assessment psychology: Potential assets and challenges. *Psychological Assessment, 10,* 279-298.

Casti, J.L. (1989). *Alternate Realities: Mathematical Models of Nature and Man.* John Wiley & Sons, 1989.

Chapman, L. J., & Chapman, J. P. (1973). Problems in the measurement of cognitive deficit. *Psychological Bulletin, 79,* 380-385.

Chapman, L. J., & Chapman, J. P. (2001). Commentary on Two Articles Concerning Generalized and Specific Cognitive Deficits. *Journal of Abnormal Psychology, 110,* 31-39.

Chechile, R. A. (1987). Trace susceptibility theory. *Journal of Experimental Psychology, 116,* 203-222.

Chechile, R.A. (1998). A new method for estimating model parameters for multinomial data. *Journal of Mathematical Psychology, 42,* 432-471.

Clark, J. M., & Paivio, A. (1989). Observational and theoretical terms in psychology: A cognitive perspective on scientific language. *American Psychologist, 44,* 500-512.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.

Delucchi, K. L. (1993). On the use and misuse of chi-square. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Statistical issues* (pp. 295-320). Hillsdale, NJ: Lawrence Erlbaum Associates.

Doob, J. L. (1953). *Stochastic processes.* New York: John Wiley & Sons.

Embretson, W. S. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93,* 179-197.

Evans, M., Hastings, N., & Peacock, B. (2000). *Statistical distributions* (3rd Ed.). New York: Wiley & Sons.

Feller, W. (1966). *An introduction to probability theory and its applications* (Vol. II). New York: Wiley.

Frackowiak, R.S.J., Friston, K.J., Frith, C.D., Dolan, R.J., Price, C.J., Zeki, S., Ashburner, J. & Penny, W. (2004). (Eds.), *Human Brain Function,* 2nd Edition. San Diego: Academic Press.

Friston, K. Harrison, L. & Penny, W. (2003). Dynamic causal modelling . *Neuroimage, 19,* 1273-1302.

García Pérez, M. A. (1994). Parameter estimation and goodness-of-fit testing in multinomial models. *British Journal of Mathematical and Statistical Psychology, 47,* 247-282.

García Pérez, M. A. (2000). Exact finite-sample significance and confidence regions for goodness-of-fit statistics in one-way multinomials. *British Journal of Mathematical and Statistical Psychology, 53,* 193-207.

Geddes, K. O., & Scott, T. I. (1989). Recipes for classes of definite integrals involving exponentials and logarithms. In E. Kaltofen & S. M. Watt (Eds.), *Computers and Mathematics* (pp. 192-201). New York: Springer-Verlag.

George, L., & Neufeld, R.W.J. (1984). Imagery and verbal aspects of schizophrenic informational-performance. *British Journal of Clinical Psychology, 23,* 9-18.

Gilden, D.L. (2001). Cognitive emission of 1/f noise. Psychological Review, 108, 33-56.

Glass, D. C., & Singer, J. E. (1972). *Urban stress.* New York: Academic Press.

Gleick, J. (1988) *Chaos: Making a new science.* NY: Penguin Books.

Gregson, R.G. & Pressing, J. (2000). Dynamic modelling. In: Cacioppo JT, Tassinary LG and Berntson G (Eds.), *Handbook of Psychophysiology, 42,* 924-948. New York: Cambridge University Press.

Hamilton, V. (1980). An information processing analysis of environmental stress and life crises. In I. G. Sarason & C. D. Spielberger (Eds.), *Stress and anxiety* (Vol. 7, pp. 13-30). Washington: Hemisphere.

Harris, B. (1966). *Theory of probability.* Reading, Mass: Addison-Wesley.

Haynes, S.N. (2001). Clinical applications of analogue behavioral observation: Dimensions of psychometric evaluation. *Psychological Assessment, 13,* 73-85.

Haynes, S. N. and O'Brien. W. O. (2000). *Principles of Behavioral Assessment: A Functional Approach to Psychological Assessment.* New York: Plenum/Klewer Press.

Herrnstein, R.J. (1979). Derivatives of matching. *Psychological Review, 86,* 486-495.

Highgate-Maynard, S., & Neufeld, R.W.J. (1986). Schizophrenic memory-search performance involving nonverbal stimulus properties. *Journal of Abnormal Psychology, 95,* 67-73.

Hintzman, D. L. (1991). Why are formal models useful in psychology? In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdoch.* Erlbaum Associates.

Hockley, W. E., & Murdock, B. B. Jr. (1987). A decision model for accuracy and response latency in recognition memory. *Psychological Review, 94,* 341-358.

Huber, M.T., Braun, H.A., & Krieg, J.C. (1999). Consequences of deterministic and random dynamics for the course of affective disorders. *Biological Psychiatry, 46,* 256-262.

Huettel, S.A., Song, A.W., & McCarthy, G. (2004). *Functional Magnetic Resonance Imaging.* Sunderland, MA: Sinauer Associates.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan (1994). *Continuous Univariate Distributions,* Volume 1 , 2nd edition, New York: John Wiley & Sons.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan (1995). *Continuous Univariate Distributions,* Volume 2, 2nd edition, New York: John Wiley & Sons.

Karabatsos, G. (2006). Bayesian nonparametric model selection and model testing. Special Issue on Model Selection, *Journal of Mathematical Psychology, 50,* 123-148.

Kenny, J. F., & Keeping, E. S. (1951). *Mathematics of statistics,* 2nd Ed. New York: Van Nostrand.

Kirk, R. E. (1994). *Experimental design: Procedures for the behavioral sciences* (3rd ed). Monterey, CA: Brooks/Cole.

Kline, M. (1985). *Mathematics and the Search for Knowledge*. Oxford: Oxford University Press.

Kline, R.B. (1988). *Structural equation modeling.* NY: Guilford Press.

Knight, R.A. & Silverstein, S.M. (2001). A process-oriented approach for averting confounds resulting from general performance deficiencies in schizophrenia. *Journal of Abnormal Psychology, 110*, 15-30

Koçak, H. (1989). *Differential and difference equations through computer experiments* (2nd edition, pp. 39-56). New York: Springer-Verlag.: Hillsdale, NJ: Erlbaum.

Kuhn, T.S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Kukde, M.P., & Neufeld, R.W.J. (1994). Facial electromyographic measures distinguish covert coping from stress response to stimulus threat. *Personality and Individual Differences, 16,* 211-228.

Lanius, R.A., Williamson, P.C., Densmore, M., Boksman, K., Neufeld, R.W.J., Gati, J.S., & Menon, R.S. (2004). The nature of traumatic memories: A 4-T fMRI functional connectivity analysis. *American Journal of Psychiatry, 161*, 36-44..

Lees, M.C., & Neufeld, R.W.J. (1999). Decision-theoretic aspects of stress arousal and coping propensity. *Journal of Personality and Social Psychology, 77,* 185-208.

Link, S. W. (1982). Correcting response measures for guessing and partial information. *Psychological Bulletin, 92,* 469-486.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

Luke, D. A., & Homan, S. M. (1998). Time and change: Using survival analysis in clinical assessment and treatment evaluation. *Psychological Assessment, 10,* 360-378.

MacCallum, R. & Ashby, F. G. (1986). Relationships between linear systems theory and covariance structure modeling. *Journal of Mathematical Psychology, 30,* 1-27.

Marley, A.A. (1989). A random utility family that includes many of the "classical" models and has closed form choice probabilities and choice reaction times. *British Journal of Mathematical and Statistical Psychology, 42,* 13-36

Marley, A. A. J., & Colonius, H. (1992). The "horse race" random utility model for choice probabilities and reaction times, and its competing risks interpretation. *Journal of Mathematical Psychology, 36,* 1-20.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

McClelland, J., Rummelhart, D. E., & the PDP Research Group (1986). *Parallel distributed processing -- Explorations in the microstructure of cognition, Vol. 1: Foundations.* Cambridge, Mass: M.I.T. Press.

McFall, R. M., & Townsend, J. T. (1998). Foundations of psychological assessment: Implications for cognitive assessment in clinical science. *Psychological Assessment, 10,* 316-330.

McFall, R. M., Townsend, J. T., & Viken, R. J. (1995). Diathesis stress model or "just so" story? *Behavioral and Brain Sciences, 18,* 565-566.

McFall, R. M., & Townsend, J. T. (1998). Foundations of psychological assessment: Implications for cognitive assessment in clinical science. *Psychological Assessment, 10,* 316-330.

McFall, R. M., Townsend, J. T., & Viken, R. J. (1995). Diathesis stress model or "just so" story? *Behavioral and Brain Sciences, 18,* 565-566.

McGill, W. J. & Gibbon, J. (1965). The general-gamma distribution and reaction times. *Journal of Mathematical Psychology, 2,* 1-18.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806-843.

Meehl, P. E., & Golden, R. R. (1982). Taxometric methods. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology.* New York: Wiley.

Miles, M.B., & Huberman, A.M. (1994). *Qualitative Data Analysis.* CA: Sage

Morrison, D. G. (1979). An individual differences pure extination process. *Journal of Mathematical Psychology, 19,* 307-315.

Morrison, M.S., Neufeld, R.W.J., & Lefebvre, L.A. (1988). The economy of probabilistic stress: Interplay of controlling activity and threat reduction. *British Journal of Mathematical and Statistical Psychology, 41,* 155-177.

Myung, I.J., Forster, M.R., & Brown, M.W. (2000). (Eds.), Special Issue on Model Selection. *Journal of Mathematical Psychology, 44,* 1-2

Neufeld, R.W.J. (1982).On decisional processes instigated by threat: Some possible implications for stress-related deviance. In R.W.J. Neufeld (Ed.), *Psychological stress and psychopathology.* New York: McGraw-Hill.

Neufeld, R.W.J. (1984). Re: The incorrect application of traditional test discriminating power formulations to diagnostic-group studies. *Journal of Nervous and Mental Disease, 172,* 373-374.

Neufeld, R.W.J. (1989). Methodological aspects of laboratory studies of stress. In R.W.J. Neufeld (Ed.), *Advances in the investigation of psychological stress.* New York: John Wiley & Sons.

Neufeld, R.W.J. (1991). Memory in paranoid schizophrenia. In P. Magaro (Ed.), *The cognitive bases of mental disorders: Annual review of psychopathology* (Vol. 1, pp. 231-261). Newbury Park, CA: Sage. *(Invited)*

Neufeld, R.W.J. (1994). *Theoretical stress and stress-proneness effects on information processing in light of mathematical models of stochastic processes.* Research Bulletin No. 720, Department of Psychology, University of Western Ontario.

Neufeld, R.W.J. (1996). *Stochastic models of information processing under stress.* Department of Psychology Research Bulletin No. 734. London, ON: University of Western Ontario.

Neufeld, R.W.J. (1999a). Quantitative schemata of decisional stress-control. Paper presented at the 32nd annual meeting of the Society for Mathematical Psychology, Santa Cruz.

Neufeld, R.W.J. (1999b). Dynamic differentials of stress and coping. *Psychological Review, 106,* 385-397.

Neufeld, R.W.J. (2002). *Stochastic modelling of stress effects on proofreading.* Presented at the Annual Summer Inter-Disciplinary Conference on Cognitive Science, Squamish, British Columbia.

Neufeld, R.W.J. & Broga, M. (1977). *Fallacy of the reliability-discriminability principle in research on differential cognitive deficit.* Department of Psychology Research Bulletin No. 360. London, ON: University of Western Ontario.

Neufeld, R.W.J., & Broga, M.I. (1981). Evaluation of information-sequential aspects of schizophrenic performance, II: Methodological considerations. *Journal of Nervous and Mental Disease, 169,* 569-579.

Neufeld, R.W.J. & Carter, J. (2000). *Formally deciphering effects of stress on cognitive performance.* Paper presented at the 33rd Annual Meeting of the Society for Mathematical Psychology, Queen's University, Kingston, Ontario.

Neufeld, R.W.J., Carter, J.R., Nicholson, I., & Vollick, D. (1999). Schizophrenia. In W.L. Marshall (Ed.), *Abnormal Psychology Perspectives* (pp. 344-372). New York: Prentice-Hall.

Neufeld, R.W.J., & Gardner, R.C. (1990). Data aggregation in evaluating psychological constructs: Multivariate and logical-deductive considerations. *Journal of Mathematical Psychology, 34,* 276-296.

Neufeld, R.W.J., & McCarty, T. (1994). A formal analysis of stressor and stress-proneness effects on basic information processing. *British Journal of Mathematical and Statistical Psychology, 47,* 193-226.

Neufeld, R.W.J., & Mothersill, K. (1980). Stress as an irritant of psychopathology. In I. Sarason & C.D. Spielberger (Eds.), *Stress and anxiety, Vol. VIII.* New York: Hemisphere.

Neufeld, R.W.J., & Williamson, P. (1996). Neuropsychological correlates of positive symptoms: Delusions and hallucinations. In C. Pantelis, H.E. Nelson, & T.R.E. Barnes (Eds.), *Schizophrenia: A neuropsychological perspective* (pp. 205-235). London: John Wiley & Sons.

Neufeld, R.W.J., Vollick, D., & Highgate, S. (1993). Stochastic modelling of stimulus encoding and memory search in paranoid schizophrenia: Clinical and theoretical implications. In R.L. Cromwell & R.C. Snyder (Eds.), *Schizophrenia: Origins, processes, treatment, and outcome: The Second Kansas Series in Clinical Psychology* (pp. 176-196). Oxford: Oxford University Press.

Neufeld, R.W.J., Carter, J.R., Boksman, K., Jetté, J., & Vollick, D. (2002). Application of stochastic modelling to group and individual differences in cognitive functioning. *Psychological Assessment, 14,* 279-298.

Neufeld, R.W.J., Vollick, D., Carter, J.R., Boksman, K., Levy, L., George, L., & Jetté, J. (2007). A mathematical process account of group and individual differences in memory-search facilitative stimulus encoding, with application to schizophrenia. In R.W.J. Neufeld (Ed.), *Advances in clinical cognitive science: Formal modeling and assessment of processes and symptoms.* Washington, D.C.: American Psychological Association.

Nicholson, I.R., & Neufeld, R.W.J. (1992). A dynamic vulnerability perspective on stress and schizophrenia. *American Journal of Orthopsychiatry, 62,* 117-130.

Nicholson, I.R., & Neufeld, R.W.J. (1993). Classification of the schizophrenias according to symptomatology: A two-factor model. *Journal of Abnormal Psychology, 102,* 259-270.

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology, 43,* 25-33.

Paivio, A. (1971). *Imagery and verbal processes.* New York : Holt, Rinehart & Winston.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory and Cognition, 3,* 635-647.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach.* New York: Oxford University Press.

Parzen, E. (1962). *Stochastic processes.* San Francisco: Holden-Day.

Patil, G. G., & Joshi, S. W. (1968). *A dictionary and bibliography of discrete distributions.* New York: Hafner Publishing Company.

Piasecki, Thomas M.; Jorenby, Douglas E.; Smith, Stevens S., Fiore, M.C., & Baker, T.B. (2003). Smoking withdrawal dynamics: I. Abstinence distress in lapsers and abstainers. *Journal of Abnormal Psychology, 112,* 3-13.

Polkinghorne, J. (2003).*The friendship of science and religion.* Plenary lecture: Veritas Forum. University of Western Ontario, London, Ontario.

Rappaport, A. (1983). *Mathematical models in the social and behavioral sciences.* New York: Wiley.

Reiman, E.M., Lane, R.D., Van Petten, C., & Bandettini, P.A. (2000). Positron emission tomography and functional magnetic resonance imaging. In: Cacioppo, J.T., Tassinary, L.G. et al (Eds.), *Handbook of psychophysiology,* 2nd ed. (pp. 85-118). New York, NY: Cambridge University Press.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95,* 318-339.

Riefer, D.M., Knapp, B., Batchelder, W.H., Bamber, D., & Manifold, V. (2002). *Psychological Assessment, 14,* 184-201

Roe, R.M., Busemeyer, J. T., & Townsend, J.T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review, 108,* 370-392.

Ross, S. M. (1996). *Stochastic processes,* 2nd ed. New York: John Wiley & Sons.

Rumelhart, D., McClelland, J. and The PDP Research Group (1987), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* (Vol. 2), MIT Press (Cambridge, Mass.).

Santor, D. A., & Ramsay, J. O. (1998). Progress in the technology of measurement: Applications of item response models. *Psychological Assessment, 10,* 345-359.

Schiffman, S.S., Reynolds, M.L, & Young, F.W. (1981) Introduction to Multidimensional Scaling: Theory, Methods and Applications.. NY: Academic Press.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review, 102,* 567-593.

Staddon, J. E. R. (1984). Social learning theory and the dynamics of interaction. *Psychological Review, 91,* 502-507.

Staddon, J. E. R. (1991) "The distemper of learning..." A review of S. B. Klein and R. R. Mowrer (Eds.) Contemporary Learning Theories: Instrumental Conditioning Theory and the Impact of Biological Constraints on Learning. Contemporary Psychology, 36, 506-507.

Steiger, J.H., & Fouladi, R.T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In L.L. Harlow, S.A. Mulaik, & J.H. Steiger (Eds.), *What if there were no significance tests?* (pp. 221-257). Mahwah, NJ: Erlbaum.

Sternberg, S. (1975). Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology, 27,* 1-32.

Strauss, M.E. (2001). Methodology for Identifying Specific Psychological Deficits: Introduction to the Special Section. *Journal of Abnormal Psychology, 110,* 4-5.

Takane, Y., & Sergent, J. (1983). Multidimensional scaling models for reaction times and same-different judgments. *Psychometrika, 48,* 393-423.

Tollenaar, N., & Mooijaart, A. (2003). Type I errors and power of the parametric bootstrap goodness-of-fit test: Full and limited information. *British Journal of Mathematical and Statistical Psychology, 56,* 271-288.

Tomarken, A.J., & Baker, A.J. (2003). Introduction to the special section on structural equation modeling. *Journal of Abnormal Psychology, 112,* 523-525.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology, 28,* 363-400.

Townsend, J.T. (1990). 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. (1994). A visual approach to nonlinear dynamics: Just cartoons or a serious pedagogical device? [Review of the 1991 book "A visual introduction to dynamical systems theory for psychology".] *American Journal of Psychology, 107,* 117-123.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3). Hillsdale, NJ: Erlbaum.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modelling of elementary psychological processes.* Cambridge University Press.

Townsend, J.T., & Fific, M. (2004, in press). Parallel versus serial processing and individual differences in high-speed search in human memory. *Perception and Psychophysics.*

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology, 39,* 321-359.

Townsend, J. T., & Wenger, M. J. (in submission). The serial-parallel dilemma: A case study in a linkage of theory and method.

Tversky, A. (1972a). Choice by elimination. *Journal of Mathematical Psychology, 9,* 341-367.

Tversky, A. (1972b). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 289-299.

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review, 7,* 424-465.

Vollick, D.N. (1994). *Stochastic models of encoding-latency means and variances in paranoid schizophrenia* (Ph.D. Thesis). London, ON: Faculty of Graduate Studies, University of Western Ontario.

Vollick, D., & Neufeld, R.W.J. (in preparation). Stochastic modelling of encoding-latency means and variances in paranoid schizophrenia.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua.* Newbury Park, CA: Sage

Weinstein, N. D. (1974). Effect of noise on intellectual performance. *Journal of Applied Psychology, 59,* 548-554.

Wenger, M.J., & Townsend, J. T. (2000). Basic response time tools for studying general processing capacity in attention, perception, and cognition. *Journal of General Psychology, 127,* 67-99.

Wiggins, J.S. (1973). *Personality and prediction: Principles of personality assessment.* Reading, Mass.: Addison-Wesley.

Wright, B.D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement, 14,* 97-116.

Table 1. Schema of Theoretical Systems, with Selected Examples.

	Static	Dynamic
Deterministic	Verbal Theory; Non-Dynamical Flow Diagrams	Deterministic Nonlinear Dynamical Systems ("Chaos-Theoretic Systems")
Stochastic	Some Decision-and-Choice Models (e.g., Selected Subjective-Utility Models)	Dynamic Extensions of Decision-and-Choice Models; Classical Stochastic Models; Nonlinear Dynamical-Systems Models, with a Stochastic Element

Note: Adapted from Busemeyer & Townsend, 1993.

Table 2. Unconditional Joint Densities of Encoding-Latency Performance Samples Under Alternate Bayesian Priors

Classification	Representative individual-participant-latency sets			
	Paranoid schizophrenia		Control	
	High-encoding load {3.476, 7.150, 9.00, 4.238}	Low-encoding load {2.177, 3.184, 2.411, 2.498}	High-encoding load {1.137, 3.204, 1.371, 2.207}	Low-encoding load {1.037, 0.554, 0.191, 6.390}
Paranoid schizophrenia				
High-encoding load[a]	.17176 (10^{-12})	.00684	..54107 (10^{-11})	.29397 (10^{-40})
Low-encoding load[b]	.999776 (10^{-11})	.00737	.70547 (10^{-9})	.23614 (10^{-32})
Control				
High-encoding load[c]	.32930 (10^{-7})	.00083	.00002	.75783 (10^{-13})
Low-encoding load[d]	Pr{ * } → 0.0	.1486 (10^{-12})	.80080 (10^{-12})	.65324 (10^{-12})

[a] Parameters of prior distributionS: r=.03735, k=2.5044; m =89.73901.
[b] Parameters of prior distributionS: r=.03735, k=2.5044; m=70.00001.
[c] Parameters of prior distributionS: r=.03735, k =2.5044; m=19.73901.
[d] Parameters of prior distributionS: r=.03735, k=2.5044; m=.00001.

Table 3. Posterior Probability of Group-membership, given Representative Individual Latency Sets.

	Representative individual-participant-latency sets			
	Paranoid schizophrenia		Control	
Classification[e]	High-encoding load {3.476, 7.150, 9.00, 4.238}	Low-encoding load {2.177, 3.184, 2.411, 2.498}	High-encoding load {1.137, 3.204, 1.371, 2.207}	Low-encoding load {1.037, 0.554, 0.191, 6.390}
Paranoid schizophrenia				
High-encoding load[a]	0.52528 (10^{-6})	0.022	0.27083 (10^{-7})	0.730536(10^{-29})
Low-encoding load[b]	0.00012	0.9511	0.00007	0.23473 (10^{-20})
Control				
High-encoding load[c]	0.99987	0.027	0.99993	0.188326
Low-encoding load[d]	Pr{g} → 0.0	0.23972 (10^{-11})	0.200186 (10^{-12})	0.81167

[a] Parameters of prior distributions: r=.03735, k=2.5044; m =89.73901.
[b] Parameters of prior distributions: r=.03735, k=2.5044; m=70.00001.
[c] Parameters of prior distributions: r=.03735, k =2.5044; m=19.73901.
[d] Parameters of prior distributions: r=.03735, k=2.5044; m=.00001.
[e]Base rates for group membership = 0.05 for Paranoid Schizophrenia/High encoding load; 0.20 for Paranoid Schizophrenia/Low-Encoding load; 0.5 for non-patient Controls/High-encoding load; and .25 for non-patient Controls/Low-encoding load.

Table 4. Process-latency base distributions and their moments with a gamma-distributed "capacity" (intensity) parameter, *v*.

Base Distribution	Distinguishing Features	nth Moment		Comments
		$k > n$	$k = n^1$	
Erlang	*k'* stages fixed, each with constant rate *v*	$\frac{r^n (k'+n-1)!\Gamma(k-n)}{(k'-1)!\Gamma(k)}$	$\frac{r^k (k'+n-1)!\ln(r+t)}{\Gamma(k)(k'-1)!}\Big]_{t=0}^{t=\infty} = \infty$	For the exponential distribution, *k'*=1.0
Compound Poisson	*k'* stages random, *v* constant across stages	$(-1)^n d^n M_T(\theta)/d\theta^n\|_{\theta=0}$ where $M_T(\theta) = \int_0^\infty f(v)[e^{m(M_T(\theta)-1)}\|v]dv$, $M_T(\theta) = \frac{v}{(v+\theta)}$, $f(v) = \frac{(rv)^{k-1} re^{-rv}}{\Gamma(k)}$, and *m* is the parameter of the Poisson distribution of *k'*. E.g., $E(T) = \frac{mr}{(k-1)}$; $Var(T) = \frac{r^2 m(2(k-1)+m)}{(k-1)^2(k-2)}$	$\infty(1-e^{-m}) = \infty$.	$E(T^n\|k') = \infty$, by results for the Erlang distribution above. $E(T^n\|k') = \infty$, hence $E(T^n) = \infty$, also by results for the General Gamma distribution, below.

Base Distribution	Distinguishing Features	nth Moment		Comments
		$k > n$	$k = n$	
General Gamma	Stage-wise rate transition	$\sum_{i=1}^{k'} \frac{C_{ik'} c_i^n \, ! \, r^n \Gamma(k-n)}{\Gamma(k)}$ where c_i is a scalar of v for the ith stage, $c_i > 0$; the rate thus in effect for stage i is $c_i v$; and $C_{ik'} = \left(\prod_{\substack{j=1 \\ j \neq i}}^{k'} 1 - \frac{c_i}{c_j} \right)^{-1}$, $c_i \neq c_j$.	∞, as $\Gamma(k\text{-}n) = \Gamma(0)$.	
Pure Death General Gamma Process with Linear Death Rate	Rate for the ith stage is $c_i v$, where $c_i=(k'\text{-}(i\text{-}1))$; i=1, 2, ..., k'.	$\sum_{i=1}^{k'} (-1)^{i+1} \binom{k'}{i} \frac{r^n n! \Gamma(k-n)}{\Gamma(k) i^n}$	$\sum_{i=1}^{k'} (-1)^{i+1} \binom{k'}{i} \frac{i r^k n! \ln(r+it)}{\Gamma(k) i^{n+1}} \Big]_{t=0}^{t=\infty}$ $= \sum_{i=1}^{k'} (-1)^{i+1} . \infty = \infty.$	
Weibull	Single stage; rate is continuous function of time.	$\frac{r^n \Gamma(n/\beta+1) \Gamma(k-n)}{\Gamma(k)}$, where $\beta>0$.	∞, as $\Gamma(k-n) = \Gamma(0)$.	

[1] Expression of moment infinity as Γ(0) or ln(∞) merely rests on integration with respect to t first, or v first, respectively.

Table 5. Moments of Weibull and Erlang Process-Latency Base Distributions, with Capacity Parameters either $\text{Weibull}_{V,B}$ or $\text{Gamma}_{r,k}$ distributed.

Mixing Distribution	Base Distribution		
	Weibull; v is Randomly Mixed	Weibull; v^{β} is Randomly Mixed	Erlang; v is Randomly Mixed
Weibull	$V^n\Gamma(n/\beta+1)\Gamma(1-n/B)$	$V^{n/\beta}\Gamma(n/\beta+1)\quad\Gamma(1-n/(\beta B))$	$\frac{V^n\Gamma(k'+n)\ \Gamma(1-n/B)}{(k'-1)!}$
Gamma	$\frac{r^n\Gamma(n/\beta+1)\Gamma(k-n)}{\Gamma(k)}$	$\frac{r^{n/\beta}\Gamma(n/\beta+1)\Gamma(k-n/\beta)}{\Gamma(k)}$	$\frac{r^n(k'+n-1)!\ \Gamma(k-n)}{(k'-1)!\Gamma(k)}$

Endnotes

1. A parameter is a variable in a mathematical expression, whose values can change without altering the structure of the expression in which it is embedded (cf. Borowoski & Borwein, 1989, p.435).

2. Consider, as a representative example, numerical versus symbolic integration, a special case being numerical solutions to differential equations composing nonlinear dynamical ("chaos-theoretic") systems, as contrasted with desired (but most often intractable) explicit solutions of the "initial-value problem" (Koçak, 1989).

3. Formal statements foster place-keeping regarding the current state of progress in the problem area. They make clear what has been formulated, whether or not new formal proposals differ, and if so, in what respects. Blind alleys are flagged, and promising leads discerned. The place of precedent is usually more fuzzy in the case of verbal statements. Even if they seem to rediscover pre-existing ideas, subsequent proposals almost always escape indictment for ignoring antecedents, and if taken to task, usually appeal to allegedly important nuances of difference.

4. These explicit starting points can always be challenged, precisely because they are explicit. Informal theories escape because they are not forced to lay bare their assumptional moorings before commencing (cf. Bandura, 1984; Staddon, 1984).

5. Neuroconnectionist modeling, entailing computer simulations of task-transacting activations among neuronal units, or unit modules, (Rummelhart, McClelland, et al, 1987), arguably has received a warmer welcome (including in the field of clinical cognitive science), because of face validity as a "brain metaphor" (notwithstanding certain challenges to biological plausibility), and because of commercial software expediting use of simulation algorithms virtually since their introduction (McClelland, Rumelhart, et al, 1986; cf. Carter & Neufeld, in submisison).

6. Added to considerations of response properties of interest, are those of their distributional properties (e.g., moments of performance-latencies, including means, and inter-trial variances, illustrated above; distribution functions, as exemplified in equation (2), above; and similar properties). The concept of "target levels of inference, or "levels of analysis", and associated boundaries of inference are underscored. Model properties and response parameters entailing task -performance latency, for example, are deemed to express dynamical properties of operations occurring at other strata, including algorithmic (neuroconnectionist), and implementational (neurophysiological; Busemeyer & Townsend, 1993, pp. 443-445; Marr, 1982). The inter-strata associations are isomorphic, in the way that a computer-machine language, and computer-programming language, are co-extensions of one another. Targeting one or the other stratum then becomes a matter of the nature of the focal problem.

7. The performance sample is represented as a set of observations, as opposed to either an ordered set, or response vector. The observations are deemed to be independent, and algebraic operations to which they are subjected does not depend on their ordering.

8. It is enticing to entertain a further mixture, of the existing mixture on *v*. The parameter *r,* for one may be considered as random rather than fixed. Stressor effects may be such as to shift the stochastic distribution, rather than a single value of *r* upwards. The net effect would still be one of increasing latencies because the influence on the distribution of capacity values *v* would be downward. Allowing the distribution of *r* to be gamma, analysis indicates that an analytical solution for the density function for $t, f(t)$ now becomes intractable. Within this distributional framework, then, there appears to be a limit to the stacking of one source of indeterminacy onto another. Tractability demands at a given point the certainty conferred by fixed parameter status. By extension, plausibility of a processing system may tolerate a limited compounding of sources of instability.

9. Although compelling in other respects to be sure, methodology developing from item-response theory (IRT) can be shown not to resolve the above dilemma, either in its own right, nor via IRT-informed measure composition (see, e.g., Santor & Ramsey, 1998).

Appendix A
Bayesian Platform for the Current Developments

Expansion of $\{Pr(g|\{*\}\}$ and $[w(\theta'|\{*\}, g\,]$ of Equation (5) results in the following expression:

$$\sum_{g=1}^{G} \Big(\{\, 1/[\sum_{g=1}^{G} Pr(g)\ ujd(\{*\}|g)] \quad Pr(g)\ ujd(\{*\}|g)\} \ [w(\theta'|g)\ cjd\ (\{*\}|\theta')/\ ujd(\{*\}|g)] \Big).$$

In this expression, $Pr(g)$ is the current base rate of group g, corresponding to the relative frequency of members of g in the participant sample at large; $ujd(\{*\}|g)$ is the unconditional joint density of the obtained latency data set, given group g, meaning the combined probability density of $\{t_1, t_2, \ldots t_N\}$, given group g, all values of θ considered; $w(\theta'|g)$ is the probability (density) of θ', given group g; and $cjd(\{*\}|\theta')$ is the conditional joint density of the obtained latency data set, given parameter-value θ'.

Note that the above expansion reveals two sets of Bayesian priors; one comprises the group-wise mixing distribution of θ. The other comprises the discrete, finite mixing probabilities of the group-wise priors being operative, corresponding to the base rates $Pr(g)$. The latter is a "meta-prior" of this "compound Bayesian formulation". Note as well that the densities of θ, including that of any specific value θ', are determined by the group at hand's prior distribution of θ.

Where θ becomes a set of parameters (e.g., v and k'; see text), each with its own independent group-related mixing distribution, one or another member of the set may be considered separately. Separate consideration occurs, for example, when constructing a posterior parameter mixing distribution following acquisition of a performance sample {*}, as exemplified in Figure 7. Computations entail aggregating over values of the set-aside parameter(s) according to its (their) group-prescribed prior probabilities (densities), while addressing specific values of the parameter that has been singled-out, also according to its group-specified prior probabilities (densities). Computational details undergirding the present developments have been presented in Neufeld, et al (2002).

Appendix B
Estimation of Base Rates of Symptom Severities

The joint probability of an array of J classifications with respect to G groups, by equation (5) and observations in **Appendix A** is

$$\prod_{j=1}^{J} \Pr(g|\{*\})_j$$

$$= \prod_{j=1}^{J} Pr(g)\ ujd(\{*\}|g)_j \,/ \sum_{g=1}^{G} Pr(g)\ ujd(\{*\}|g)_j . \qquad \text{(B.1)}$$

This product can be show to be a reduction of the joint multinomial probability of the J observed classifications under the model $Pr(g|\{*\})$, each probability being an expression of the multinomial likelihood of classification j (see, e.g., García Pérez, 1994 ; Reifer & Batchelder, 1988).

Maximum-likelihood estimates of the $Pr(g)$, g = 1,2, ..., G - 1 (recalling that $Pr(G - 1) = 1 - \sum_{g=1}^{G-1} Pr(g)$) are obtained by differentiating (B.1) with respect to each term in question, setting each derivative to 0, and simultaneously solving the G - 1 equations in the G - 1 unknowns. A numerical search algorithm will be required to replace the analytical solutions as computations become unwieldy with increasing J and/or G.

Note that equation (B.1) beguilingly invites a χ^2 test for empirical goodness of fit of the prevailing model of $Pr(g|\{*\})$. Specifically, -2 ln(B.1), which here can be shown to equal -2 ln (B.1)/1.0, is tantamount to -2 ln(multinomial-likelihood ratio), an expression that stands to be distributed as χ^2 with its associated degrees of freedom (cf Carter & Neufeld, 1999). In fact, this expression is dubiously distributed as $\chi^2_{df=J-(G-1)}$ in the present case because the current data array suffers from the problem of "extreme sparseness of cell-wise observations" (see Delucchi, 1993; Tollenaar & Mooijaart, 2003).

Appendix C

C.a. Values of Distribution Moments

The *n*th moment of a distribution of latencies *t*, including those of the mixtures discussed in the text, will be infinite if its integral does not converge:

$$\int_0^\infty f(t)\, t^n\, dt = \infty.$$

Equivalently,

$$(-1)^n\, E(T^n) = d^n/d\theta^n\, M_T(\theta)|_{\theta=0} = \infty.$$

Here, $M_T(\theta)$ is the moment-generating function

$$\int_0^\infty f(t)\, e^{-\theta t}\, dt,$$

whose *n*th derivative with respect to θ encounters a singularity at $\theta = 0$.

$E(T^n)$ also can be written as

$$n \int_0^\infty (1 - F(t))\, t^{n-1}\, dt$$

$$= \quad n \int_0^\infty \bar{F}(t)\, t^{n-1}\, dt,$$

where $F(t)$ is the cumulative probability distribution function (see Figure 12 and related text), and $1 - F(t) = \bar{F}(t)$ is the survivor function. This version of $E(T^n)$ relates distribution moments, and their availability as finite values, to "process-completion potential". The latter is defined as

$$\int_0^t f(t')/\bar{F}(t')\, dt', \qquad \text{(C.a.1)}$$

where $f(t')/\bar{F}(t')$ is the hazard function $H(t')$, indicating the conditional rate of process completion at time t', given non-completion by t'. Because $\bar{F}(t) = \exp(-\int_0^t (H(t')\, dt'),$ (C.a.1) readily is seen to be simply $-\ln(\bar{F}(t))$ (see discourses of Townsend & Ashby, 1983; Wenger & Townsend, 2000). This expression of process-completion potential, and its association with $E(T^n)$ is called upon in developments presented in the text under *Interplay with Capacity Utilization*, and Appendix section **C.c**, below.

C.b Moments of Selected Base Distributions Mixed on a Gamma-Distributed "Capacity" (Intensity) Parameter

Mixture-model latency distributions considered in this section are composed of alternate process-latency base distributions, whose "capacity" (distributional intensity or scale) parameters are randomly distributed according to a gamma distribution whose own intensity or scale parameter is denoted r and whose shape parameter is denoted k. Moments are computed for the case where $k>n$, n being the order of the moment, and where $k=n$.

Three of the base distributions, the exponential, Erlang, and Weibull are nested in the Generalized Gamma Distribution, whose probability density function is

$$\beta(vt - va)^{\beta k' - 1}/(\Gamma(k')) \; v \exp(-(vt - va)^{\beta}). \qquad \text{(C.b.1)}$$

The probability density function for the Erlang distribution is obtained from (C.b.1) by setting β equal to 1.0, and a equal to 0. The density function for the Weibull distribution results when $k' = 1.0$ and $a = 0$. The exponential distribution, in turn is nested in both the Erlang and Weibull distributions, as its density function is derived from (C.b.1) when $k' = \beta = 1.0$, and $a = 0$. Loosely speaking, each base distribution has a conjugate prior, in that both prior and base entail exponentials, making for mathematical tractability (see, e.g., Geddes & Scott, 1989).

The probability density function of mixture-model latencies $f(t)$ for the Erlang base distribution, with its capacity (also rate) parameter v being distributed as $\text{gamma}_{r,k}$ is

$$\frac{r^k t^{k'-1}\,\Gamma(k+k')}{\Gamma(k)(k'-1)!\;(r+t)^{k+k'}} \quad .$$

Then $\int_0^{\infty} f(t)t^n dt \;=\; E(T^n)$

$$= \qquad \frac{r^n\;(k'+n-1)!\Gamma(k-n)}{\Gamma(k)(k'-1)!} < \infty \qquad \text{if } k>n$$

$$\left\{ \qquad \frac{r^k(k'+n-1)!\ln(r+t)}{\Gamma(k)(k'-1)!}\right]_{t=0}^{t=\infty} = \infty \qquad \text{if } k=n$$

The next base distribution is the compound Poisson (e.g., Feller, 1966; Ross, 1996), of which the special case at hand is composed of an Erlang distribution, whose stage (distributional shape) parameter *k*' is Poisson distributed, with parameter *m* (see Figures 5 and 6, and equation (8) of text, and related prose). The Erlang rate parameter of this base distribution in turn is gamma distributed, with parameters *r* and *k*. A concise method of obtaining the moments of this second-order mixture capitalizes on its moment-generating function $M_T(\theta)$; the *n*th-order derivative of the latter, evaluated at $\theta = 0$, and scaled by $(-1)^n$ yields the *n*th-order moment (for an especially lucid and still current exposition of the moment-generating function, see Kenny & Keeping, 1951).

Proceeding accordingly,

$$M_T(\theta) = \int_0^\infty f(v)\ (\mathrm{M}_\mathrm{T}(\theta)|v)\ dv$$

$$= \int_0^\infty f(v) \sum_{k'=0}^{k'=\infty} \frac{m^{k'}}{k'!}\ e^{-m}(M_T(\theta)|k' \cap v)\ dv$$

$$= \int_0^\infty f(v) \left[\sum_{k'=0}^{k'=\infty} \frac{m^{k'}}{k'!}\ e^{-m}(M_{T_i}^{k'}(\theta) - e^{-m})|v \right] dv$$

$$= \int_0^\infty f(v) \left[(e^{m(M_{T_i}(\theta)-1)} - e^{-m}|v \right] dv,$$

where $M_{T_i}(\theta)$ is the moment generating function for each of the *k*' stage intercompletion times, or $v/(v+\theta)$, for the Erlang distribution with rate parameter *v*. As the *k*' intercompletion times are identically and independently distributed for a given value of *v*, $M_T(\theta)|k' \cap v = M_{T_i}^{k'}(\theta)|v$. Note that

$$M_T(\theta)|(k'=0) = 0.$$

Thus,

$$M_T(\theta) = \int_0^\infty \frac{(rv)^{k-1}}{\Gamma(k)}\ re^{-rv+m(v/(v+\theta)-1)} dv - e^{-m}$$

and

$$E(T^n) = (-1)^n d^n M_T(\theta)/d\theta^n|_{\theta=0}.$$

The n^{th} order moment of this second-order compound Poisson distribution will be infinite if *k*=*n*, as follows.

$$E(T^n) = \sum_{k'=0}^{k'=\infty} \frac{m^{k'}}{k'!} e^{-m} E(T^n|k')|_{k=n}$$

$$= \sum_{k'=0}^{k'=\infty} \frac{m^{k'}}{k'!} e^{-m} .\infty \; - e^{-m} .\infty \Big|_{k=n}$$

(by results for the Erlang base distribution, above)

$$= e^{m} \; e^{-m} \; . \; \infty - e^{-m} \; .\infty \Big|_{k=n} \; = \; \infty \; (1 - e^{-m}) \; \Big|_{k=n}$$

$$= \infty \Big|_{k=n} \; .$$

By similar reasoning, $E(T^n) = \infty|_{k=n}$ also for the general-gamma base distribution, below.

Turning to the general gamma distribution, unlike that of the Erlang, the rate parameter is stage-specific. In the present implementation, the rate for stage i is $c_i v$, $c_i > 0$, $i = 1,2, \ldots, k'$. For this distribution,

$$M_T(\theta) \; = \; \int_0^{\infty} f(v) \; M_T \; (\theta) | v \; dv$$

$$= \int_0^{\infty} \frac{(rv)^{k-1}}{\Gamma(k)} \; re^{-rv} \sum_{i=1}^{k'} \; \frac{C_{ik'} \; c_i v}{c_i v + \theta} \; dv$$

where C_{ik} is defined in Table 4 of the text.

$$E(T^n) \; = \; (-1)^n \; d^n / d\theta^n \; M_T(\theta) \Big|_{\theta=0}$$

$$= \sum_{i=1}^{k'} \; C_{ik} \; c_i \; n! \; r^n \; \frac{\Gamma(k-n)}{\Gamma(k)} \; .$$

For the special case, where $c_i = (k'-(i-1))$, the density function $f(t)$ can be computed from developments supplied by Morrison (1979), as

$$\sum_{i=1}^{k'} \; \frac{(-1)^{i+1} \; \binom{k'}{i} \; r^k \; i \; k}{(r+it)^{k+1}} \; .$$

From here, it is a simple matter to obtain

$$E(T^n) \; = \; \int_0^{\infty} f(t) \; t^n \; dt$$

$$= \begin{cases} \sum_{i=1}^{k'} (-1)^{i+1} \binom{k'}{i} r^n n! \frac{\Gamma(k-n)}{\Gamma(k)\, i^n} < \infty & \text{if } k>n \\ \sum_{i=1}^{k'} (-1)^{i+1} \binom{k'}{i} \frac{i\, r^k\, n!}{\Gamma(k)\, i^{n+1}} \ln(r+it) \Big]_{t=0}^{t=\infty} = \infty & \text{if } k=n \end{cases} .$$

The probability-density function for the Weibull distribution is

$$\beta v(vt)^{\beta-1} \exp(-(vt)^{\beta}),$$

whereby $E(T^n|v) =$

$$\Gamma(n/\beta + 1)/v^n.$$

$E(T^n)$, in turn, is readily available as

$$\int_0^{\infty} f(v)\, E(T^n|v)\, dv$$

$$= \frac{r^n\, \Gamma\,(n/\beta\, +\, 1)\, \Gamma(k-n)}{\Gamma(k)} \quad .$$

which again is severed from finite values if $k=n$.

C.c. Formulae underlying Figure 13

For the nonhomogeneous Poisson process, with dynamic rate $vt^{\beta-1}$, v having been $\text{gamma}_{r,k}$ distributed,

$$f(t) = \frac{\beta r^k\, \Gamma(k'+k)\, t^{\beta k'-1}}{(k'-1)!\, \Gamma(k)\, (t^{\beta}+r)^{k'+k}}$$

and

$$\bar{F}(t) = \sum_{j=0}^{k'-1} \frac{\Gamma(j+k)\, r^k\, t^{\beta j}}{j!\, \Gamma(k)\, (t^{\beta}+r)^{j+k}}.$$

The hazard function $H(T) = f(t)/\bar{F}(t)$, enters into the index of process-completion-potential as follows: $\int_0^t H(t')\, dt'$. Happily this integral is available as $-\ln(\bar{F}(t))$, especially considering the possibility of complex $f(t)$ and $\bar{F}(t)$, such as those above (see Appendix **C.a**).

Figure 1. Probability of completing *k*' stages in 4 time units (lower surface), and 7 time units (upper surface), with each stage transacted at rate *v* (see Equation 2 in text); *k*' ranges from 1 to 10, and *v* ranges from .01 to 3.0.

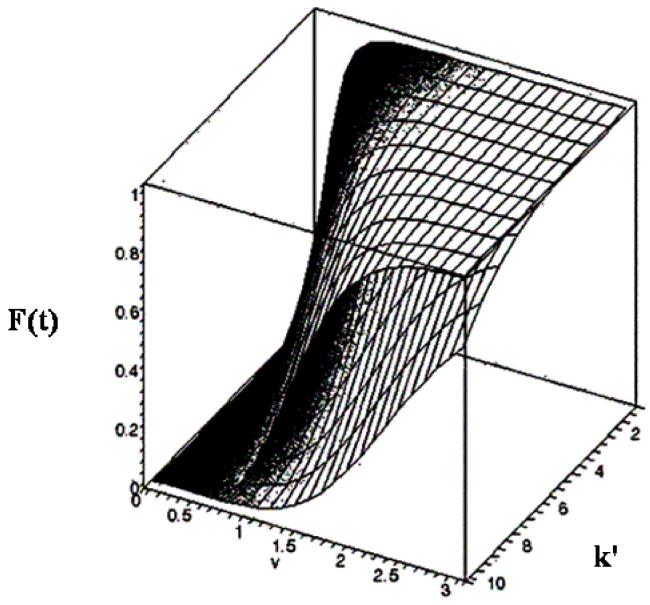

Figure 2. Transitions from basic cognitive science among normals to cognitive assessment among patients.

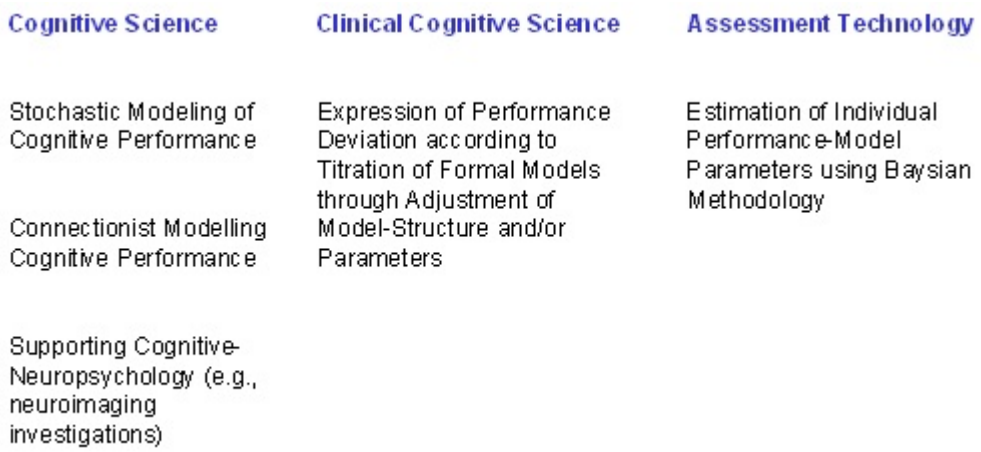

Figure 3. Erlang distribution, with various pairs of its parameters.

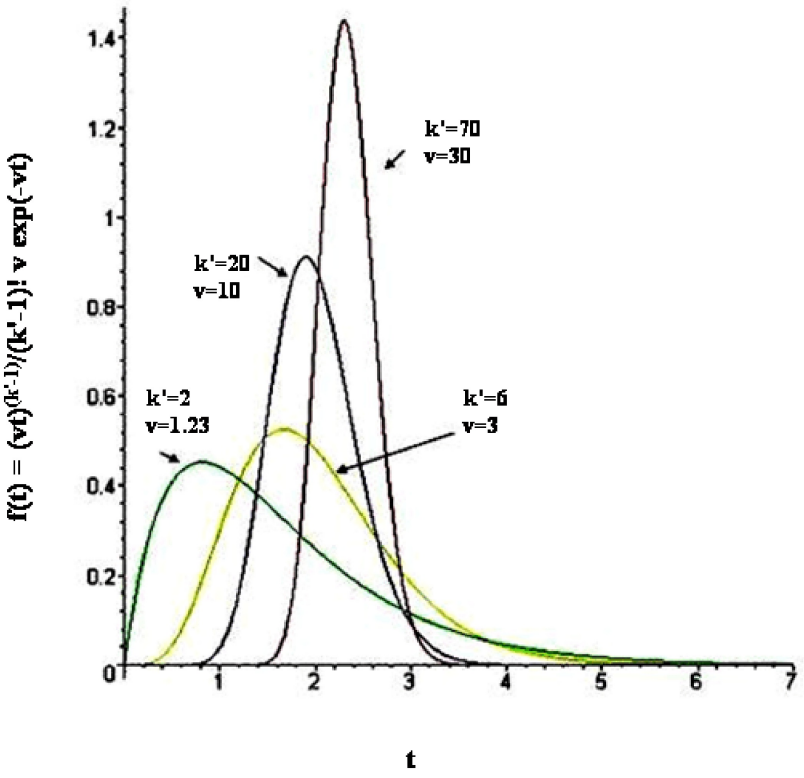

Figure 4. Gamma distribution of *v*.

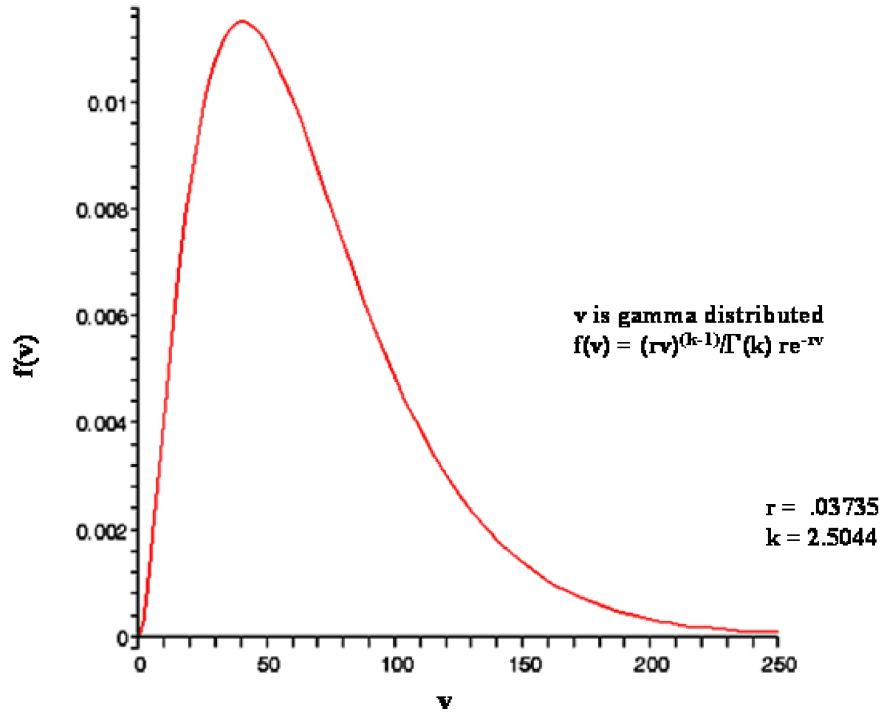

Figure 5. Poisson distribution of *k*' (smoothed), for differing values of *m*.

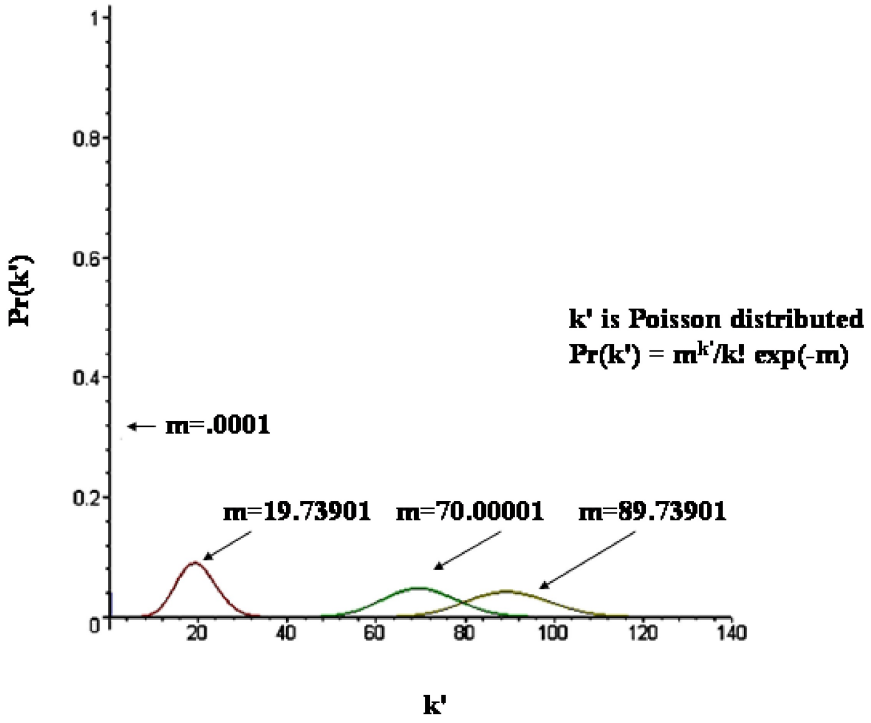

Figure 6. Design of mixture model.

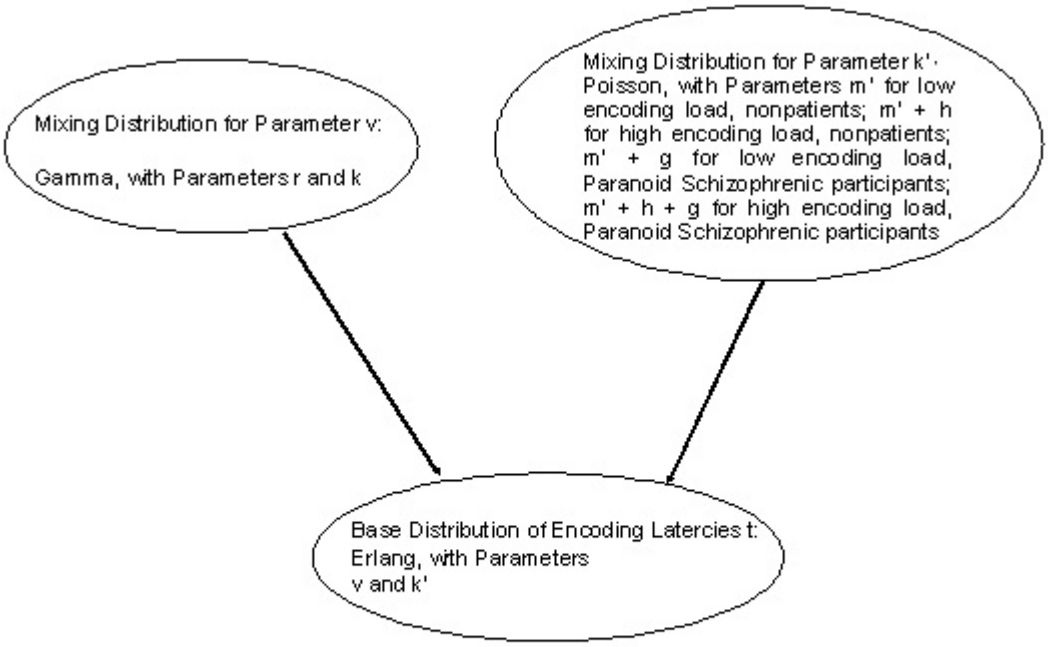

Figure 7. Posterior distribution of *k'*, *Pr*(*k*'|{1.037, .554, .191, 6.39}; *m* = 70.00001), and prior distribution of *k'*, *Pr*(*k*'; *m* = 70.00001).

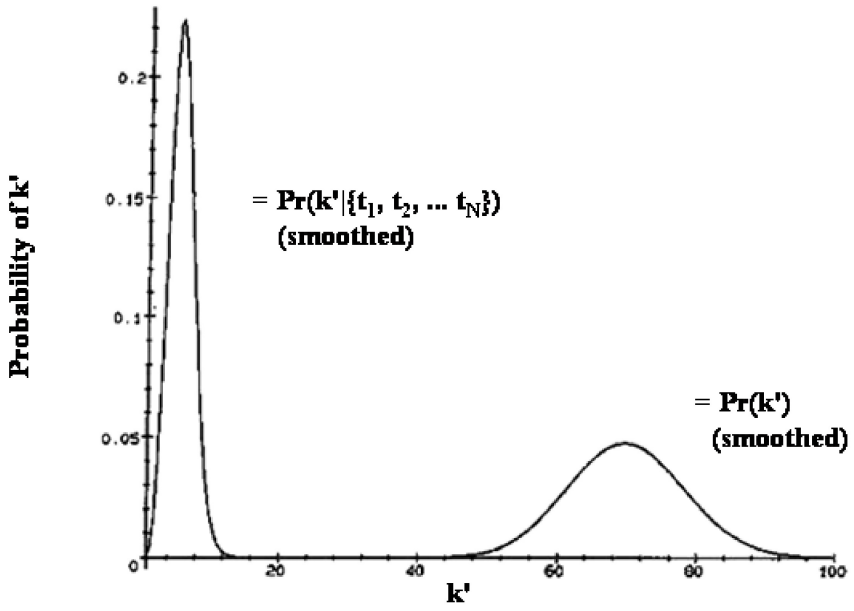

Figure 8. Posterior latency distributions $f(t|\{2.177, 3.184, 2.411, 2.498\}; g)$.

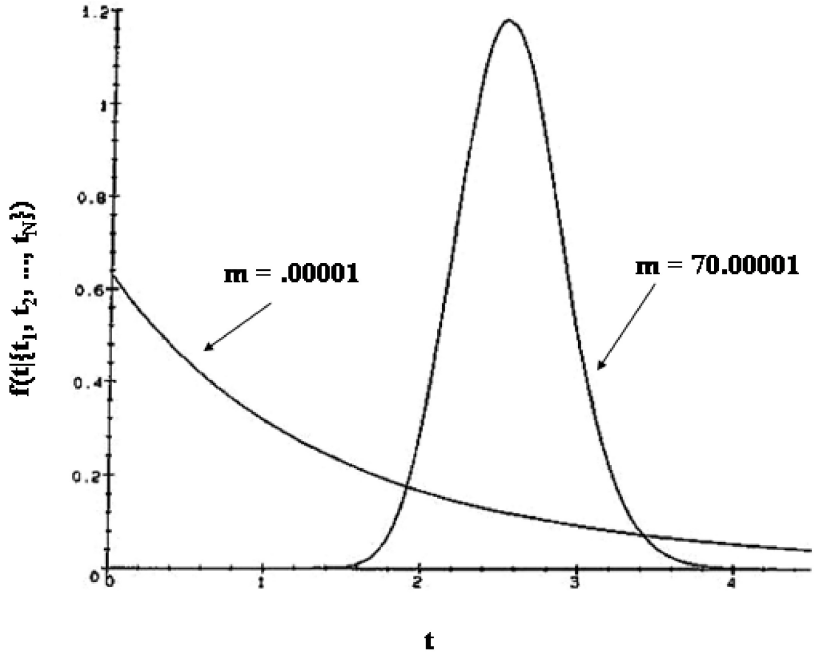

Figure 9. Posterior latency distributions $f(t|\{2.177, 3.184, 2.411, 2.498\};g)$.

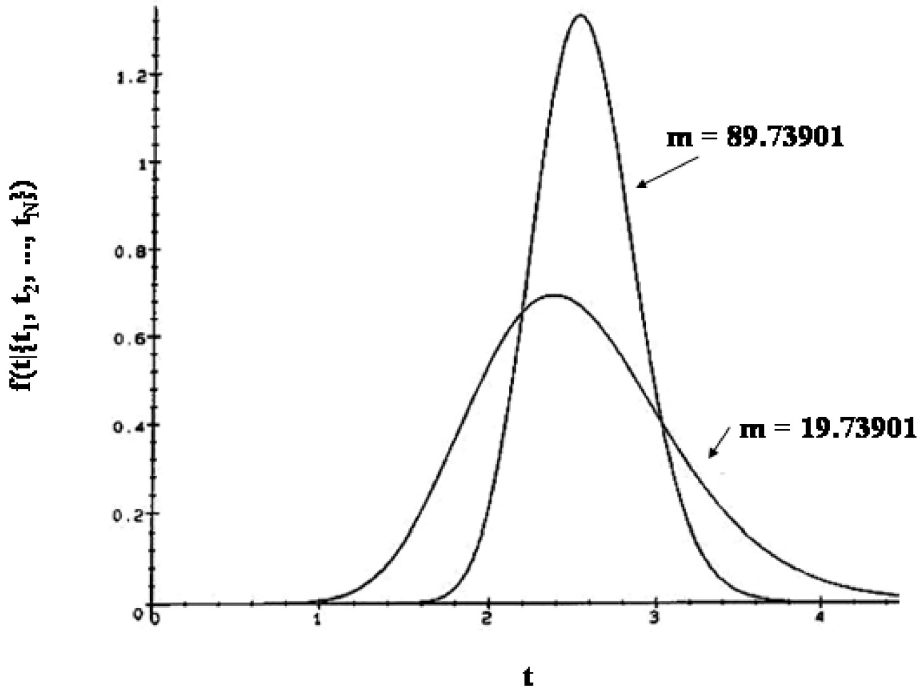

Figure 10. Mixture-model latency distribution $f(t|g)$.

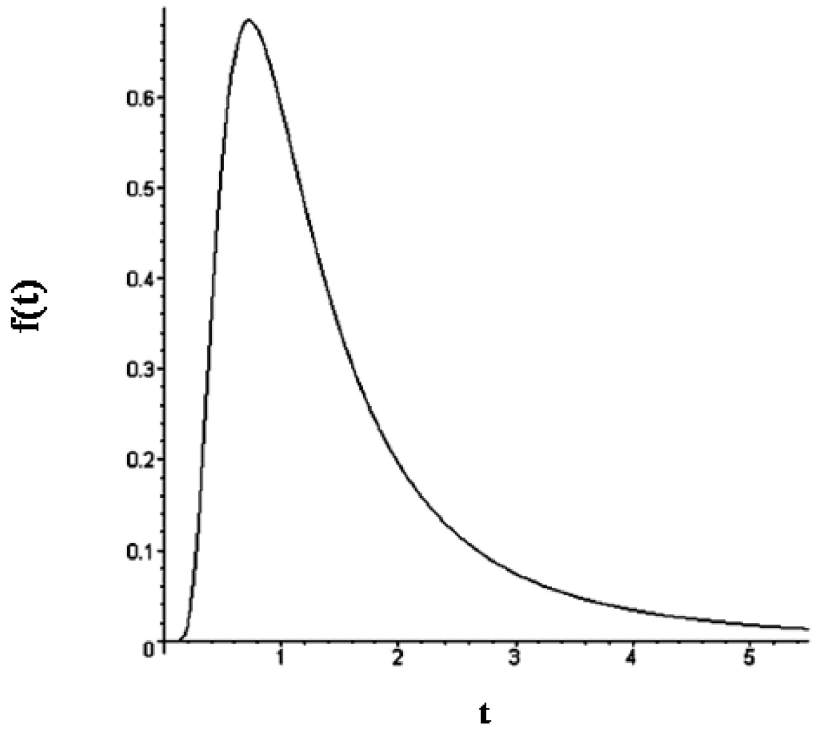

Figure 11. Mixture-model latency distribution $f(t|g)$.

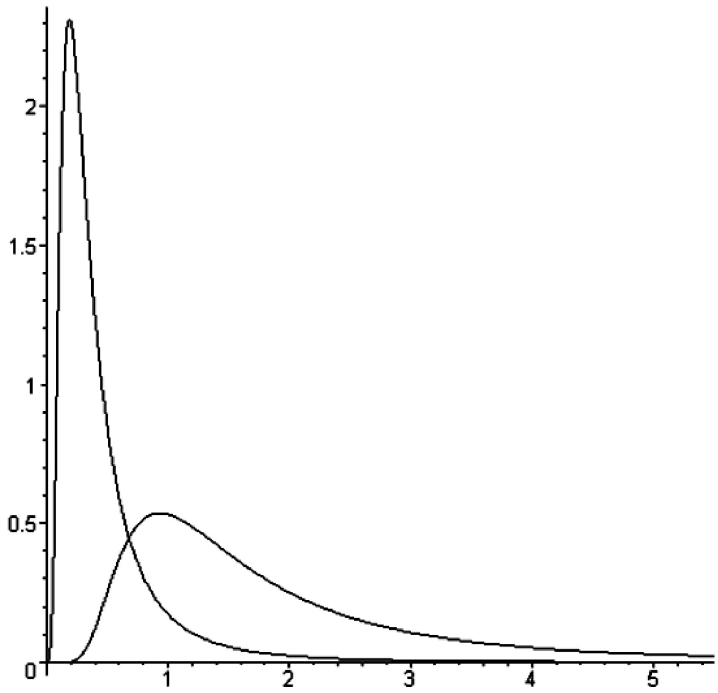

Figure 12. Cumulative probability distributions corresponding to probability-densities of Figure 9.

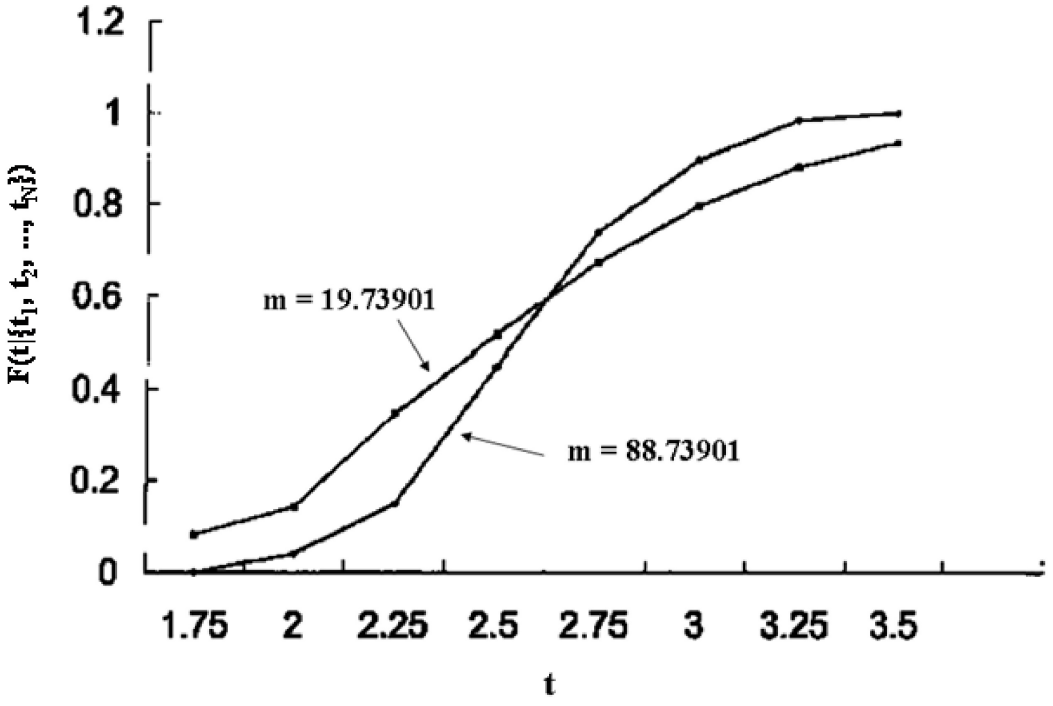

Figure 13. Values of process-completion potential -ln($F(t)$) for the nonhomogeneous Poisson distribution, with rate parameter $vt^{\beta-1}$ and v gamma$_{r,k}$ distributed, for k varying from 1 to 3 and β varying from 1 to 5; r = .03735; k' = 18; t = 2.25.