Behavioral Development Bulletin
2015, Vol. 20, No. 2, 137–149

1942-0722/15/$12.00 http://dx.doi.org/10.1037/h0101380

Analyzing Single Subject Data for Showing Intervention Effectiveness

Michael Lamport Commons
Harvard Medical School

Patrice Marie Miller
Harvard Medical School and Salem State University

Leonard Sidney Miller
University of California, Berkeley

Although individual charting can be an effective way to demonstrate progress, it does not allow for comparisons of effectiveness using traditional statistical standards. Due to the increasing need for evidence of effectiveness of interventions it is important that there be a way to compare interventions. In this paper a model of change in behavior along a behavioral-developmental sequence is proposed and assessed, and how it can be used to evaluate interventions is demonstrated. First, an individual's progress is documented along a behavioral-developmental sequence, using the model of hierarchical complexity (MHC). A behavioral aim can then be selected and behavior can be tracked depending on whether developmental tasks are completed. This paper then lays out a statistical model for combining sections of charts. This model may be generalized to take into account charts of tasks of different difficulties due to stage subtask difficulty and subsubtask difficulty, as well as individual differences and subdomain differences. It can also be generalized to charts of different people's performances, and to different chart supervisors and programs. This is simply done by adding more independent variables to the model. The implications for using this method to evaluate interventions are discussed.

Keywords: charting, precision teaching, group designs, policy, evaluation

The current methods of showing effectiveness on the basis of individual charting are compelling in their own right. However, they do not meet generally agreed upon standards for effectiveness used by behavioral scientists. The major problem is there is not a statistical method of evaluating successful behavior based on charts. That makes it impossible to analyze the effectiveness of interventions across people. Combining individually-based data with group data will allow for greater communication with other scientists and with policymakers.

There are a number of reasons group data are needed. Most importantly, there is a policy issue. How effective is a program of interventions? Increasingly, interventions, whether purely medical or behavioral, will have to be justified in terms of their effectiveness. The increasing need for studies demonstrating efficacy is largely due to new policies on evidence-based practice in health care settings. For example, some devices used in intervention programs require FDA and other agency approval. Professional standards and accreditation are also being increasingly applied to education and educational programs. These programs rely heavily on transparency for effectiveness and will want to monitor their own progress and success. They will also seek ways to communicate that success to others. Combining individ-

Michael Lamport Commons, Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School; Patrice Marie Miller, Beth Israel Deaconess Medical Center, Harvard Medical School and Department of Psychology, Salem State University; Leonard Sidney Miller, School of Social Welfare, University of California, Berkeley.

We thank Joseph Pear, whose comments fundamentally changed the direction of the paper for the better. Additional thanks to Dristi Adhikary and Kyle Featherston for their arduous work in editing through numerous versions.

Correspondence concerning this article should be addressed to Michael Lamport Commons, Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, 234 Huron Avenue, Cambridge, MA 02138-1328. E-mail: commons@tiac.net

ual chart data into group data makes it possible to apply standard statistical tests to interventions so that statistical significance about efficacy may be assessed.

This paper lays out a statistical model for combining sections of charts. This model may be generalized to take into account charts of tasks of different difficulties due to developmental stage, as well as individual differences, group differences, and subdomain differences. This is done by simply adding independent variables. This allows for great flexibility of what kind of individual, tasks, interventions, and group comparisons may be made.

As part of the process of converting individual charts to group data, this paper argues that the first step is to situate individual charts within a behavioral developmental sequence. In addition to the need for group data discussed above, evaluating progress in terms of development has further benefits. The acquisition of individual, possibly helpful behaviors is clearly important. However, showing overall progress in terms of sequences of behaviors that are acquired is also important. A set of behaviors from the same point in a developmental sequence may be acquired, but such acquisition does not always represent development. Development is when there is movement up in the behavioral developmental task sequence. Through developmental testing, one can tell if the participants actually develop, or if they just acquire, some isolated and possibly helpful behavior. Technically, hierarchical complexity of items predicts with up to an $r = .985$ whereas simple horizontal complexity predicts with an $r = .7$ (Commons, Giri, & Harrigan, 2014). How behavioral development sequences are defined by the MHC is discussed next.

The Model of Hierarchical Complexity

The MHC is a nonmentalistic, neo-Piagetian and *quantitative behavioral development theory*. This model has been applied widely around the world (Bernholt, Parchmann, & Commons, 2009; Commons et al., 2008; Commons et al., 2005; Dawson, 2002). It offers a standard method of examining the universal pattern of development of increasingly successful completions of tasks, which, in some domains, could be referred to as "smartness." It shows that development proceeds across general *sequences* of behavior. These sequences exist in every *subdomain* including social, interpersonal, mathematical, logical, scientific, moral, and so on.

The Three Definitions of the MHC

Order of hierarchical complexity (OHC) is an analytic measure applied to tasks. Higher order actions: (a) are defined in terms of two or more task actions from the next lower OHC, (b) organize two or more next lower order complex actions, and (c) are carried out in a nonarbitrary way. The notion of combining elements as discussed in definition (a), is similar to Binder's combinations of elements into tasks (Binder, 2000). Applying these definitions we have shown that tasks can be ordered from simpler to more complex. Because the more complex tasks are composed of the simpler ones, the simpler ones must be acquired first. The MHC identifies 17 orders of hierarchical complexity (Table 1). Based upon the OHC of a task that an organism successfully completes, they are then said to perform at a particular *behavioral stage of development*. The behavioral stage of development has the same name and number as the

Table 1
Orders of Hierarchical Complexity

Order number	Order name	Order number	Order name
0	Computational	9	Concrete
1	Automatic	10	Abstract
2	Sensory or Motor	11	Formal
3	Circular Sensory Motor	12	Systematic
4	Sensory-Motor	13	Metasystematic
5	Nominal	14	Paradigmatic
6	Sentential	15	Crossparadigmatic
7	Preoperational	16	Meta-crossparadigmatic
8	Primary		

OHC of the task that it correctly completes. However, there is a difference between the OHC of tasks and the corresponding stage of performance on those tasks. *Stage* is an operationally defined performance measure of the most hierarchically complex task in a subdomain solved by the person in question.

Because the MHC is content- and context-free, and is not a direct measure of participant behavior but of "ideal" task actions, it applies to all participants, including people, animals, and machines. This means that it should be possible to aggregate performance of any individual on different charts. A more detailed explanation of how to identify OHC of different tasks will occur later in the paper, but first we introduce the standard method of tracking behavior in interventions, and why it can benefit from a developmental perspective using the MHC.

The Celeration Chart

The paper will now introduce a sample celeration chart for a single individual (Figure 1). The chart makes it possible for the intervener to quickly assess if a person's performance (changes in behavior) is accelerating or decelerating through time. The term celeration describes both. If the behavior is desirable, the measure of improvement is the amount of positively changing behavior (acceleration). An *aim* is the goal for a terminal rate of performance. The rate would go up if the aim for ultimate rate was met. If the behavior is not desirable, the measure of improvement is the amount of decrease in undesirable behavior (deceleration). By seeing the rate of improvement or the lack thereof, interveners may quickly adjust the task on which the person is working.

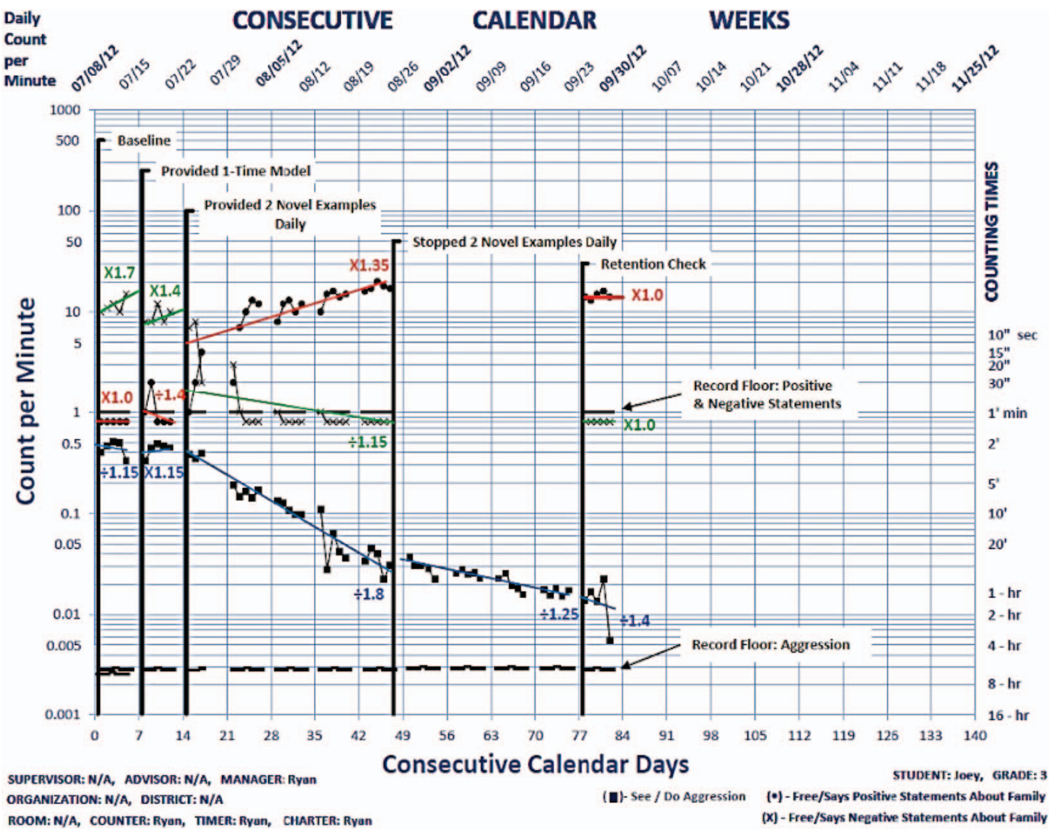

Figure 1. Standard Celeration Chart for a single individual. This chart demonstrates a Celeration chart for a student called Joey, by a counter named Ryan. Aggression and positive and negative statements about the family are charted on a per minute basis across 81 days.

Consider the following example. Note there are five sections to this sample chart: (a) Baseline, (b) Provided 1-Time Model, (c) Provided 2-Novel Examples Daily, (d) Stopped 2-Novel Examples Daily, and (e) Retention Check. In this example, there were two aims of the intervention. The first was to reduce the rate of aggression. The second was to raise the rate of saying positive statements about their family. From the chart, one sees that both aims were achieved. It is important to note that this is just one example of a celeration chart and is not meant to be representative of all celeration charts, which can be used in a variety of manners.

Strengths and Weakness of the Chart

The change in frequency of behavior (usually termed rate in behavioral sciences), which can be seen by the celeration on the chart, is what is of interest. The rate of behavior indicates whether the intervention is working to produce development on a task within a domain. Without tracking frequency (change) rate, participant and staff time will be wasted because the intervention will be less effective. A low rate of change indicates that tasks either too easy or too developmentally advanced are being used. This main purpose of the chart can also be seen as a major strength.

Another strength of the celeration chart is that the y-axis is logarithmic (base 10) so that celeration is easy to see. That is, on a logarithmic scale, progress will be shown roughly as a straight line with a positive slope. The rate of a behavior is displayed on the y-axis and graphed as "Log y" There is a label which is a nominal variable identifying the precise intervention which consisted of what was used. There is also the number x of the intervention,

$$y = \beta x.$$

Because this relationship between frequency of behavior and successive interventions is shown as a straight line, it is intuitively clear whether progress is being made or not. In this case, a logarithmic graph also accommodates frequencies ranging from 1 per day to 1,000 per minute. There are cases of the standard chart that require such a huge range. We will not address the issue of what happens when the frequency is 0. When the chart is not used longitudinally, a large accumulation of individual charts may result, each of which can be looked at separately. However, these individual charts do not provide clear evidence about overall development, the rate of development, or confinement to domains. Evaluating the charts using developmental theories allows us to look at the issue of development without using traditional longitudinal data collected from incomplete cross-sectional tests. Further, using the MHC allows us to look at the rate of development and confine this to domains.

In line with this, behavior can be analyzed by the *difficulty of tasks* that an individual successfully addresses. Task properties that influence item difficulty have two overall parts: (a) OHC of the items in a task, and (b) aspects of task content that are nonstructural—such as language, culture or country, and familiarity, to name a few. The most important predictor of difficulty is classifying each task by its OHC.

What Are the Means to Change Behavior Effectively?

The purpose of the Standard Celeration Chart is to measure the effects of each specific intervention (changes in rates of behavior due to interventions). As stated above, the chart has many advantages in measuring the effectiveness of interventions as the interventions are taking place. In this section of the paper, a very simple model of how interventions produce behavioral change is introduced. One additional assumption is introduced—that the probability of getting an answer right depends on the intervention being correctly placed in the developmental sequence of the individual. If the difficulty of the task is too low in the sequence, no behavior acquisition goes on because the Aim has already been achieved. If the difficulty of the task is too high, no acquisition occurs because the task is too difficult.

The means to change behavior effectively may be represented mathematically is as follows: changes in behavior are simply the product of the time that a participant actively engages in getting correct answers in a task and the probability of getting an answer correct when the task is correctly placed in the developmental sequence, as shown in Equation 1,

$$\Delta B = t_{\text{on task actively engaged in}} \times pl_{\text{getting answer right when placed in sequence correctly}}. \quad (1)$$

Or without the subscript notion

$$\Delta B = t \times pl$$

where ΔB = change in behavior, t = total time on task actively engaged in, and pl = probability of getting answer right when placed in sequence correctly.

Note that $pl = f$ (the task being in the right place in the developmental sequence). That function includes: (a) the OHC of the task, (b) its relationship to the overall sequence of tasks, and (c) perhaps the organism's previous experience in that sequence and subdomain.

Further, where S is the contingency for reinforcement for correct answers:

$$t = f(S)$$

Or, in other words, time engaging actively on a task is dependent on the contingent reinforcement of correct responses. Note that active engagement depends not only on the consequences of making correct responses but the value of those consequences to the participant. Together, they improve the chances of the person being actively engaged in the task,

$$B_{\text{correct}} \rightarrow S^{R+}$$

Every task that can be administered to a participant is a member of a set of task sequences. If Equation 1 correctly explains changing participant behavior, then effective training requires the determination of where in the sequence of tasks lies the most difficult task to which the participant may correctly respond. How is this accomplished?

We answer this by beginning with an example illustrating its accomplishment and then explain the example in general terms. An example is teaching MHC naïve participants through a computer-based instructional method (Commons, Owens, & Will, 2015). In this example, MHC was taught using six different modules. The OHC of the task remained the same within a module, and each subsequent module increased in hierarchical complexity. During the first section, participants were required to match stage names to stage numbers for each of the orders being learned about (5 through 13). This task was at Primary Order 8. The second section's task will not be described in full here. However, it is important to note that the second section's task was at Concrete Order 9, and so on, increasing to Metasystematic Order 13 for the fifth section's task. The workshop was programmed to supply information that is lower than the highest stage at which a person may act. This information is necessary to do the higher order tasks. However, by tracking the participants' progress along a series of tasks, it is possible to see what order tasks they complete and with which they struggle.

In more general terms, the determination of the best particular task to provide an individual simultaneously requires: (a) determining the task sequence to which the particular task belongs (as shown in Tables 2 and 3; e.g., task sequences in the development of language and arithmetic, respectively), (b) finding the OHC of the tasks in which the participant meets the aim, and (c) finding the OHC for the current task. Each of these requirements will be explained next.

Choosing the Correct OHC of the Task and Subtask

Identifying the difficulty of the tasks and subtasks within an order maximizes the student's performance and learning since it allows tasks to be presented in the most effective order. Many actual charts consist of a record of multiple behaviors. The example celeration chart in Figure 1 had two behaviors: (a) aggression, and (b) positive and negative statements about the family. Sometimes within a chart, these different behaviors are ordered by increasing difficulty. This is more likely to be done for a series of charts of behavior change in the same domain.

It is important to explain at this point more detailed components of MHC that will make aggregating charts more understandable. First, because most tasks, except the simplest ones, are combinations of other, lower order, tasks, we define two additional types of tasks: subtasks and subsubtasks. A *subtask action* is defined in terms of (a) only one same order action, and (b) another lower order action. It is not a higher order task action because there is only one next lower order task action that the task action operates on, not two. One *subsubtask* action is

Table 2
Task Sequence of Arithmetic Development

OHC	Name	Example
0	Calculatory	Simple machine arithmetic on 0's and 1's
1	Sensory or motor	Move limbs, lips, toes, eyes, elbows, head; view objects or move
2	Sensory and motor	Either seeing circles, squares, etc. or instead, touching them. O #
3	Circular Sensory-motor	Reaching and grasping a circle or square. O #
4	Sensory-motor	A class of filled in squares may be formed # # # # #
5	Nominal	That class may be named, "Squares"
6	Sentential	The numbers, 1, 2, 3, 4, 5 may be said in order
7	Preoperational	The objects in Row 5 may be counted. The last count called 5, five, cinco, etc * * * * * # # # # # O O O OO
8	Primary	There are behaviors that act on such classes that we call simple arithmetic operations $1 + 3 = 4$ $5 + 15 = 20$ $5(4) = 20$ $5(3) = 15$ $5(1) = 5$
9	Concrete	There are behaviors that order the simple arithmetic behaviors when multiplying a sum by a number. Such distributive behaviors require the simple arithmetic behavior as a prerequisite, not just a precursor $5(1 + 3) = 5(1) + 5(3) = 5 + 15 = 20$
10	Abstract	All the forms of five in the five rows in the example are equivalent in value, $x = 5$. Forming class based on abstract feature
11	Formal	The general left-hand distributive relation is $x * (y + z) = (x * y) + (x * z)$
12	Systematic	The right-hand distribution law is not true for numbers but is true for proportions and sets: $x + (y * z) = (x * y) + (x * z)$ $x \subsetneq (y \subsetneq z) = (x \subsetneq y) \subsetneq (x \subsetneq z)$
13	Metasystematic	The system of propositional logic and elementary set theory are isomorphic: x & $(y$ or $z) = (x$ & y) or $(x$ & $z)$ Logic $\subsetneq x \subsetneq (y \subsetneq z) = (x \subsetneq y) \subsetneq (x \subsetneq z)$ Sets T(False) $\subsetneq \varphi$ Empty set T(True) $\subsetneq \Omega$ Universal set

Note. OHC = order of hierarchical complexity.

defined in terms of one or more actions two orders down. (For a more detailed discussion of subtasks, see Boom, 2012; Commons, 2014).

There is just one principal that describes the organizing of subtask actions. That is, one subtask action serves as a *prerequisite* for the next subtask action. For example, a primary order task is to coordinate preoperational task actions. A Preoperational Order 7 task action coordinates Sentential Order 6 task actions. Coordinating the Preoperational 7 task with a Sentential Order 6 task within a primary task action would be a subtask. In contrast, the subsubtask actions have a weaker relationship and may be arbitrarily organized, or one may serve as *precursor* and may be only *sufficient but not necessary* for the next subsubtask action. The sequence of the subsubtask actions acquired often depends on the sequence that the teacher provides for the student. And that subsubtask may require more horizontal complexity than the preceding task, as in the case of adding more

Table 3
Task Sequence of Speech Development

Stage	Age (month)	Action
2	Birth	Eye contact and listening
3	5	First sounds are uttered
4	12	Sounds progress into words
5	18	Names given to concepts
6	24	Named concepts progress into sentences
7	40	Sentences progress into paragraphs

than two numbers together. Next is a more detailed example.

Arithmetic at the Primary Order

As an example of OHC and how subtasks make up an order, we will discuss arithmetic performance at the primary order (see Table 3 for a more complete task sequence of arithmetic development). At the Primary Order 8, which is the order at which *true counting* appears, two or more actions from the Preoperational Order 7 are coordinated. The first Primary Order 8 subtask actions may organize counts of organized objects, a Preoperational Order 7 task, and apply them to very large numbers of randomly organized sets of objects. This is done by not only using the counting of objects from the Preoperational Order 7, but keeping track of what has been counted, another task from the Preoperational Order 7. By definition, a next order task has to organize to adjacent lower tasks. Therefore organizing just one higher order task and one lower order task is just a subtask. A subtask is transitional to the next stage. Within the "counting" subtask action, there are a number of subsubtask actions. The first subsubtask action could be to count disordered objects that are the same. For example, objects that are all circles. The second would then be to count disordered objects that are not the same. The total number counted indicates the size of the set. For example, for five objects, the size of the set would be "5." The third subsubtask in true counting is counting very large numbers with randomly organized sets of objects. Hence, children performing at the Concrete Stage 9 can count hundreds of objects as opposed to between 10 and 12 or so found in Chimpanzees. Until the concrete stage, it is not over practiced enough to get the accuracy up. In the Primary Stage 8, they learn the subtask actions of addition, subtraction, and then multiplication (Van der Ven, Boom, Kroesbergen, & Leseman, 2011) and their inverses. This can connect ordinality—the ordering of numbers, to cardinality—the labeling of sets of items as a number.

How Would Notions From MHC Be Used as Part of Examining Group Data

Behavior analysts are highly trained in task analysis and going from simple to more difficult tasks and looking at the celeration to see if progress is being made. However, it is important to understand this process not just from a behavioral analytic perspective but also from a behavioral developmental perspective. It has to be emphasized that these task sequences must be looked at from this perspective in both performing and evaluating the intervention.

There are a few steps for getting the OHC of a task to match closely enough to the present performance of the person being trained. First, the long-term aim is selected. The intermediate aims are found by looking at the entire developmental sequence in a subdomain which the long-term aim is a part of. In order to know what the normal developmental sequence looks like, charts need to be combined so that the normal pattern of development can be found. This is the reason for combining charts, so that an individual chart can be compared to the developmental sequence. Then some sample tasks at different orders of hierarchical complexity are given as pretests until one determines the highest OHC task that is successfully completed in each developmental sequence. The results of this behavioral development assessment yield a developmental profile that may be compared to the "normal" developmental profile. Training is then conducted starting at that order to raise the subtask and subsubtask difficulty that the participant successfully addresses. If the rate is close to the aim for that behavior, then a task from the next order is trained.

To Use the Chart in Group Data, Just Use the Stage of the Action

In evaluative studies of intervention programs, there is the necessity of analyzing not just individual data but overall performance of all the individuals in a group. Here, group analysis depends on a longitudinal analysis of individual participant performance improvements made through using a set of interventions. It is best to aggregate data from a single subdomain. This way, the subdomain induced variability is eliminated.

Choosing the correct subdomain is an important factor in a successful intervention. In order to choose the most productive subdomain, one must create a profile of the participant behavioral stage in each subdomain. This is done by assessing stage of performance in all the different subdomains. There is an ordering of difficulty across subdomains. In other words, one has to ascertain that proficiency exists in one subdomain in order to complete tasks in a different subdomain. In the example of language development, development in the subdomain of reading facial expressions and gestures is important before the development of language can progress (see Table 2). This can be a problem for children with autism. It is important to start at the subdomain that a participant performs lowest at and is a prerequisite for other tasks.

For example, one might examine the reading of gestures and emotions as would exist in the first year of life. The best way to see changes in these behaviors is to use multiple regression analysis. If the change in behavior is predicted by session number and the placement of the person's performance in the right place in the sequence as suggested by Equation 1 (presented earlier), then there would be an understandable prediction of acquisition. Change in stage of performance of behaviors from a sequence of tasks in a given subdomain is regressed on time or trial number of the intervention. The independent variable is the trial number within a chart or number in a sequence of charts. If one wants to compare the progress to that of a waiting list control group, time of pretest and posttest would be equated. If there is a control, the performance for the control group would be an assessment on the same sequences. A given stage is met when a task of a given OHC is correctly performed. This usually means meeting the aim.

There would be many example points. Each chart consists of many possible intervention periods. Each has many possible values of one of the variables that is being used as a predictor, or values or possible values, of one of the variables being predicted. There are usually many charts of behavior in a given developmental sequence.

Logistic Regression to See What Progress There is in a Chart

In this section, the applications of a single sequence of tasks in generating a detailed way of combining charts will be presented. Figure 2 represents a computer-generated Standard Celeration Chart from a participant of a workshop

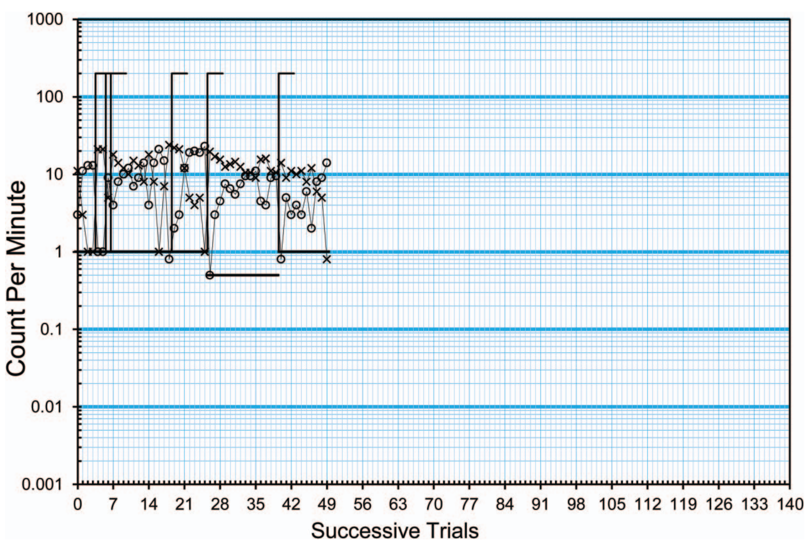

Figure 2. Standard Celeration Chart for a single individual. This chart demonstrates a celeration chart for an SRAD workshop participant. The individual performance reached 100% correct (○) and 0% incorrect (×) upon completing Sections 1, 2, and 3. Vertical lines indicate when participants change modules. The sequence of modules completed for the participant is indicated in the captions under each chart. Sections 4 and 5 show an increase in frequency of correct responding, but also reflect below chance level responding. Order followed: 1, 2, 1, 2, 3, 4, 5. Note that the first attempt at Module 2 and the second attempt at Module 1 are not used in the analysis.

of a Society for Research in Adult Development (SRAD). The content of the workshop sections can be found in Table 4. The different elements of a Standard Celeration Chart, time period and precision teaching, can summarize whether or not correct stage questions were asked. In Figure 2, the circles depict correct answers and the "×" depicts incorrect answers. Vertical lines indicate when participants change sections. Each section contained one module. Note that the last module does not have a vertical line after it. These modules are numbered 1–5, with Modules 1 and 2 being attempted twice. The first attempt at Module 2 and the second attempt at Module 1 were not used in the analysis. The independent variable is whether or not there is a change in stage. Stage change is a two-level variable, $S = \{1, 0\}$. That is, $S = 1$ when there was stage change and $S = 0$ when there was no stage change. In Figure 2, there are vertical lines. What module is being presented is indicated by being to the left of each vertical line, reading from left to right. Within each module the points are numbered from right to left, with the last attempt, as indicated by being closest to the left of the vertical line, being labeled 1.

Table 4
Section, Exercise, Content, and Instruction of MHC Workshop

Section	Exercise	Material and content presented	Instruction
1	Identify corresponding numbers of the MHC order presented	Materials and content presented MHC order name and corresponding number	Each presentation lasts one minute, so work as fast as you can. Do not read everything if you do not need to.
2	Enter order of MHC when presented with definitions	Materials and content included descriptions and examples of orders of the MHC	Each presentation lasts one minute, so work as fast as you can. Do not read everything if you do not need to.
3	Score order of presented task and provide rationale	Materials included axiom rules; example tasks that explained which axioms are violated. Content was experimenter-generated tasks.	Each presentation lasts one minute, so work as fast as you can. Do not read everything if you do not need to.
4	Score order of presented task and provide rationale	Example tasks were drawn from Counselor-Patient vignette	In the following short groups of sentences, please determine the stage. After you correctly identify the stage, you will be directed to two questions. The first question will ask you to identify the reasons why that sentence or statement was at that particular stage. The second question will ask you why that sentence or statement was not at a higher stage. Each presentation lasts two minutes minute, so work as fast as you can. Do not read everything if you do not need to.
5	Score order of presented task and provide rationale	Example tasks were taken from Counselor-Patient interaction vignette	Read all five vignettes in the Counselor-Patient Interaction carefully. Score the stage of each vignette. Identify components of each vignette: Variables, Relations, Systems, Relations among systems. Each presentation lasts one minute, so work as fast as you can. Do not read everything if you do not need to.
6	Score novel tasks and provide rationale	Example tasks included experimenter-generated tasks that included a variety of domain in tasks	Each presentation lasts two minutes minute, so work as fast as you can. Do not read everything if you do not need to.

Note. Content covers the MHC orders from Nominal (5) through Metasystematic (13). Content was presented prior to exercises.

The number of trials per module varied from a low of 4 for Module 1 to a high of 14 for Module 4.

This analysis will simply attempt to support the hypotheses in this paper about the means to change behavior. As Equation 1 posits earlier in this paper, change in behavior should be a result of the difficulty of the task as demonstrated by the behavioral developmental order of difficulty, as well as the time actively engaged in the task. For the data in Figure 2, the participant was put in the correct developmental sequence, since each module increased in OHC by one. Therefore, their behavior should change, represented by reaching the aim, based on these two parameters. The task difficulty was represented by the number of the module and the time actively engaged was represented by the number of attempts in the sequence.

The first hypothesis is that the more difficult the module, the lower the probability of having a stage change in any particular trial in that module. The second is that the more trials one undertakes in a module, the higher is the probability of achieving a stage change on any particular trial. The specification $y^* = (b0 + b1\ M + b2\ s)$, where M is the difficulty of the module and s is the sequence number of the trial, is used to test whether the data is or is not consistent with these hypotheses.

The probability that a trial is a success is the cumulative logit error from minus infinity to y^*, as in Equation 2,

$$Prob[y* > 0] = 1/(1 + \mathrm{Exp}[-(b0 + b1\ M + b2\ s)])$$

and (2)

$$Prob[y* \leq 0] = 1/(1 + \mathrm{Exp}[(b0 + b1\ M + b2\ s)])$$

The Hessian of the problem is calculated at the estimated coefficient values (Rose & Smith, 2013). The square root of the diagonal elements of the Hessian estimates the standard errors of the coefficient estimates.

The results of this analysis are shown in Table 5. Because each successive module is more difficult, the probability of stage advancement ought to decline as the module number rises. Hence the estimate of β_1 should be negative, and statistically significant. In fact, it estimates as negative, and statistically different from zero ($\beta_1 = -0.712368$, $t = -2.42817$).

If the individual learns by doing, then the probability of stage advancement should go up with more attempted trials. Hence the estimate of β_2 should be positive. In fact, it estimates as positive and statistically different from zero ($\beta_2 = 0.292425$, $t = 2.57761$).

Therefore, the empirical analysis showed that the hypotheses were consistent with the data.

Discussion

This paper attempts to lay out a model of change in behavior through the use of a behavioral developmental model, the MHC. This model should prove to be effective in both implementing and evaluating behavioral interventions. The paper posits that change in behavior is evidenced by change in behavioral developmental stage. This change in behavioral stage will be dependent on the (a) difficulty of the task, as evidenced by the OHC of the task; and (b) the time spent actively engaged. However, this is only true if one is put in the correct place along a behavioral developmental task sequence. Details about how to identify subtasks and subsubtasks in given sequences in order to properly implement sequences and specific examples of arithmetic and language learning have been discussed. Using these examples, it should be possible to create other behavioral developmental task sequences that will allow for effective interventions. The goal is to eventually specify the content, sequence, and rate of development for the paths of development within a subdomain.

The paper offers a statistical analysis of a chart from a single participant in the MHC workshop. In this example, the tasks had been designed to have been placed in a correct sequence and all correct answers were rewarded through a point system and scoring leaderboard. The analysis supports the hypotheses that success, in terms of

Table 5
Results of Poisson Model Evaluation

	Coefficient estimate	Standard error of coefficient	t-value estimate
β_0	.635955	.975248	.652096
β_1	−.712368	.293376	−2.42817
β_2	.292425	.113448	2.57761

reaching the aim, was dependent on difficulty of task (module number) and time spent on the task (number of attempts). The assumption in this analysis is that the tasks are already in the correct behavioral developmental sequence, and that this sequence allows for the best learning environment. The workshop was designed by practitioners who were well trained and practiced in MHC. As discussed in this paper, the MHC is a well-studied model and it is central to the theory in this paper that placement in a behavioral development sequence is necessary for proper learning. Additionally, the nature of the learning in this particular example is ambiguous. Stage advancement is measured by going from failure on a task to success on a task. There are many theories about learning and it is an overly determined phenomenon. The empirical tests performed do not distinguish the nature of learning in this case.

One unaddressed question in this paper is how to prioritize what sequences on which to work. This is a separate paper. What is addressed here is how effective a kind of intervention is within a subdomain and a task sequence. There are additional variables that can be added to this calculation. With multiple participants and charts, the task of analyzing the data only becomes slightly more difficult. However, the sample analysis in this paper demonstrated that, if properly implemented, using the model proposed in this paper, an intervention was effectively be assessed and shown to work well. It is just a start for combining charts and properly assessing interventions.

References

Bernholt, S., Parchmann, I., & Commons, M. L. (2009). Kompetenzmodellierung zwischen Forschung und Unterrichtspraxis. *Zeitschrift für Didaktik der Naturwissenschaften, 15,* 219–245.

Binder, C. (2000, May). *Component/composite analysis and programming from the real world*. Paper presented at the annual meeting of the Association for Behavior Analysis, Washington, DC.

Boom, J. (2012). *New model for strategy development: combining categorical data analysis with growth modeling*. Paper presented at Biennial Meeting of the International Society for the Study of Behavioral Development, Edmonton, Canada.

Commons, M. L. (2014). What are the relationships between four notions of stage change? *Behavioral Development Bulletin, 19,* 15–21. http://dx.doi.org/10.1037/h0100584

Commons, M. L., Giri, S., & Harrigan, W. J. (2014). The small effects of non-hiearchical complexity variables on performance. *Behavioral Development Bulletin, 19,* 31–36. http://dx.doi.org/10.1037/h0101079

Commons, M. L., Goodheart, E. A., Pekker, A., Dawson, T. L., Draney, K., & Adams, K. M. (2008). Using Rasch scaled stage scores to validate orders of hierarchical complexity of balance beam task sequences. *Journal of Applied Measurement, 9,* 182–199.

Commons, M. L., Goodheart, E. A., Pekker, A., Dawson-Tunik, T. L., Cyr, E., & Rodriguez, J. A. (2005). The relationship between orders of hierarchical complexity and Rasch scaled stage scores: Balance beam, laundry, and counselor-patient task sequences. *Journal of Applied Measurement, 10,* 1–11.

Commons, M. L., Owens, C. J., & Will, S. M. (2015). Using a computer-based precision teaching program to facilitate learning of complex material: The case of the model of hierarchical complexity. *Behavioral Development Bulletin, 20,* 207–226. http://dx.doi.org/10.1037/h0101379

Dawson, T. L. (2002). A comparison of three developmental stage scoring systems. *Journal of Applied Measurement, 3,* 146–189.

Rose, C., & Smith, M. D. (2013). *Mathematical statistics with Mathematica*. New York, NY: Springer.

Van der Ven, S. H. G., Boom, J., Kroesbergen, E. H., & Leseman, P. M. (2011). Relations between the development of strategy selection in learning single-digit multiplication and working memory: A microgenetic study. *Journal of Experimental Child Psychology*. Manuscript submitted for publication.

(*Appendix follows*)

Appendix

Proposed Equations for Combining Charts

There are two forms of difficulty:

1. The first is the task number for a given OHC.
2. The second is the subtask number.

T_j represents the number of the task *j*, presented to a student;

T_j has different classes of subsubtasks and S_{jk} is the number of activities of subsubtasks type *jk*, presented to the student;

RS_{jk}/S_{jk} is the percentage of times activity S_{jk} was appropriately reinforced;

PS_{jk}/S_{jk} is the percentage of times activity S_{jk} was the appropriate activity presented to the student, given their performance.

$\rho_{11}\ RS_{11}/S_{11}$ = the contribution of appropriately reinforced (and possibly not appropriately placed) trials to T_1;

$\pi_{11}PS_{11}$ = the contribution of appropriately placed (and possibly not appropriate reinforced) trials to T_1;

$\zeta_{11} = RS_{11}(PS_{11}/S_{11})S_{11}$ = the contribution of the proportion of appropriately placed and reinforced trials to T1;

β_{jk} is the contribution of a subtask *k* to task *j* to the change in OHC, whether or not it is appropriately given and appropriately reinforced. A possible hypothesis is that all β_{jk} are equal to zero.

γ_{jk} is the contribution to the change in stage when only the appropriately provided subtask *k* of task *j* is made;

δ_{jk} is the contribution to the change in stage when only the appropriately reinforced subtask *k* of task *j* is made;

η_{jk} is the contribution to the change in stage when both the provision and the reinforcement of subtask *k* of task *j* is made; and

ε is a random error whose expected value is equal to zero.

The following equations describe how to combine charts across tasks and across people. Equation A1 describes the change in performance on a task of a given order of higher complexity as a function of the tasks engaged in with a student's interest.

$$\text{The propensity to } \Delta\text{Stage} = \tau_1 T_1 + \tau_2 T_2 + \ldots + \tau_n T_n + \varepsilon \qquad \text{(A1)}$$

where the coefficients τ_j are unknown coefficients. τ_j is the contribution of one task of type T_j to the change in performance stage.

Equation 1 describes the counting of appropriately given and appropriately reinforced tasks provided,

$$T_1 = (\gamma_{11} + \rho_{11}R_{11} + \pi_{11}P_{11} + \zeta_{11}R_{11}P_{11})S_{11} + (\gamma_{12} + \rho_{12}R_{12} + \pi_{12}P_{12} + \zeta_{12}R_{12}P_{12})S_{12} + \ldots,$$

$$T_2 = (\gamma_{21} + \rho_{21}R_{21} + \pi_{21}P_{21} + \zeta_{21}R_{21}P_{21})S_{21} + (\gamma_{22} + \rho_{22}R_{22} + \pi_{22}P_{22} + \zeta_{22}R_{22}P_{22})S_{22} + \ldots. \qquad \text{(A2)}$$

(*Appendix continues*)

Substituting Equation A2 into Equation A1 yields Equation A3,

$$\begin{aligned}\Delta\text{Stage} =&\\ \tau 1\{&(\gamma_{11} + \rho_{11}R_{11} + \pi_{11}P_{11} + \zeta_{11}R_{11}P_{11})S_{11} + (\gamma_{12} + \rho_{12}R_{12} + \pi_{12}P_{12} + \zeta_{12}R_{12}P_{12})S_{12} + \ldots\} +\\ \tau 2\{&(\gamma_{21} + \rho_{21}R_{21} + \pi_{21}P_{21} + \zeta_{21}R_{21}P_{21})S_{21} + (\gamma_{22} + \rho_{22}R_{22} + \pi_{22}P_{22} + \zeta_{22}R_{22}P_{22})S_{22} + \ldots\} + \ldots\\ \tau_n\{&(\gamma_{n1} + \rho_{n1}R_{n1} + \pi_{n1}P_{n1} + \zeta_{n1}R_{n1}P_{n1})S_{n1} + (\gamma_{12} + \rho_{12}R_{12} + \pi_{12}P_{12} + \zeta 1_{n2}R_{n2}P_{n2})S_{n2} + \ldots\} + \varepsilon\end{aligned} \tag{A3}$$

Simplifying notation yields Equation A4:

$$\begin{aligned}\Delta\text{Stage} =&\\ &\beta_{11}S_{11} + a_{11}R_{11}S_{11} + \delta_{11}P_{11}S_{11} + \eta_{11}R_{11}P_{11}S_{11} + \ldots\\ &\quad \beta_{12}S_{12} + a_{12}R_{12}S_{12} + \delta_{12}P_{12}S_{12} + \eta_{12}R_{12}P_{12}S_{12} + \ldots\\ &\beta_{21}S_{12} + a_{21}R_{21}S_{21} + \delta_{21}P_{21}S_{21} + \eta_{21}R_{21}P_{21}S_{21} + \ldots\\ &\quad \beta_{22}S_{22} + a_{22}R_{22}S_{22} + \delta_{22}P_{22}S_{22} + \eta_{22}R_{22}P_{22}S_{22} + \ldots + \ldots +\\ &\beta_{n1}S_{n1} + a_{n1}R_{n1}S_{n1} + \delta_{n1}P_{n1}S_{n1} + \eta_{n1}R_{n1}P_{n1}S_{n1} +\\ &\quad \beta_{n2}S_{n2} + a_{n2}R_{12}S_{n2} + \delta_{n2}P_{n2}S_{12} + \eta_{n2}R_{n2}P_{n2}S_{n2} + \ldots + \varepsilon\end{aligned} \tag{A4}$$

where:

$$\beta_{jk} = \tau_j\gamma_{jk};$$
$$a_{jk} = \tau_j\rho_{jk};$$
$$\delta_{jk} = \tau_j\pi_{jk};$$
$$\eta_{jk} = \tau_j\zeta_{jk}.$$

If ε is assumed to be distributed as a *N*[0,1], the coefficients can be estimated by probit analysis. If ε is assumed to have a logistic[0,1] distribution, the coefficients can be estimated by logistic regression analysis.

Received August 28, 2015
Revision received November 12, 2015
Accepted November 12, 2015 ■