

The relative accuracy and interpretability of five sequential analysis methods: A simulation study

Blair P. Lloyd¹ · Paul J. Yoder^{1,2} · Jon Tapp² · Johanna L. Staubitz¹

Published online: 20 October 2015
© Psychonomic Society, Inc. 2015

Abstract A variety of sequential analysis methods exist to quantify close temporal associations between events from direct observation data. In the present study, we compared the relative accuracy and interpretability of five sequential-analysis methods using simulated data. The methods included three existing approaches (*event lag*, *concurrent interval*, and *time window*) and two proposed modifications of the event lag approach (*event lag with contiguous pauses* and *event lag with noncontiguous pauses*) designed to address limitations of the existing approaches. We evaluated accuracy on the basis of the extent to which the mean contingency estimates produced by each method approximated a known mean (i.e., zero). We evaluated interpretability on the basis of the extent to which the contingency estimates produced by each method were independent from chance estimates of the two-event sequence. The results indicated that the *event lag with contiguous pauses* method produced the most accurate and interpretable estimates of contingency. This modified method prevents the total number of event types from influencing contingency estimates, thus solving a problem associated with the traditional event lag method.

Keywords Sequential analysis · Contingency · Event lag · Concurrent interval · Time window · Simulation

✉ Blair P. Lloyd
blair.lloyd@vanderbilt.edu

¹ Department of Special Education, Peabody College, Vanderbilt University, Box 228, Nashville, TN 37203, USA

² Vanderbilt Kennedy Center, Vanderbilt University, Nashville, TN, USA

Sequential analysis combines measurement, statistical knowledge, and theory to determine whether the occurrence of one event is associated with the subsequent occurrence of another event (Bakeman & Quera, 2011; Yoder & Symons, 2010). Sequential analyses provide quantitative indices of contingency between events as they occur in real time. Though sequential analyses produce correlational data, the identification of close temporal relations between two events may be used to inform subsequent studies in which the immediate influences of one event on another are evaluated experimentally. For example, these data may help investigators eliminate or rule out events or behaviors that lack close temporal relations, thereby reducing the expense of conducting numerous experiments on each potential causal event. Sequential associations themselves also may be used as primary dependent variables in experimental studies (e.g., Symons et al., 2001; Yoder, Davies, Bishop, & Munson, 1994).

Sequential analysis methods have been used to address research questions emerging from multiple disciplines and methodologies, many of which relate to patterns of interaction between actors. In developmental psychology, researchers have applied sequential analysis to study parent–child interactions as they relate to language development in young children with developmental disabilities and communication delays (e.g., Vernon, 2014; Yoder, Davies, & Bishop, 1994). Behavior-analytic researchers have used sequential analysis to identify potential reinforcing consequences of problem behavior for individuals with intellectual and developmental disabilities (e.g., Martens, Gertz, Werder, & Rymanowski, 2010; Symons, Hoch, Dahl, & McComas, 2003; Vollmer, Borrero, Wright, Van Camp, & Lalli, 2001). Though it is less common, researchers also have used sequential analysis to estimate the temporal relations between events *within* an actor (e.g., successive episodes of self-injury [Marion, Touchette, & Sandman, 2003] or epileptic seizures and problem behavior

in adults with developmental disabilities [Roberts, Yoder, & Kennedy, 2005]). In relation to research methodologies, sequential analyses have been applied to research questions focusing on both between-group (e.g., Pineda, Cole, & Bruce, 2007; Wilde, Silva, & Oliver, 2013) and within-participant (e.g., Symons et al., 2003; Vollmer et al., 2001) comparisons.

Existing methods of sequential analysis

Because contingencies between events cannot be directly observed, a variety of methods have been used to estimate them. Different methods of estimating contingencies are defined on the basis of their behavior sampling and data representation approaches. The *behavior sampling* approach refers to how the direct observation data are collected. For example, an entire observation session may be coded by indicating the times at which key events occur in the session (i.e., continuous, timed-event sampling). Alternatively, an observation session may be divided into fixed intervals (e.g., 10 s) and the presence or absence of each key event may be recorded for each interval (i.e., interval sampling). *Data representation* refers to whether the data are represented for analysis as events in sequence or as coded time units. When the research question requires testing the sequential association between two events in a sequence or within some unit of time, a 2×2 contingency table is used to organize the sequential data. The tallied units may be event pairs or time units (i.e., fixed intervals or seconds) and are classified into one of four cells, which are defined by the sequence of key events or the presence or absence of key events within specified units of time.

For research questions specifying an analysis of event onsets, three existing methods of sequential analysis are the *event lag*, *concurrent interval*, and *time window* approaches. Each method represents a different combination of behavior sampling and data representation approaches.

The *event lag* approach involves continuous event sampling with pairs of events as the coded unit. In Fig. 1 and throughout this article, we refer to Event X followed by Event Y as the sequence of interest. A 2×2 contingency table representing the event lag approach is presented in the top left panel of Fig. 1. Units with Event X followed by Event Y are tallied in Cell A; units in which only one event is in the observed sequence of interest are tallied in Cells B and C; and units in which neither event is in the observed sequence of interest are tallied in Cell D. This approach to tallying event pairs in the 2×2 contingency table has a long history in addressing research questions in which the sequence of events is critical (Bakeman & Quera, 2011). The “not X” and “not Y” labels in Fig. 1 may be defined as any other non-key events in a coding system. In event lag sequential analysis, the total number of event types that a researcher codes impacts the B, C, and D cell counts, which impacts the contingency estimate.

This issue has been described as the “D cell problem” (Yoder & Symons, 2010). In the present article, we refer to this issue as the *total event types* problem, because the issue does not only affect the D cell.

The total number and cell placement of *time units* are not influenced by the total number and types of non-X and non-Y events being coded. Thus, interval-based sequential analysis has been proposed as a potential solution to the total event types problem (Yoder & Symons, 2010). In addition, using intervals as the time unit represents a resource-efficient sampling method. The cost-saving nature of interval behavior sampling may be especially relevant when observational data are being collected in natural settings (e.g., classrooms) or when video-recording is not possible. The *concurrent interval* approach involves interval sampling and tallies intervals according to the presence or absence of key events within the same interval. As shown in the top right panel of Fig. 1, units with both key events present are tallied in Cell A; units in which one key event type occurs without the other are tallied in Cells B and C; and units in which both key events are absent are tallied in Cell D. That is, the concurrent interval approach tallies intervals in the 2×2 contingency table according to the *co-occurrence* of events within an interval.¹ Any interval in which both key events occur is tallied in the A cell, regardless of the sequence of these events within the interval. The efficiency of interval sampling over continuous behavior sampling, therefore, may be outweighed by the loss of information on the event sequence within intervals (Bakeman & Quera, 2011; Martin & Bateson, 2007). Regardless, this method has been used to address research questions concerning event sequence (e.g., Eckert, Martens, & DiGennaro, 2005).

A third method of sequential analysis is the *time window* approach. This approach requires continuous event sampling with seconds as the coded units of analysis (Bakeman & Quera, 1995; Yoder & Symons, 2010). A 2×2 contingency table representing the time window approach is presented in the bottom right panel of Fig. 1. Time units (seconds) within a specified window after Event X with Event Y present are tallied in Cell A; seconds within the time window after Event X with Event Y absent are tallied in Cell B; seconds outside the specified window after Event X with Event Y present are tallied in Cell C; and seconds outside the specified window after Event X with Event Y absent are tallied in Cell D. Thus, the sum of all four cell counts equals the total number of time units (seconds) in the observation. Like the event lag method, the time window method preserves the sequence of key events, and like the

¹ The *interval lag* sequential analysis, an alternative interval-based method, addresses the extent to which one key event occurs in the interval *directly following* the other (Bakeman & Quera, 2011). Because the interval lag method is most appropriate for intervals that are short enough to contain single events, this method undermines the rationale for using interval sampling as a practical alternative to continuous behavior sampling. Thus, we did not include this method in the simulation.

		Event Lag		Concurrent Interval	
		Event Y is second event	Event Y is not second event	Event Y present	Event Y absent
Event X is first event	A	[X, Y]	B	[X, not Y]	
Event X is not first event	C	[not X, Y]	D	[not X, not Y]	

		Event Lag With Contiguous Pauses		Time Window	
		Event Y is second event	Event Y is not second event	Second with Event Y present	Second with Event Y absent
Event X is first event	A	[X, Y]	B	[In window, Y]	[In window, not Y]
Event X is not first event	C	[Y, Y] [P, Y]	D	[Y, X] [P, X] [Y, P] [P, P]	[Outside window, Y] [Outside window, not Y]

Fig. 1 Organization of tallied units among the four cells of contingency tables for the *event lag* (top left), *concurrent interval* (top right), *time window* (bottom right), and *event lag with contiguous pauses* (bottom left) approaches. X = Event X, Y = Event Y, P = pause

concurrent interval method, the use of time as the coded units also prevents the number and types of non-key events from influencing the cell counts. However, as we discuss in a subsequent section (Methods for Estimating Simple Probabilities of Key Events), there is an asymmetry in the conceptualization of the simple probabilities for Event X and Event Y (i.e., 1st row and 1st column marginals). This asymmetry has the potential to produce a biased contingency estimate (i.e., a mean contingency different from the known mean). Though an extant simulation supported the accuracy of the time window approach using a 2-s window (Yoder & Tapp, 2004), the accuracy of contingencies produced using the time window method with longer temporal windows is unknown.

Characteristics of a useful contingency estimate

Contingency estimates can be calculated from the four cell values of the 2×2 contingency tables produced from each sequential analysis approach described above. The utility of contingency estimates from the different sequential analysis approaches may be judged according to their relative accuracy and interpretability. Contingency estimates are considered accurate if they are trivially different from the *known* or *true* contingency value. Outside of simulation studies modeling a zero mean contingency, we are not aware of any scenarios in which the *true* contingency value may be known. In addition, the accuracy of contingency estimates should not change depending on the characteristics of the observational data collected. For example, contingency estimates that are accurate across conditions in which the total frequency of coded events varies would be more desirable than contingency estimates

that are accurate only under conditions in which the total frequency of events is at a particular level.

Contingency estimates are considered interpretable if they are influenced by the probability of the target sequence of the key events, yet independent from the simple probabilities of each event, and thus independent from the chance occurrence of the sequence of interest (Bakeman & Quera, 2011; Yoder & Symons, 2010). The simple probability is the count of the event or time unit type divided by the total number of events or time units. The chance estimate of the probability that two events will occur together is calculated as the product of the simple probabilities of the two events. That is, we would expect the likelihood of two events occurring in sequence to increase by chance alone as the simple probabilities of the events increase. Because the goal of sequential analysis is to identify close temporal links that are *not* due to chance, contingency estimates that are independent from the chance sequential occurrence or co-occurrence of the key events are considered to be more interpretable than those that are not.

Another aspect of the relative interpretability of contingency estimates is the match between the approach used to quantify the contingency and the research question. Preserving the sequence of key events is necessary to address most research questions involving hypothesized immediate effects of one event on another. That is, we often need a contingency estimate that increases as the number of coded units with both events *in sequence* increases relative to that expected by chance. In addition to preserving the sequence of key events, it is also important to account for the temporal proximity of key events. If we only coded the two events of interest in sequence without accounting for the time between the key

events, then “sequences” separated by several minutes would be treated the same (i.e., tallied in the A cell) as sequences separated by several seconds. Thus, for most research questions involving sequential analysis, it is critical to preserve both the temporal sequence and proximity of the two events of interest.

Methods for estimating simple probabilities of key events

The methods used to estimate the simple probabilities of each event type depend on the method of data representation used. The event lag method produces base probabilities of each event of interest (i.e., the number of instances of each key event divided by the total number of coded events). The concurrent interval approach produces interval base rates (i.e., the number of intervals with each key event divided by the total number of intervals). As we mentioned earlier, the time window approach estimates simple probabilities differently for the X event than for the Y event. Base rates of Event X (the first event) are estimated by dividing the number of seconds *within a specified window after Event X* by the total number of seconds. Base rates of Event Y (the second event) are estimated by dividing the number of seconds with Event Y present by the total number of seconds. Because the time window is invariably longer than one time unit, the estimated base rate of Event X would be larger than the estimated base rate of Event Y even with equal occurrences of the two events.

Because the method of estimating the simple probabilities of each event will influence the estimate of the two events occurring in sequence due to chance alone, the method of conceptualizing simple probabilities may play an important role in the interpretation of contingency estimates. When various sequential analysis methods produce discrepant contingencies, those that are independent from their estimates of chance sequential occurrence or co-occurrence are considered more interpretable than those that are not (Bakeman & Quera, 2011; Lloyd, Kennedy, & Yoder, 2013; Yoder & Tapp, 2004). To date, however, no studies have examined the extent to which each method of estimating simple probabilities, and thus the chance sequential occurrence or co-occurrence of key events, is associated with the contingency estimates produced by the different sequential analysis methods.

Modifications to existing methods of sequential analysis

The limitations identified for each of the sequential analysis methods described above present an opportunity to evaluate modifications to existing sequential analysis methods that may improve upon existing approaches. To summarize, the event lag method accounts for event sequence but does not account for temporal proximity; the concurrent interval method accounts for temporal proximity but does not account for event sequence; and the time window approach accounts for

both event sequence and temporal proximity, but estimates the simple probabilities of Events X and Y differently, which may impact the interpretability of contingency estimates. In the present article, we propose two potential modifications of the traditional event lag method to account for both event sequence and temporal proximity. The proposed modifications each involve (a) stripping out the non-X and non-Y events from continuous timed-event data streams and (b) inserting *pauses* as a separate event when neither X nor Y has occurred for a specified period of time (e.g., Yoder, Davies, & Bishop, 1994). Stripping out all non-X and non-Y events would seemingly solve the *total event types* problem, because no other events would be considered when computing the contingency estimate. That is, if this method resulted in accurate contingency estimates that were independent from the chance occurrence of the two events of interest, it would enable event lag sequential analysis regardless of the type and number of non-X and non-Y events coded.

When inserting pauses as a separate event type, a relevant question is whether to insert (a) contiguous, fixed-duration pause events or (b) single, noncontiguous pause events to represent the passage of time at or beyond a specified duration without either key event occurring. For example, if Event X was coded at 1 min 30 s, Event Y was coded at 2 min 5 s, and the minimum pause duration was set at 10 s, one could insert (a) three contiguous 10-s pause events between Event X and Event Y or (b) a single, noncontiguous pause event. To address this question empirically, we included two variations of the modified *event lag with pauses* approach: (1) *event lag with contiguous pauses* (EL-CP) and (2) *event lag with non-contiguous pauses* (EL-NCP). The data representation for the EL-CP method is shown in the bottom left panel of Fig. 1. The only difference between the data representations for the EL-CP and EL-NCP methods is that in the former, it is possible for pause events to follow each other in sequence, and these sequences would be tallied in Cell D, whereas in the latter, it would not be possible for pause events to follow each other. This article constitutes the first proposal of event lags with computer-inserted pauses and the first evaluation of the relative accuracy and interpretability of the three existing sequential analysis methods and the modified event lag approaches presented in Fig. 1.

Advantages of simulation studies

Simulation studies can provide empirical guidance regarding the accuracy and interpretability of the contingency estimates produced by different sequential analysis approaches (e.g., Bakeman & Dorval, 1989; Lloyd et al., 2013; Yoder & Tapp, 2004). First, when large numbers (e.g., 5,000) of timed-event streams are generated using a random process to sequence events and assign events to time units, a known (i.e., zero) mean contingency is produced. Thus, methods

producing the mean contingencies closest to zero may be interpreted as the most accurate. Second, large numbers of randomly generated data streams provide opportunities for analyzing subsets of data representing different levels of a variable of interest (e.g., data streams with varying total frequencies of events). Third, estimating simple probabilities of each event on the basis of each sequential analysis approach allowed us to address whether and to what extent the contingencies produced by each method are associated with simple probabilities and chance associations. Fourth, continuous timed-event streams may be generated, then re-presented as event or interval streams, which enables multiple methods of sequential analysis to be applied to the same data set. Modifications to existing sequential analysis approaches also may be applied to evaluate the effects such modifications have on the accuracy and interpretability of contingency estimates.

Study purpose and research questions

The purpose of the study was to compare the following sequential analysis approaches on the basis of the accuracy and interpretability of their contingency indices: *event lag*, *concurrent interval*, *time window*, *EL-CP*, and *EL-NCP*. We conducted the present simulation study to address the following research questions: (1) Across sequential analysis approaches, which approach produces the most accurate index of contingency? (2) Does the relative accuracy of the contingency estimates produced by each sequential analysis approach depend on the total frequency of key events? and (3) Which method(s) produce indices of contingency that are most interpretable (i.e., least correlated with the chance occurrence of the two-event sequence)?

Method

Data generation

We generated 5,000 timed-event data streams consisting of two key events with no durations (i.e., onsets only). Only two events were generated because we sought to model the effects of stripping out non-X and non-Y events. For each key event, we randomly selected a number from a uniform distribution with a minimum of .01 and a maximum of .60. This number represented the simple probability of each event type. We then multiplied each probability by a randomly selected whole number value between 100 and 1,800. For example, if .10 and .50 were randomly selected for Events X and Y, respectively, and the randomly selected whole number value was 1,000, then 100 X events and 500 Y events were included in the behavior stream (i.e., a total of 600 events). Finally, we randomly assigned each event to one of 3,600 positions within

the event stream to provide a representation analogous to a time of occurrence. This step allows codes to follow themselves (i.e., repeatable events). More on this topic will be provided in the section on contingency table construction. In this way, 5,000 random sequences of Event X and Event Y were generated. These event streams were conceptualized as representing onset-only, timed-event streams from a 1-h observation with a mean contingency of zero. Using randomization to generate and sequence the events produced a known contingency (i.e., zero) on which judgments of accuracy could be based. Unlike possible attempts to model non-zero contingencies with a known mean, randomly generating these data does not favor one method of sequential analysis over another.

Contingency table construction

The five sequential analysis approaches primarily differ in terms of how analysis units are tallied among the four cells of the 2×2 contingency tables (Fig. 1). For the *event lag* approach, we applied the following logic to the generated timed-event streams. Pairs of events in the stream were tallied using “overlapping” pairs, in which the second event in one pair was the first event in the next pair (see Yoder & Symons, 2010, for a detailed description). This method has been shown to double the number of tallied units without inflating contingency estimates (Bakeman & Dorval, 1989). Because the only two events in the stream were the two events of interest (Event X and Event Y), four pairs of events were possible (i.e., [X, Y], [X, X], [Y, Y], and [Y, X]), with each possible pair being tallied in a different cell of the contingency table (Cells A, B, C, and D, respectively). For research questions involving repeatable events (i.e., events that can follow themselves in sequence), allowing events to follow themselves in sequence is necessary for the data to be represented in a mutually exclusive and exhaustive manner. Excluding events that follow themselves from contingency tables has been shown to produce biased contingency estimates (Yoder & Symons, 2010). Thus, the contingency tables that we generated apply only to the case of repeatable events.

For the *concurrent interval* approach, we imposed 10-s intervals over the generated timed-event streams. Thus, each event stream was divided into 360 intervals, and all intervals were tallied among the four cells of the contingency table according to the presence and absence of each key event per interval. As is shown in the top right panel of Fig. 1, intervals containing both key events, regardless of the event sequence, were tallied in Cell A.

To construct contingency tables for the *time window* approach, event streams were divided into 3,600 s, and all seconds were tallied among the four cells of the contingency table (see the bottom right panel of Fig. 1). Seconds within a 10-s window following Event X with Event Y present were tallied in Cell A; seconds within this window but with Event Y absent

were tallied in Cell B; seconds outside a 10-s window following Event X with Event Y present were tallied in Cell C; and seconds outside a 10-s window following Event X with Event Y absent were tallied in Cell D. Note that the conceptualization of the simple probability (i.e., base rate) of Event X, defined by the 1st row marginal, is different from that of Event Y, defined by the 1st column marginal. Rather than representing the number of seconds with Event X present divided by the total seconds, the base rate of Event X was calculated as the number of seconds *within a 10-s window following Event X* divided by the total number of seconds. The base rate of Event Y, however, was calculated as the number of seconds with Event Y present divided by the total number of seconds.

To apply the EL-CP approach, we added pauses as a third event type to be inserted throughout the same timed-event streams for every 10 s in which neither key event occurred. Just as multiple instances of the same event could be coded in sequence, pause events also could be coded after other pause events when extended intervals of time elapsed without either key event. With the addition of this third event type, a total of nine possible pairs of events were tallied in the contingency table, as is shown in the bottom left panel of Fig. 1. To apply the EL-NCP approach, we inserted single pause events when 10 s or more had elapsed without either key event occurring. Because pause events could not be coded after other pause events, eight possible pairs of events were tallied in the EL-NCP contingency table.

Contingency index

From the 2×2 contingency tables produced by each sequential analysis approach, we calculated risk difference (RD; Higgins & Green, 2011) as the index of contingency. RD is the difference between two conditional probabilities: the probability of Event Y given the presence of Event X ($A/A + B$) minus the probability of Event Y given the absence of Event X ($C/C + D$). RD ranges from -1 to 1 , with values between 0 and 1 indicating positive contingencies (i.e., Event Y is *more* likely to occur given Event X than in the absence of Event X) and values between -1 and 0 indicating negative contingencies (i.e., Event Y is *less* likely to occur given Event X than in the absence of Event X); zero values indicate no contingency (i.e., Event Y is equally likely given the presence or absence of Event X). RD has been used extensively in the epidemiological literature (Rothman, Greenland, & Lash, 2008). More recently, this index also has been termed the “operant contingency value” (OCV) in the behavior-analytic literature, since this difference in conditional probabilities is consistent with how operant contingencies of reinforcement are defined (Martens, Gertz, Werder, Rymanowski, & Shankar, 2013). RD is derived from a formula similar to that of another popular index of sequential association, Yule’s Q

(Yule & Kendall, 1957). RD and Yule’s Q are thus strongly and positively correlated (Lloyd et al., 2013). We selected RD as the index of contingency for the simulation study because Yule’s Q has an undesirable tendency to reach its maximum or minimum value when one or more contingency table tallies are zero, whereas RD does not (Lloyd et al., 2013; Martens et al., 2013).

Analysis methods

To evaluate the accuracy of the contingency indices produced from each approach, we calculated estimates of the effect size (Cohen’s d), in which zero was the reference value. The sequential analysis approach producing the smallest effect size (i.e., closest to zero) was interpreted as being the most accurate. To evaluate whether the accuracy of the contingency estimates produced by each method varied with the total frequency of events, we created three subgroups of event streams using the frequency values corresponding to the 25th and 75th percentiles, and computed the one-sample effect sizes for each subgroup and each sequential analysis method.

To evaluate the interpretability of the contingency indices produced by each sequential analysis approach, we calculated Pearson correlation coefficients between the RDs produced from each approach and the chance occurrence of the two-event sequence (i.e., the expected probability of the A cell), as estimated by the method’s estimate of the chance sequence or co-occurrence of the two key events. For each sequential analysis approach, the expected probability of the A cell was calculated as the product between the simple probabilities of Event X and Event Y. Indices that were least correlated with the expected probability of the A cell (i.e., r closest to zero) were identified as being the most interpretable.

Results

Description of event streams

The random process generated event streams with an average of 450 total events ($SD = 271$, Range = 11–1,110). Table 1 depicts the simple probabilities of each event and the expected A cell probabilities estimated by each sequential analysis method. With the exception of the time window approach, each method of sequential analysis produced similar mean simple probabilities of each event. The time window method produced discrepant simple probability estimates for Events X and Y because of the distinct estimations of the base rates of each event that are unique to this approach. In addition, the insertion of pauses as a third event type produced lower mean simple probabilities of each event for the modified event lag approaches (EL-CP and EL-NCP), relative to those estimated by the traditional event lag method. This occurred because

Table 1 Simple probabilities and expected A cell probabilities across sequential analysis approaches

Approach	Event X SP			Event Y SP			Expected A Probability		
	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	Range
EL	.51	.26	(.02–.98)	.49	.26	(.02–.98)	.18	.06	(.02–.25)
CI	.40	.28	(.003–.99)	.39	.27	(.003–.98)	.14	.14	(.0002–.67)
TW	.40	.28	(.003–.98)	.06	.06	(.0003–.30)	.02	.02	(.00002–.13)
EL-CP	.33	.26	(.002–.96)	.32	.25	(.003–.96)	.07	.06	(.0002–.23)
EL-NCP	.39	.25	(.005–.96)	.38	.25	(.01–.96)	.09	.05	(.003–.23)

SP = simple probability, EL = event lag, CI = concurrent interval, TW = time window, EL-CP = event lag with contiguous pauses, EL-NCP = event lag with noncontiguous pauses

adding the pause events increased the number of total events analyzed relative to the event lag method.

The frequency values corresponding to the 25th and 75th percentiles were 218 and 658, respectively. The low- and high-frequency subgroups each had 1,250 event streams, and the remaining 2,500 event streams were in the moderate-frequency subgroup. The low-frequency, moderate-frequency, and high-frequency subgroups of event streams had means of 127 (*SD* = 53.7), 424 (*SD* = 124.2), and 825 (*SD* = 115.5) events per stream, respectively.

Primary results

Means, standard deviations, and *d* statistics for the RDs produced by each sequential analysis approach are displayed in Table 2. Regardless of frequency subgroup, the effect sizes of RDs produced from the EL-CP method were the closest to zero (range in *d* values = 0.004–0.02). For the event lag and EL-NCP approaches, the effect sizes of the RDs were negatively biased, with the least accurate results for the low-frequency subgroup. For the concurrent interval and time window approaches, the effect sizes of the RDs were also negatively biased, with the least accurate results for the high-frequency subgroup. We found the rankings of relative accuracy to be the same when subgroups were created on the basis of Event X frequency or Event Y frequency (rather than total event frequency).

Pearson correlation coefficients between the RDs and the expected probabilities of the A cell as estimated by each type of contingency table are presented in Table 3. The contingencies estimated from the EL-CP approach were not correlated with the chance sequencing of Event X and Event Y (i.e., expected probabilities of the A cell; *r* = .001). The contingencies estimated from the remaining approaches were associated with chance sequencing of Event X and Event Y to a nontrivial degree (Cohen, 1988, identified *r* values >|.10| to be nontrivial). The event

lag, concurrent interval, and time window approaches produced contingences that were negatively correlated with chance sequencing, whereas the EL-NCP approach produced contingencies that were positively correlated with chance sequencing.

Table 2 Descriptive statistics for risk difference across sequential analysis approaches and total event frequency subgroups

Approaches per Subgroup	RD		
	<i>M</i>	<i>SD</i>	<i>d</i>
All Data			
Event lag	-.1755	.20	-0.88
Concurrent interval	-.0328	.07	-0.45
Time window	-.0083	.01	-0.59
Event lag with contiguous pauses	.0006	.06	0.01
Event lag with noncontiguous pauses	-.0745	.11	-0.69
Low-Frequency Subgroup			
Event lag	-.1984	.20	-0.98
Concurrent interval	-.0158	.09	-0.18
Time window	-.0018	.01	-0.17
Event lag with contiguous pauses	.0016	.09	0.02
Event lag with noncontiguous pauses	-.1871	.13	-1.46
Moderate-Frequency Subgroup			
Event lag	-.1860	.21	-0.90
Concurrent interval	-.0341	.06	-0.53
Time window	-.0073	.01	-0.63
Event lag with contiguous pauses	.0002	.06	0.004
Event lag with noncontiguous pauses	-.0530	.07	-0.76
High-Frequency Subgroup			
Event lag	-.1313	.17	-0.76
Concurrent interval	-.0469	.06	-0.73
Time window	-.0170	.02	-1.00
Event lag with contiguous pauses	.0004	.04	0.01
Event lag with noncontiguous pauses	-.0048	.04	-0.13

RD = risk difference

Table 3 Pearson correlation coefficients between risk difference and chance occurrence of the sequence of interest as estimated by each contingency table type

Approach	$r(\text{RD}, \text{ExpA})$
Event lag	-.11
Concurrent interval	-.21
Time window	-.45
Event lag with contiguous pauses	.001
Event lag with noncontiguous pauses	.26

RD = risk difference; ExpA = Expected probability of A cell calculated from each type of contingency table

Discussion

We conducted the present simulation study to compare the accuracy and interpretability of the contingency indices produced by five methods of sequential analysis (i.e., three existing methods and two proposed modifications of the event lag method). Our results support the following three conclusions: (1) across sequential analysis methods, the EL-CP approach produces the most accurate index of contingency; (2) the EL-CP approach is the only approach for which the accuracy of the contingency estimates does not depend on the total frequency of events; and (3) the EL-CP approach produces indices of contingency that are least correlated with the chance occurrence of the two-event sequence.

Recommendations and implications

We recommend the EL-CP method as the sequential analysis approach most likely to produce accurate and interpretable contingency estimates. This approach requires (a) the use of a continuous timed-event method of data collection and (b) the identification of a pause duration, decided by the investigator. The pause duration should reflect the upper limit of temporal proximity between the two events of interest and will likely be specific to the research question and behaviors or events under investigation. A primary advantage of this method is that it solves the *total event types* problem associated with the traditional event lag approach. That is, investigator decisions on the number of event types or behaviors to code would not influence the accuracy or interpretability of each contingency analysis. One would need a computer program allowing multiple events to be coded, with an option for users to indicate the two key events for each sequential analysis. The program would then disregard non-key event types and insert pauses of a user-specified duration before generating each 2×2 contingency table. A recently modified version of the Multi-Option Observation System for Experimental Studies software (MOOSSES; Tapp, Wehby, & Ellis, 1995) enables the EL-CP approach (<http://mooses.vueinnovations.com/>).

It is worth noting the potential implications of using the sequential analysis approaches that we have found to produce less accurate and interpretable contingency indices than the EL-CP approach. Our results suggest that stripping out non-X and non-Y events and estimating contingencies using the event lag (without pauses) or EL-NCP approaches may produce negatively biased estimates of contingency, especially under conditions in which the total frequency of key events is relatively low. The concurrent interval and time window approaches also may produce negatively biased estimates of contingency, but especially under conditions in which the total frequency of key events is relatively high. In the common scenario that researchers hypothesize positive contingencies between two events, the implication of using these less accurate methods is that they might be more likely to identify a negative contingency or no contingency when a zero or positive contingency exists (i.e., an elevated Type II error rate). In addition, our results suggest that any practical benefit of using interval sampling over continuous timed-event sampling is outweighed by the less accurate and interpretable estimates of contingency produced from the concurrent interval method.

Limitations

As is the case with most simulation studies, the extent to which our results apply to observational data with different characteristics than those of the data streams we generated is unknown. The original event streams we generated produced mean frequencies of 230 for Event X and 220 for Event Y, which represent relatively high base rates of key events (3.83 and 3.67 per minute, respectively). In addition, our generated event streams represented 1 h of observation time, which might be considered on the high side of typical session durations for behavioral observation studies. Although we randomly selected simple probabilities of each key event from uniform distributions between .01 and .60 when generating the data, we acknowledge that in some situations the simple probabilities of key events could be lower than those reflected in the present data set. In addition, we defined subgroups based on the 25th and 75th percentiles of the total frequency of key events in the original event streams. We chose to use the total frequency of key events instead of simple probabilities because using the latter would result in selecting different event streams among sequential analysis methods, which would produce an ambiguous reason for differences in mean contingencies among the five methods. Our subgroup analysis allowed comparisons of accuracy across higher and lower frequencies relative to one another; the low-frequency subgroup of event streams represented streams with mean frequencies of 64 for both Event X and Event Y (1.07 per minute for each key event). Although we cannot know the extent to which our results would generalize to observational sessions

with lower key event frequencies or of shorter durations, we are aware of the difficulties of estimating contingencies with few instances of the key event types (Yoder & Symons, 2010). Thus, one might question the appropriateness of conducting sequential analyses from particularly brief sessions or for key events occurring at particularly low base rates.

The present simulation does not speak to the relative accuracy and interpretability of sequential analysis methods for event streams in which the *duration* of events is relevant. For example, researchers may wish to estimate the contingency between engaging in a particular state (e.g., on-task behavior or active engagement) after a particular event type in which duration is not thought to be relevant (e.g., a discrete teacher prompt) or during an event type in which duration is thought to be relevant (e.g., 1:1 instructional format). In the present study, we only modeled the onset of event types. Thus, our results may not apply to sequential analyses in which the duration of one or both events is relevant to the contingency of interest. Finally, the simulation does not inform sequential analyses of non-repeatable events. We recognize that some researchers believe that some research questions are best addressed using methods that enable the use of structural zeros in cells that would ordinarily be used for sequences of repeating codes (Bakeman & Quera, 1995).

Future research

The results of this study suggest several avenues for future research. First, because the EL-CP method was shown to produce the most accurate estimates of contingency in the present simulation, evaluating strategies to empirically determine appropriate pause durations given the research questions and events under investigation may be worthwhile. Nearly all research questions involving sequential analysis require the researcher to define how temporally proximal the two key events must be to be considered an instance of the sequence of interest. For the EL-CP method, this boundary of temporal proximity is defined when selecting the duration of pauses, or the amount of time elapsing between Event X and Event Y that should prevent the event pair from being tallied in Cell A. In previous sequential analysis research, including our own, temporal proximity often has been defined arbitrarily. The selected pause duration will, however, influence how event pairs are tallied in the 2×2 contingency table, and thus will impact the resulting contingency estimate. Of course, the “appropriateness” of pause durations almost certainly will depend on the type of research question, the events or behaviors under investigation, and the context in which behavior is being measured.

Second, additional simulation studies may be needed to examine the relative accuracy and interpretability of the contingencies produced by the five sequential

analysis methods for situations in which event duration is important. After generating streams of event onsets as described in the present study, the duration of events could be randomly generated within a range that is realistic for research questions involving contingencies in which event duration is relevant.

Third, applications of the EL-CP approach to actual observational data will be needed for further validation, including conditions in which the total frequencies of key events are lower than those modeled in the present study. Because we do not know the “true” contingencies that exist in a given clinical or educational setting, it is likely not possible to address questions of the *absolute accuracy* of non-zero contingency estimates produced from sequential analyses in applied settings. Even in the case of simulation studies, we are not aware of a way to generate a set of behavior streams with a known, non-zero mean without relying on one of the very methods under comparison. Thus, it is likely that the absolute accuracy of the contingency indices produced by sequential analysis approaches can be evaluated only when the known mean is zero. Future research may, however, evaluate the *construct validity* of the EL-CP method when used to estimate contingencies in applied settings. Because one common purpose of sequential analysis is to identify event sequences that potentially represent an immediate causal relation (i.e., the occurrence of Event X has an immediate influence on the occurrence of Event Y; Yoder & Symons, 2010), the EL-CP method could be applied to pairs of events for which functional relations have and have not been identified. The extent to which the EL-CP method (a) produces positive contingencies for event pairs that do represent a functional relation and (b) produces zero or negative contingencies for event pairs that do not represent a functional relation could further inform the utility of this approach.

Conclusion

For many years, sequential analysis methodology has offered researchers a tool to capture an additional dimension of observational data beyond summary levels of behavior. Such methods allow the quantification of temporal relations between two events or behaviors as they occur in real time. Because several methods of sequential analysis now exist, each involving different combinations of behavior sampling, data representation, and 2×2 contingency table construction, it is important to develop a clearer understanding of the strengths and limitations of each approach so that new and improved methods may be developed. The present study takes a step in this direction by proposing the EL-CP approach as a promising new alternative that may improve the accuracy and interpretability of contingency estimates.

Author note This research was supported in part by National Institute of Child Health and Human Development Grant P30HD15052 to the Vanderbilt Kennedy Center for Research on Human Development.

References

- Bakeman, R., & Dorval, B. (1989). The distinction between sampling independence and empirical independence in sequential analysis. *Behavioral Assessment, 11*, 31–37.
- Bakeman, R., & Quera, V. (1995). Log-linear approaches to lag-sequential analysis when consecutive codes may and cannot repeat. *Psychological Bulletin, 118*, 272–284. doi:10.1037/0033-2909.118.2.272
- Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge, UK: Cambridge University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Eckert, T. L., Martens, B. K., & DiGennaro, F. D. (2005). Describing antecedent-behavior-consequence relations using conditional probabilities and the general operant contingency space: A preliminary investigation. *School Psychology Review, 34*, 520–528.
- Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions* (Version 5.1.0 [updated March 2011]). London, UK: The Cochrane Collaboration. Available from www.cochrane-handbook.org
- Lloyd, B. P., Kennedy, C. H., & Yoder, P. J. (2013). Quantifying contingent relations from direct observation data: Transitional probability comparisons versus Yule's Q. *Journal of Applied Behavior Analysis, 42*, 479–497.
- Marion, S. D., Touchette, P. E., & Sandman, C. A. (2003). Sequential analysis reveals a unique structure for self-injurious behavior. *American Journal on Mental Retardation, 108*, 301–313.
- Martens, B. K., Gertz, L. E., Werder, C. S., Rymanowski, J. L., & Shankar, K. H. (2013). Measures of association in contingency space analysis. *Journal of Mathematical Psychology, 59*, 114–119.
- Martens, B. K., Gertz, L. E., Werder, C. S., & Rymanowski, J. L. (2010). Agreement between descriptive and experimental analyses of behavior under naturalistic test conditions. *Journal of Behavioral Education, 19*, 205–221.
- Martin, P., & Bateson, P. (2007). *Measuring behavior: An introductory guide* (3rd ed.). Cambridge, UK: Cambridge University Press.
- Pineda, A. Q., Cole, D. A., & Bruce, A. E. (2007). Mother–adolescent interactions and adolescent depressive symptoms: A sequential analysis. *Journal of Social and Personal Relationships, 24*, 5–19.
- Roberts, C., Yoder, P. J., & Kennedy, C. H. (2005). Descriptive analysis of epileptic seizures and problem behavior in adults with developmental disabilities. *American Journal on Mental Retardation, 110*, 405–412.
- Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.
- Symons, F. J., Hoch, J., Dahl, N. A., & McComas, J. J. (2003). Sequential and matching analyses of self-injurious behavior: A case of overmatching in the natural environment. *Journal of Applied Behavior Analysis, 36*, 267–270.
- Symons, F. J., Tapp, J., Wulfsberg, A., Sutton, K. A., Heeth, W. L., & Bodfish, J. W. (2001). Sequential analysis of the effects of naltrexone on the environmental mediation of self-injurious behavior. *Experimental and Clinical Psychopharmacology, 9*, 269–276.
- Tapp, J., Wehby, J. H., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instruments, & Computers, 27*, 25–31.
- Vernon, T. W. (2014). Fostering a social child with autism: A moment-by-moment sequential analysis of an early social engagement intervention. *Journal of Autism and Developmental Disorders, 44*, 3072–3082.
- Vollmer, T. R., Borrero, J. C., Wright, C. S., Van Camp, C., & Lalli, J. S. (2001). Identifying possible contingencies during descriptive analyses of severe behavior disorders. *Journal of Applied Behavior Analysis, 34*, 269–287.
- Wilde, L., Silva, D., & Oliver, C. (2013). The nature of social preference and interactions in Smith–Magenis syndrome. *Research in Developmental Disabilities, 34*, 4355–4365.
- Yoder, P. J., Davies, B., Bishop, K., & Munson, L. (1994a). The effect of adult topic continuing wh-questions on conversational participation in children with developmental disabilities. *Journal of Speech and Hearing Research, 37*, 193–204.
- Yoder, P. J., Davies, B., & Bishop, K. (1994b). Reciprocal sequential relations in conversations between parents and children with developmental delays. *Journal of Early Intervention, 18*, 362–379.
- Yoder, P., & Symons, F. (2010). *Observational measurement of behavior*. New York, NY: Springer.
- Yoder, P. J., & Tapp, J. (2004). Empirical guidance for time-window sequential analysis of single cases. *Journal of Behavioral Education, 13*, 227–246.
- Yule, G. V., & Kendall, M. G. (1957). *An introduction to the theory of statistics*. London, UK: Griffin.