



Exploratory Factor Analysis of the VB-MAPP: Support for the Interdependency of Elementary Verbal Operants

Jordan Belisle¹ · Mark R. Dixon² · Albert Malkin² · Joshua Hollie² · Caleb R. Stanley³

Accepted: 23 October 2020 / Published online: 1 January 2021
© Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Competing viewpoints on the independency or interdependency of Skinner's verbal operants have been discussed in the literature and with empirical support for both positions generated using single-case research methods. Our study provides support for the interdependency of the verbal operants using items contained in the Verbal Behavior Milestones Assessment and Placement Program as a measure of broader skill acquisition in each verbal operant category and across skill complexity levels. The result of an exploratory factor analysis conducted across 85 participants with autism suggested that items did yield factors consistent with the verbal operants, rather items appeared to cluster in terms of skill complexity producing a best-fit 2 factor model. Together with prior research showing untrained cross-operant transfers, results fail to support Skinner's verbal behavior taxonomy distinguishing between the verbal operant categories as independent constructs, with implications for how behavior scientists and analysts describe language development and assess and treat language deficits of individuals with autism.

Keywords Verbal behavior · Interdependency · Factor analysis · VB-MAPP exploratory factor analysis of the vb-mapp: support for the interdependency of elementary verbal operants

✉ Jordan Belisle
jbelisle@missouristate.edu

¹ Psychology Department, Missouri State University, Springfield, USA

² Southern Illinois University, Carbondale, USA

³ Utah Valley University, Orem, USA

Introduction

During the past 60 years, Skinner's Verbal Behavior (1957) has received attention within and beyond the field of behavior analysis. Several behavior analytic training assessment protocols available that are grounded in Skinner's verbal operant framework (Dixon, Belisle, McKeel, et al. 2017) and some commonly used assessment protocols that focus on Skinner's verbal operants include the Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP; Sundberg 2008), the Assessment of Basic Language and Learning Skills Revised (ABLLS-R; Partington 2008), and the PEAK Direct Training (Dixon 2014a) and Generalization (Dixon, 2014b) modules. Each of these assessments operate under the assumption, first proposed by Skinner (1957), that the verbal behavior of a speaker emerges through reinforcement from a listener and that verbal behavior can be discussed scientifically as emerging in verbal operant categories, such as tacts, mands, and intraverbals. Recent psychometric research has emerged evaluating the validity of several of these assessments. Malkin, Dixon, Speelman, and Luke (2017) provided data supporting the convergent validity between participant results on the PEAK Direct Training assessment and ABLLS-R, and Dixon, Belisle, Stanley, et al. (2014) conducted a similar study correlating participant scores on the PEAK Direct Training and Generalization assessments with scores on the VB-MAPP. In addition, intervention procedures developed from these assessments can result in the acquisition of new language skills (e.g., McKeel, Rowsey, Belisle, Dixon, and Szekely 2015), corresponding with overall growth in empirical research teaching elementary verbal operants (Dymond, O'Hora, Whelan, and O'Donovan 2006), and primarily conducted with individuals with autism (Dixon, Small, and Rosales 2007; Sundberg and Michael 2001).

Although Skinner's analysis has undoubtedly led to research and technologies for teaching elementary verbal skills (DeSouza, Akers, and Fisher 2017; Dymond et al. 2006), some debate exists as to whether the verbal operants are indeed functionally independent (Fryling, 2017; Gamba, Goyos, and Petursdottir 2015). If the operants are independent, then the development of a given operant skill, such as manding, should not correspond with the development of other unrelated skills, such as tacting or intraverbal responding. Assuming verbal operants are independent is central in Skinner's analysis, as is captured in his quote, "When the response Doll! has been acquired as a mand, however, we do not expect that the child then spontaneously possesses a corresponding tact of similar form" (Skinner 1957, p. 187). Some early single case studies have supported the potential functional independence of verbal behaviors. Partington and Bailey (1993) demonstrated that reinforcing tact responses in preschoolers reliably led to increases in the probability of tact responses, but did not lead to corresponding increases in intraverbal responses of the same topography. Shillingsburg, Kelley, Roane, Kisamore, and Brown (2009) demonstrated that when three boys were taught the verbal responses "Yes" and "No" as one of either a tact, mand, or intraverbal response, the participants did not demonstrate a functional transfer to the remaining two untaught operants. Conversely, several single case studies have

supported the functional interdependence of the verbal operants; or, when a verbal topography is taught as one operant, the topography will be used as another operant without direct training. May, Hawkins, and Dymond (2013) demonstrated that when adolescents with autism were taught several tact responses, intraverbal responses of the same established topographies emerged without direct training. Grannan and Rehfeldt (2012) also showed that when tact training was conducted in a match-to-sample arrangement, intraverbal responding emerged. Several additional studies have been conducted with this same general finding (e.g., Grow and Kodak 2010), such as by demonstrating untrained tact to mand transfers (Wallace, Iwata, and Hanley 2013).

Although some authors have discussed the potential for consolidating the functional interdependence of verbal operants with Skinner's original theory (Carr and Miguel 2013), if learning a verbal topography as one operant leads to the untrained use of the same topography in another operant context, then the utility of Skinner's taxonomy delineating the operants as independent constructs is threatened. As discussed by Fryling (2017), constructs are the words that scientists use to describe events, where the independent operants discussed by Skinner are tools used by some¹ behavioral scientists to describe how language develops. In science, constructs are only useful insofar as the constructs describe distinct events, and if the verbal operants are interdependent (e.g., learning to tact corresponds with learning to mand and to respond intraverbally), then verbal operants are not distinct constructs. Whereas single case studies have led to mixed findings regarding functioning independence and interdependence (May, Hawkins, and Dymond 2013; Shillingsburg et al. 2009), an alternative strategy involves the use of inferential statistical procedures to determine whether the operants indeed represent independent constructs. Exploratory factor analyses (EFAs) are designed for use with substantially larger participant samples than those used in single-case research designs and can determine the degree to which independent observations represent separate constructs, or factors.

For example, Nicholson, Konstantinidi, and Furniss (2006; extending previous research by Paclawskyj, Matson, Rush, Smalls, and Vollmer, 2000) conducted an EFA of items contained in the Questions About Behavior Function (QABF; Matson and Vollmer, 1995) assessment tool that is used to isolate the environmental function for a given challenging behavior, where an assumption inherent in the development of the assessment was that the five identified functions of behavior are independent or distinct constructs (e.g., a behavior is generally either attention-maintained or escape-maintained, instead of simultaneously maintained by all of the potential functions). Results of the EFA supported that items endorsing a given function clustered with other items endorsing the same function and diverged sufficiently from items endorsing the remaining other functions. Therefore, the results support the utility in delineating functions of challenging behavior as independent constructs.²

¹ We specify "some" behavioral scientists because of disagreements regarding the usefulness of Skinner's account in general (e.g., Hayes, Barnes-Holmes, and Roche, 2001).

² See Belisle, Stanley, and Dixon (2017) for data suggesting that distinguishing between functions of challenging behavior in this way may provide less utility as language develops and behavior becomes increasingly rule-governed.

Given the emergence of several comprehensive verbal behavior assessments over the past decade, we can now assess for more broad verbal behavior repertoires of participants and in substantially larger samples than have participated in single-case research on the functional independence of verbal operants.³ If the operants represent distinct constructs, we would expect that assessment items that endorse a given operant, such as tacting, are distinctly related to other items that endorse the same operant (i.e., tacting). An advantage of using a quantitative statistical approach rather than a narratively described single-case design is that the results of such an analysis are not left up to the interpretation of the researcher and susceptible to biases of one interpretation of Skinner's account over another; rather, this approach allows for a mathematical interpretation as to whether the operants are independent constructs, or if instead the operants do not emerge as unique events, limiting the construct validity of Skinner's account in describing how verbal behavior emerges.

The VB-MAPP is one assessment that appears to be well suited for an analysis of the independency or interdependency of Skinner's verbal operants. As noted by the developer (Sundberg 2008), the assessment was developed from Skinner's analysis. Items contained in the VB-MAPP's milestones assessment are arranged in two ways: by verbal operant category (e.g., tact, mand) and by complexity level (e.g., Level 1 mand and Level 2 mand). Each item is scored as either 0, ½, or 1 based on direct testing, observation, or some combination, and items are then summed to arrive at a score ranging from 0 to 5 in each domain (i.e., level/operant category). Because the VB-MAPP delineates the skills by operant categories consistent with Skinner's theoretical interpretation, unlike other assessments such as the PEAK Direct Training and Generalization modules, conducting an EFA on participant scores for each item can determine whether items contained in each operant category in fact represent distinct factors or constructs. In addition, because the VB-MAPP also differentiates items by levels of complexity, an alternative outcome that may be expected from an EFA is that items do not cluster based on operant category, but by skill complexity. Indeed, when Rowsey, Belisle, Dixon, and colleagues (Rowsey, Belisle, and Dixon 2015; Rowsey, Belisle, Stanley, Daar, and Dixon 2017) conducted an EFA on the PEAK Direct Training and Generalization modules, both of which have been correlated with the VB-MAPP (Dixon, Belisle, Stanley, et al. 2015), results generated 4-factor models for each assessment that progressed in skill complexity, rather than by the type of operant assessed.⁴ Neither of the PEAK assessments were designed to measure independent verbal operant categories, nor were the assessments exclusively based on Skinner's account (Dixon, 2014a, b), so both were less well suited

³ We do not intend to downplay the single-case research that has been conducted to answer this research question; rather, that a larger, statistical analysis can supplement these studies and add additional depth to answering this research question.

⁴ Although the PEAK Direct Training and Generalization assessments do not offer scores for each "verbal operant," several items do note the operant "e.g., *tacting* planets," and results of the EFAs of these PEAK modules do not indicate that items in the same operant class necessarily fall into the same factor (Rowsey et al. 2015, 2017).

Table 1 Participant demographic information

Characteristics	N	%
Age		
5–8	15	17.65
9–12	22	25.88
13–16	23	27.06
17–20	25	29.41
Male	71	83.53
Female	14	16.47
Diagnosis	85	100

$M = 13.36$; $SD = 4.432$; $Skewness = -0.1423$; $Kurtosis = -1.118$

than the VB-MAPP to an EFA of the independency/interdependency of the verbal operants.

Therefore, the purpose of the present study was to conduct an EFA of items contained in the VB-MAPP to provide a statistical analysis of the independency or interdependency of Skinner’s verbal operants. We used an exploratory analysis to determine whether items in the VB-MAPP factored in terms of the verbal operant being assessed for, or in terms of skill complexity, without an a priori theoretical model as would be imposed in the confirmatory factor analysis. Factor loading in terms of verbal operant category would support the functional independence of the verbal operants, or that more elementary forms of verbal operant behavior (e.g., echoic, mand) emerge earlier and independent of more complex forms of verbal operant behavior (e.g., tact, intraverbal). Alternatively, we anticipated that items may instead cluster in terms of the complexity of the skill being assessed, more precisely described in terms of the “levels” in the VB-MAPP rather than operant categories. Such an outcome would suggest that, although verbal behavior may be developed through operant conditioning, the distinct and presumably independent “verbal operant” constructs developed in Skinner’s taxonomy may not provide the best possible analysis. In our study, we involved only participants with autism to approximate a sample that is most likely to benefit from verbal behavior training technologies (Dymond et al., 2006; Dixon, Small, & Rosales, 2007), and as such are most likely to benefit from the VB-MAPP or related assessments.

Method

Participants, Setting, and Materials

This study was based on data collected from participants recruited from a public school located in the Midwest of the USA. Data from participants included 85 students, including 14 females and 71 males. Participants’ ages ranged from 5 to 20 years ($M = 13.5$, $SD = 4.4$; Table 1). Participation in the study was limited to school-aged children diagnosed with autism. The above population was recruited based on the assumption that children with autism would likely benefit from and

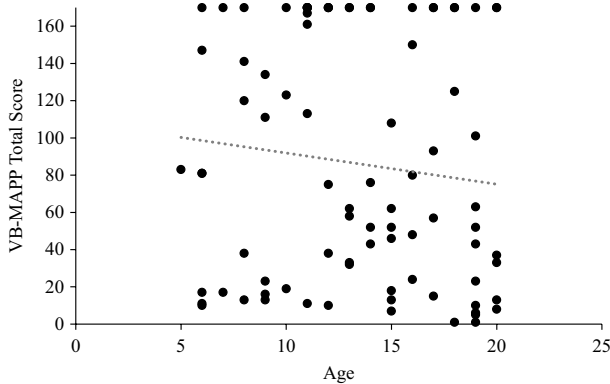


Fig. 1 Correlation between VB-MAPP score and age. $R^2=0.0131$

are representative of the population that most commonly participates in language-based behavior analytic interventions and research and are the most likely to benefit from clinical interventions derived from Skinner’s analysis of verbal behavior and the VB-MAPP (Dixon et. al. 2007). Informed consent provided by participants’ primary caregivers, which was a prerequisite for inclusion in the sample. We state here that readers should be cautious when interpreting the generality of these results outside of the current sample given all participants were recruited from the same school and area and the sample size is smaller than is generally recommended within EFA research (Costello and Osbourne 2005). As noted by Costello and Osbourne (2005), the general rule “more is better” (p. 5) applies to generality, but the obtained outcomes speak more locally to the appropriateness of a sample within a given study. Therefore, we address this point in more detail in “**Discussion**” section.

Additional consideration was given to participant inclusion criteria; given that participants’ chronological ages did not correspond with the age-related levels of the VB-MAPP (e.g., 0–18, 18–30, and 30–48 mos.; described in further detail below) and were skewed toward late adolescence, overall scores were analyzed using Pearson product correlation to determine whether age and overall score were related. A weak negative correlation was found using a two-tailed analysis ($r = -0.11$) (Fig. 1). Thus, it appears that age did not significantly influence language abilities in the recruited sample, which implies that the participants may benefit from additional language-based assessment and intervention, based on the VB-MAPP (Figs. 2, 3).

The materials used in the study consisted of the VB-MAPP, which is a criterion-referenced tool that serves as an assessment, curriculum guide, and skill tracking system (Dixon, Belisle, Stanley, et al. 2014). The VB-MAPP contains 170 skills, divided into 3 levels based on developmental levels associated with typical language development across childhood, ranging in age from 0 to 48 months (Sundberg 2008). The first level is based on language development that occurs within the first 18 months; items assessed within the first level include verbal capabilities that exemplify Skinner’s elementary verbal operants (e.g., mand, tact, listener responding, and echoic repertoires) and additional skills that serve as prerequisites for more

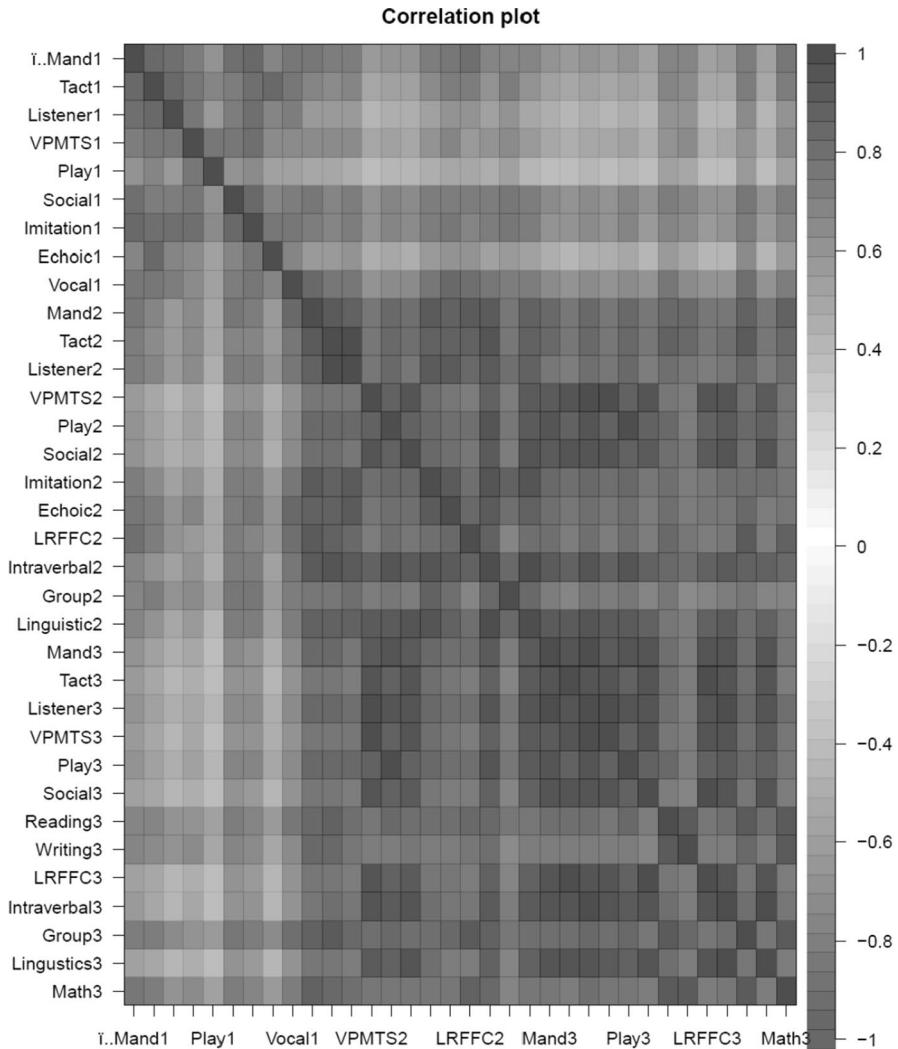


Fig. 2 Correlation plot with all items

complex language and social behavior (e.g., visual performance/matching to sample, imitation, play, social skill, and vocal repertoires). The subsequent level of the VB-MAPP has a focus on language that typically develops between the ages of 18 and 30 months; these skills expand on the prerequisite skills of level one and focus on additional requesting and simple conversational skills (Sundberg 2008). The skills targeted in addition to the categories of skills in level one include: listener responding by feature, function, and class, intraverbal, group and classroom behavior, and the linguistic structure of verbal responding (Sundberg 2008). Finally, the third level of the VB-MAPP aims to assess skills typically developed between 30 and 48 months of age, including an increased level of complexity with the elementary

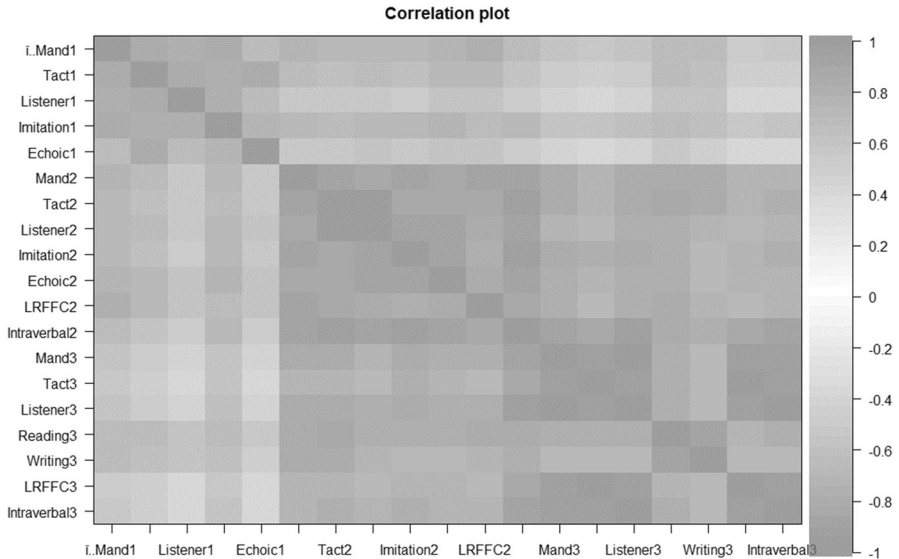


Fig. 3 Correlation plot with VB items only

verbal operants included in the previous level, excluding imitation and echoic behavior, along with additional academic skills (e.g., math, reading, and writing).

A variety of materials related to the test items were used, including flashcards, toys, and other common stimuli typically found in a home or classroom. For example, the assessment item 5-M requires that participants tact 10 items; including pictures of common objects, people, body parts, or pictures. Stimuli was prepared specific to each test item to directly observe participants responses in the presence of specified contextual arrangements.

All assessments were carried out by graduate students who had experience providing instruction directly to the participants for at least six months prior to the start of the current study. Additionally, graduate students met the recommendations of assessor experience, knowledge, and practice. Specifically, students were familiar with and able to apply assessment procedures related to Skinner’s (1957) analysis of verbal behavior, motivating operations, prompt levels, and basic linguistic structure and development. Student assessors were provided with an assessment sheet that contained the title, brief description, and criterion for mastery for each of the 170 items within the VB-MAPP. Each assessment item had two boxes to check, to indicate either a “yes,” or “no” regarding the participant’s ability to complete the item successfully. The student assessors were blind to the purpose of the present study.

The student observers were instructed to score skills that were directly observed and/or contrived. Skills were all observed in the natural environment and through instruction in a classroom setting. Observers followed assessment procedures unique to each skill, as indicated by the VB-MAPP guide; skills were either “observed,” “tested,” “either observed or tested,” or “timed,” based on the assessor’s history of observation of each individual repertoire. The guide indicates that skills that are

clearly within the participants' repertoire may be endorsed without testing and may be endorsed; however, if uncertainty regarding the participants' ability to perform a skill, direct testing must be conducted. Additionally, testing was terminated if less than 2 out of 5 items were endorsed on the previous level. Each item within the VB-MAPP was assigned a score of either 0 or 1. The VB-MAPP allows for items to receive partial scores, however, to decrease the likelihood of subjective scoring, increase consistency, and maintain conservative scoring, data collectors were instructed to score skills "1" if the participant met the full criteria; otherwise, skills received a score of "0." The purpose of the analysis was to determine the dependence or interdependence of the verbal operant categories represented in the data, rather than determining the validity of the assessment; thus, endorsing assessment items as either present or absent was determined to more closely align with this aim. Conversely, endorsing partial points may result in the effect of masking distinctions between latent variables. The above approach was most likely to yield meaningful results either confirming or refuting the relationships between verbal operant categories.

Data Analysis

Initial correlation analyses were carried out using R Studio (Version 1.0.136) to determine the correlation between each item of the VB-MAPP. Pearson correlation was found to be significant at the 0.01 level (two-tailed) for each item of the assessment. The data were subsequently assessed for suitability for a factor analysis via inspection of the determinant of the correlation matrix by calculating the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and the Bartlett's Chi-square test of sphericity. The determinant of the correlation matrix was 0.000, and the matrix was indicated as "not positive definite." Variables must be correlated to perform a factor analysis; however, high correlation can be problematic (Field, 2009). The above results indicate that the correlation matrix is singular. Multicollinearity is the probable cause given the significantly highly correlated assessment items.

A common recommendation to address multicollinearity is proceeding with the analysis with the elimination of any variables that may be highly correlated (e.g., $R > 0.8$) and/or reducing the number of variables analyzed. Given that the primary interest of the current study is the interdependence, or lack thereof, between the elementary verbal operants, the investigators eliminated all categories of the VB-MAPP from assessment, that arguably, do not explicitly target the assessment or acquisition of verbal behavior. The categories removed from further analysis include all levels of the following items: visual and match to sample, play, social, group, linguistic, and math. Following the reduction of variables, subsequent analyses of the determinant of the correlation matrix, KMO, and Bartlett's Chi-square test of sphericity were conducted again. The data were judged to be factorable.

The logic of factor analysis was used to determine the relationship between the assessment items represented in the VB-MAPP. Specifically, the EFA identifies variables that are highly related and unrelated to other variables; the highly related

variables are merged to create a new variable, to represent performance on the combined variables as a single score; ultimately, the resulting analysis and interpretation are parsimonious due to the reduction of many variables into a smaller number of meaningful categories (Huck 2012). The current study made use of the EFA to determine the clustering of items in the VB-MAPP in the hope of either the identification of the relationship of items according to verbal operant categories, complexity of verbal operant responses, or the potential that neither pattern may provide a robust account of the skills represented in the VB-MAPP. The three stages of conducting an EFA involve (1) extraction, (2) rotation, and (3) interpretation; each method will be described below.

The process of extraction refers to methods that aim to determine the number of factors that explain the observed covariation matrix; this process is was carried out by employing a principal component analysis (PCA) to extract the underlying factors among assessment items. Visual analysis was used to determine the number of factors above the elbow point of the plot. The factors to the left of the elbow of the Scree plot explain a greater degree of the variance observed. The process of rotation allows for the creation a simple structure, by reducing the dimensions of the data. The investigators noted that the although the number of variables was reduced to a sufficient level, we suspected that the factors would nonetheless be correlated. An appropriate method of rotation under the above circumstances is the Oblimin rotation. Specifically, the PCA used a direct Oblimin rotation and Kaiser normalization; we retained factors with an eigenvalue greater than 1 and subsequently evaluated factor loading, based on finding that the PCA retained two factors. Finally, Chronbach's alpha analyses were conducted to assess reliability among items that loaded onto each factor.

Results

An EFA was conducted on the 34 items of the VB-MAPP, the PCA using an Oblimin rotation with Kaiser normalization; the above method revealed that the Pearson correlation (two-tailed) between items was significant at the 0.01 level (Fig. 1). However, the KMO measure did not verify the sampling adequacy for the analysis, the correlation matrix had a determinant=0.000 (which falls below the minimum recommended value of 0.00001), and the matrix was indicated as “not positive definite.” The initial PCA extracted two components (Fig. 4). However, given the lack of stability in the data, nonessential items were removed from subsequent assessment (Fig. 5).

A second EFA was conducted on 19 items of the VB-MAPP, again, making use of PCA and Oblimin rotation with Kaiser normalization analyses; the Pearson correlation (two-tailed) between all items was again significant at the 0.01 level ($N=85$) (Fig. 4). However, the KMO measure met the criteria for sampling adequacy for the analysis, the correlation matrix had a determinant=Determinant=8.33 (which falls above the minimum recommended value of 0.00001.), the KMO measure of 0.939 and Bartlett's test of sphericity was significant $\chi^2(171)=2844.705, p<0.001$. According to Huck (2012), data are “factorable if the determinant is greater than

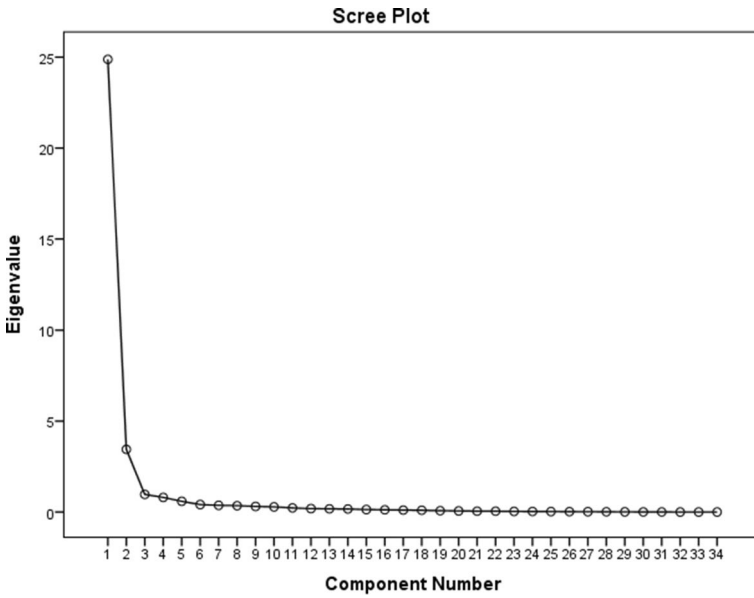


Fig. 4 Scree plot (All items). 2 components extracted. Extraction method: principal component analysis

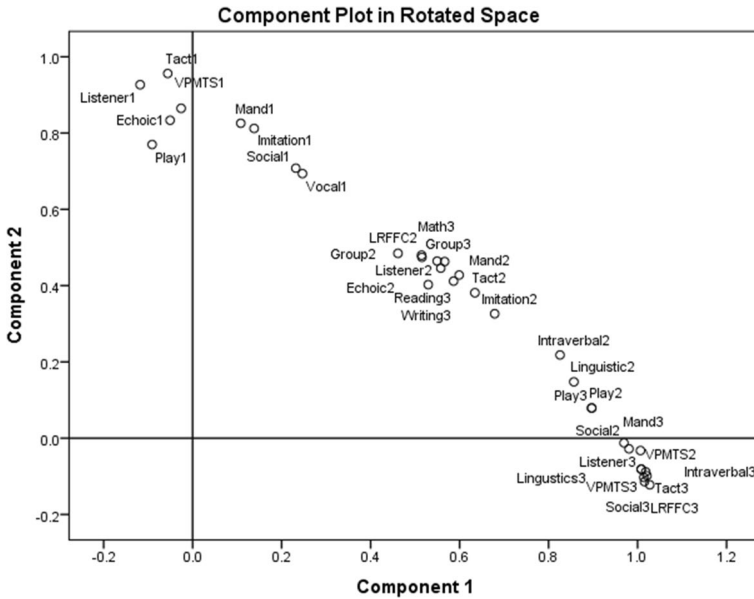


Fig. 5 Component plot in rotated space (all items)

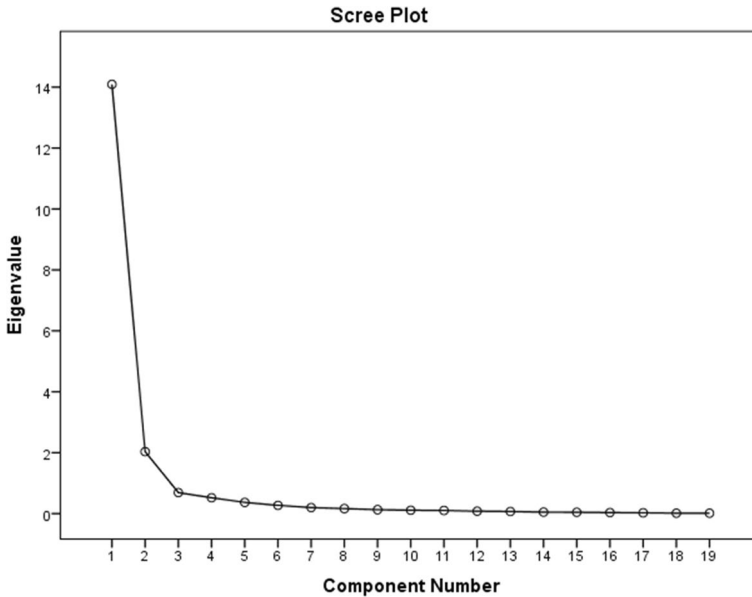


Fig. 6 Scree plot (VB items only). 2 components extracted. Extraction method: principal component analysis

0.00001, if the KMO measure of sampling adequacy is greater than 0.60” (p.487). The PCA indicated that two components had eigenvalues greater than 1; the components explained 74.168%, 10.694% of the total variance, respectively (Fig. 6). Table 2 shows the factor loadings after rotation (Oblimin with Kaiser normalization; converged in 6 iterations). The items that cluster on the same components suggest that components one and two represents complex and foundational verbal behavior repertoires, respectively.

It is noteworthy that the initial analysis, which included academic, group, and social skills, specified that those skills loaded onto component two. Visual inspection of the component plot in rotated space (Fig. 7) seems supports the above interpretation. However, it appears that there may be a relation between a third cluster of skills, although not a statistically significant relationship. However, the derived model does seem to be stable; reliability statistics indicated Chronbach’s alpha of 0.943 ($n = 14$) for component 1; and Chronbach’s alpha of 0.985 ($n = 21$) for component 2 items. Both components exceed the minimum criteria alpha of 0.7, indicating good overall reliability.

Discussion

The results reported in the current study extend research on the independency or interdependency of Skinner’s verbal operants (Grannan and Rehfeldt 2012; Shillingsburg et al. 2009) by conducting a larger-scale statistical analysis of broader

Table 2 Pattern matrix elementary verbal operants

	Component	
	1	2
Tact3	1.033	-.151
Intraverbal3	1.033	-.145
LRFFC3	1.029	-.167
Listener3	1.025	–
Mand3	1.011	–
Intraverbal2	.898	.128
Imitation2	.760	.241
Tact2	.741	.278
Mand2	.711	.314
Reading3	.679	.315
Listener2	.664	.351
Echoic2	.641	.379
Writing3	.635	.294
LRFFC2	.605	.405
Tact1	–	.930
Listener1	–	.908
Echoic1	–	.863
Mand1	.184	.789
Imitation1	.202	.769

Pattern matrix (elementary verbal operants only): extraction method: principal component analysis. Rotation method: oblimin with Kaiser normalization; a rotation converged in 6 iterations

verbal behavior development using the VB-MAPP. An EFA affords the advantage of isolating constructs that are independent (i.e., correlated with other factor-consistent items), and our results failed to provide support for the independency of the verbal operant categories of the VB-MAPP. For example, performance on level 1 tact items was more likely to be related to performance on level 1 mand items than to level 2 tact items; therefore, the results support that the operants appear to be interdependent as broader areas of skill development, at least psychometrically. We encourage readers to interpret this conclusion with caution because this does not discount the utility to training verbal operants per se; rather, it questions the utility of this taxonomy in describing differentiated assessment outcomes among learners. Although the verbal topography “Doll!” may have been directly taught in every possible way, we believe the common-sense view expressed by Michael (1988, p. 7) may be more likely, that “based on experience with normal children and adults, once a person has learned what an object is called ... it is reasonable to assume that when the object becomes important the learner will be able to ask for it without further training.” This becomes even more likely considering the tact to intraverbal transfers demonstrated by May, Hawkins, and Dymond (2013) and tact to mand transfers demonstrated by Wallace, Iwata, and

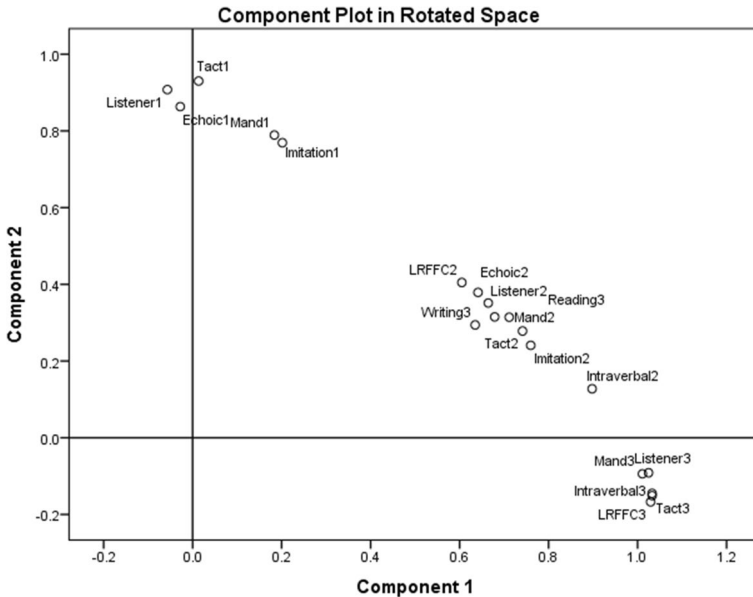


Fig. 7 Component plot in rotated space (elementary verbal operants only)

Hanley (2006). If true, this has potential implications regarding how educators go about establishing the verbal operants. As a learner’s performance on assessments such as the VB-MAPP improves, it may be the case that targeting single operants is less efficient than targeting multiple operants in order to build verbal complexity—as complexity appears to be the most consistent determinant of factor structure within now multiple assessments of verbal operant development.

This may threaten Skinner’s verbal behavior taxonomy of the independent operants (Grannan and Rehfeldt 2012; May, Hawkins, and Dymond 2013) at a theoretical⁵ level given any taxonomy provides a set of constructs that may or may not prove useful scientifically. To use manding as an example of this conceptual problem, although a mand *may* be acquired in the presence of a motivating operation and through the delivery of a specified reinforcer (Skinner, 1957) as has been demonstrated in prior research (Davis, Kahng, and Coryat, 2012; Endicott and Higbee, 2007; Howlett et al. 2011), there are also near infinite other ways that mands could be acquired. A learner may be taught to identify the “requested object” receptively or vocally as a tact, where a mand occurs as a cross-operant transfer (Kooistra, Buchmeier, and Klatt 2012; Wallace, Iwata, and Hanley 2006). An echoic may also adventitiously become a mand, in that the learner shows a delayed echoic in a novel

⁵ We use “theory” here to describe any account that extends some prior research to develop a complete set of constructs or descriptions that extend considerably from the prior research prior to empirical testing. Given Skinner’s *Verbal Behavior* (1957) was published in the absence of direct experimental testing of basic assumptions (e.g., independency of verbal operants), this qualifies as a theory in this sense.

context (e.g., hearing “Ball!” on television and then saying “Ball!” at the park) that produces the item despite the absence of a specific motivating operation or reinforcement history. These simple and contrived examples are given without consideration for the potentially vast complexity of verbal behavior, such as when a person combines several established tacts to produce an abstract mand, like “I want to be the kind of person that people go to with their problems and can help them achieve the life that they value.” As noted by Fryling (2017), distinguishing between the verbal operants, if they are interdependent, may simply provide a description of context in which a verbal response occurs (e.g., a *mand* is simply a description of a scenario in which a person makes a verbal response and a listener provides a specified reinforcer), rather than a complete analysis of how verbal behavior emerges (Skinner 1953, 1957).

Skinner’s account has led to the development of several evidence-based approaches that have been effective at teaching new skills to individuals with disabilities (Dymond et al. 2006; Sundberg and Michael, 2001), and we do not intend to discount this work. The results reported here and in prior studies on the interdependency of the verbal operants do not refute this body of research, as the observed events undoubtedly occurred; rather, the results question the conceptual interpretation of these results. More contemporary theoretical accounts offered in stimulus equivalence (Sidman 1971; Sidman and Tailby 1982) and relational frame theory (Hayes, Barnes-Holmes and Roche 2001) rarely refer to the verbal operants, but put forward constructs such as symmetry, transitivity, combinatorial entailment, and transfers and transformations of stimulus function that appear to be more so based on the complexity of a verbal behavior, rather than on the functional actions of the listener (see Hayes, Barnes-Holmes, and Roche 2001 for a detailed criticism of Skinner’s operants and the pitfalls of defining any behavior by the actions of another person). Unlike the verbal operants, research on relational responding has shown a clear progression from simple forms of derived responding earlier in life to more complex forms of relational responding later (Dymond et al. 2006; Hayes, 1996; Lipkens, Hayes, and Hayes 1993), supporting that the constructs in both equivalence and relational frame theory are distinct. Although a more robust statistical analysis of these constructs is required as in the present study, we use equivalence and relational frame theory only as examples of alternative accounts offering a taxonomy separate from Skinner’s analysis that appear to hold greater construct validity by emphasizing complexity.

Beyond conceptual implications that we describe above, the results have immediate practical implications for how we interpret and evaluate behavior analytic approaches to language training technologies. First, there are several behavioral language assessments that have been developed over the last decade, but only recently has any research been conducted to evaluate the validity of these assessments or the effectiveness of emergent intervention strategies (Dixon, Belisle, McKeel, et al. 2017). Thus, support for these assessments has largely been based on conceptual correspondence with Skinner’s account of verbal behavior (Sundberg 2008) instead of rigorous statistical analyses of assessment data. Given our results and prior research call into question the utility of Skinner’s account in delineating independent

or useful constructs, conceptual correspondence with this account is insufficient. Where research has been done on the construct validity of the VB-MAPP (current study) and other assessments (i.e., PEAK), the validated constructs are those that appear to be based on verbal complexity, not operant categorization. Thus, although our results do support a degree of construct validity for the levels of the VB-MAPP,⁶ differentiating items based on categories like “tact” and “mand” could be unnecessary psychometrically. We do contend, however, that this taxonomy has educational utility by delineating the different contextual features that could give rise to the target verbal behavior, such as arranging the opportunity or contriving motivating operations to evoke mand responses to strengthen this operant. Second, differentiating between the verbal operant categories is considered a strength of the VB-MAPP and some other assessments (Sundberg, 2008); however, given the operants are unlikely independent constructs, distinguishing between the operants on these assessments may be irrelevant or even illusory. A potentially better way to describe the skills assessed for in these assessments may be to provide component scores in terms of progressions in verbal behavior complexity, as is done in the PEAK Direct Training and Generalization modules and in the level system of the VB-MAPP.

Despite the above findings, there are several limitations that may be addressed in future research. First, although our sample size is approximately equal to the sample reported on EFAs conducted for the PEAK Direct Training (Rowsey, Belisle, and Dixon, 2014a, b) and Generalization (Rowsey et al. 2017) modules that are similar in construction to the VB-MAPP, more robust EFAs call for greater sample sizes to ensure generality of the results. That is not to say, however, that these results necessarily lack internal validity due to the sample size (Costello and Osborne, 2005). Generally, larger samples produce greater stability in the obtained model, but obtained values within the EFA ultimately speak to the stability of the obtained factors. As noted in the results, the KMO measure met the threshold of stability. In addition, as noted by Costello and Osborne (2005), factor loadings are considered adequate and stable if item communalities exceed 0.40 and if greater than 5 items appear within each factor, which was achieved in the present study. Greater samples in future research may, however, serve to reduce the probability of a type II statistical error. Another advantage of utilizing a greater sample size is that each of the 170 items in the VB-MAPP could be entered independently into the analysis, rather than as a single score for each level operant. Doing so would allow for a more precise analysis of factor loadings in terms of the operant categories within levels as well as across levels and should be a priority in future research.

A second limitation is that our sample skewed in terms of older participants, where the mean age was approximately 13 years. The VB-MAPP was developed for use with children below this age (i.e., ages 0–4 years); however, prior research has utilized the VB-MAPP with participants who exceed 4-years of age (Dixon, Belisle, Stanley, et al. 2014; Geiger, LeBlanc, Dillon, and Bates 2010) that experience

⁶ The multicollinearity observed in our results suggest that eliminating items that appear to approximately equally correspond in both factors may lead to greater construct validity for the VB-MAPP as an assessment instrument.

diminished language functioning such as is expected in an autism sample. Despite the high ages reported in the current study, we observed variability in total scores across the sample, suggesting that the complexity of the assessment items was appropriate in evaluating the language skills of the participants. In addition, we failed to observe a correlation between assessment scores and age, corresponding with prior research in this area (Dixon et al. 2014), further suggesting that the appropriateness of this assessment or others should be determined by language ability rather than participant age when used with an autism sample. One advantage, however, of replicating this procedure with younger participants is that verbal operants may show greater independency at younger ages when language is first emerging, but due to rapid cross-operant transfer, result in the clustering observed in the current study with this sample. Therefore, we may anticipate different results if this study were to be replicated specifically with children with or without disabilities under the age of 4 years.

A third limitation that corresponds with the second limitation is that all participants had a diagnosis of autism. We used a purely autism sample to ensure homogeneity between the participants, and because Skinner's verbal behavior theory has been most commonly applied with participants with autism (Dixon, Small, and Rosales 2007). We find it important to note, however, that Skinner's theory was never intended to be an autism language developmental theory, rather a general theory that can be applied to language development of all speaking humans. Therefore, although our results suggest that verbal development is likely interdependent for individuals with autism, research with typically developing children will be required to determine whether the verbal operants are interdependent in general. Such a finding would go a long way in either validating or invalidating the verbal operant constructs proposed in Skinner's taxonomy and assumed in the development of many behavior analytic language training technologies.

A final limitation is that the scoring system that we used assumed values of 0 or 1 for each item to remove ambiguity in the scoring of items. That is, raters only had to rate if the score was present or absent. In larger sample research, a less conservative scoring criterion that includes all possible scores for all items as described in the VB-MAPP could lead to more precise values for each item. Whether the gain in precision occurs without cost in the reliability of the obtained scores could also be further evaluated by using an item-by-item agreement analysis between two independent raters using both scoring methods. Such a strategy would also serve to evaluate the reliability of the VB-MAPP items and scoring system in addition to the internal consistency of the measure.

Corresponding with the above limitations, future research should first look to replicate and extend our procedures with a larger and more robust sample that includes within it typically developing children between the ages of 0 and 4. The VB-MAPP is conceptually structured in such a way that it is highly amenable to supporting or refuting the independency of the verbal operants, providing an opportunity for statistically interpretations of broader verbal operant development to bear upon the independency/interdependency debate. Our results are only the first in a line of research that could be conducted to empirically validate Skinner's fundamental assumptions that were used in the development of this and similar

instruments. Second, the results reported in the current study and in EFAs conducted by Rowsey et al. (2015) and Rowsey et al. (2017) all seem to suggest that verbal skills cluster in terms of skill complexity, rather than in terms of verbal operant categories. In the current study, we assume that “complexity” increases across successive skill levels of the VB-MAPP. In the EFAs conducted by Rowsey and colleagues, the conceptual structure of the PEAK modules progress from simple to more complex items, where earlier and later items in the assessments cluster together (i.e., in terms of complexity). If simple versus complex verbal operants represent independent constructs, this suggests there may be utility in isolating what exactly is meant by complexity. For example, prior research has suggested that greater training may be required for a learner to respond correctly to compound stimuli relative to stimuli differing along only a single dimension (e.g., Ribeiro, Miguel, and Goyos 2015).

Another example may be found using accounts of language development that emphasize derived responding, where greater complexity refers to the nodal distance of relational derivations in an equivalence or relational frame theory account (e.g., Arntzen and Holth 2000). Therefore, although tacts, mands, and the other “verbal operants” do not seem to represent unique constructs, responding in terms of compound stimuli, or deriving relations, may be generalized operants that are independent constructs. If true, there may be pragmatic utility in teaching verbal skills not as tacts or mands by contriving specific contextual events that surround the emission of the verbal behavior, rather by progressively increasing the complexity of verbal operant skills in terms of compound stimulus arrangements or derived relations and evaluating whether the function of these more complex skill transfers in traditional tact and mand arrangements. Each of these avenues for future research extend well beyond the ambit of the data presented in the current study, but speak to the importance of leading with data not theory in developing a valid account and approach to language development, with special utility in application with participants with diminished language skills such as those with autism and related disabilities.

In summary, because of the emerging technologies that are being made available based on Skinner’s conceptual account of verbal behavior, we need to be cautious of assuming the validity of this approach in the absence of data. Several discussions and research studies have taken place that question one aspect of the theory, namely the independence of the verbal operant categories put forth by Skinner. The current study, which extends upon prior research by conducting a statistical EFA, fails to support the independence of the verbal operant categories using the assessment items contained in the VB-MAPP. This finding, when considered in the context of prior research on this topic, suggests that the verbal operants may not be representative of independent constructs, potentially necessitating alternative accounts of operant language development that distinguish verbal behaviors in terms of complexity or other variables. This outcome has potential implications for how language assessments are developed by behavior analysts for use in educational or training settings with individuals with disabilities with diminished language skills.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

- Arntzen, E., & Holth, P. (2000). Probability of stimulus equivalence as a function of class size versus number of classes. *The Psychological Record, 50*, 79–104.
- Belisle, J., Stanley, C. R., & Dixon, M. R. (2017). The relationship between derived mutually entailed relations and the function of challenging behavior in children with autism: Comparing the PEAK-E-PA and the QABF. *Journal of Contextual Behavioral Science, 6*, 298–307. <https://doi.org/10.1016/j.jcbs.2017.07.004>
- Carr, J. E., & Miguel, C. F. (2013). The analysis of verbal behavior and its therapeutic applications. In G. J. Madden (Ed.), *APA handbook of behavior analysis Translating principles into practice*. Worcester, MA: American Psychological Association.
- Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research, and evaluation, 10*, 7–16. <https://doi.org/10.7275/jyj1-4868>
- Davis, B. J., Kahng, S., & Coryat, K. (2012). Manipulating motivating operations to facilitate the emergence of mands for a child with autism. *The Analysis of verbal behavior, 28*, 145–150. <https://doi.org/10.1007/BF03393116>
- DeSouza, A. A., Akers, J. S., & Fisher, W. W. (2017). Empirical application of skinner's verbal behavior to interventions for children with autism: A review. *The Analysis of Verbal Behavior, 33*, 229–259. <https://doi.org/10.1007/s40616-017-0093-7>
- Dixon, M. R. (2014a). *PEAK relational training system: Direct training*. Carbondale, IL: Shawnee Scientific Press.
- Dixon, M. R. (2014b). *PEAK relational training system: Generalization*. Carbondale, IL: Shawnee Scientific Press.
- Dixon, M. R., Belisle, J., McKeel, A., Whiting, S., Speelman, R., Daar, J. H., & Rowsey, K. (2017). An Internal and critical review of the PEAK relational training system for children with autism and related intellectual disabilities: 2014–2017. *The Behavior Analyst, 40*, 493–521. <https://doi.org/10.1007/s40614-017-0119-4>
- Dixon, M. R., Belisle, J., Stanley, C., Rowsey, K., Daar, J. H., & Szekeley, S. (2015). Toward a behavior analysis of complex language for children with autism: Evaluating the relationship between PEAK and the VB-MAPP. *Journal of Developmental and Physical Disabilities, 27*, 223–233. <https://doi.org/10.1007/s10882-014-9410-4>
- Dixon, M. R., Small, S. L., & Rosales, R. (2007). Extended analysis of empirical citations with skinner's Verbal Behavior: 1984–2004. *The Behavior Analyst, 30*, 197–209. <https://doi.org/10.1007/BF03392155>
- Dixon, M. R., Whiting, S. W., Rowsey, K., & Belisly, J. (2014). Assessing the relationship between intelligence and the PEAK relational training system. *Research in Autism Spectrum Disorders, 8*(9), 1208–1213. <https://doi.org/10.1016/j.rasd.2014.05.005>
- Dymond, S., O'Hora, D., Whelan, R., & O'Donovan, A. (2006). Citation analysis of skinner's verbal behavior: 1984–2004. *The Behavior Analyst, 29*, 75–88. <https://doi.org/10.1007/BF03392118>
- Endicott, K., & Higbee, T. S. (2007). Contriving motivating operations to evoke mands for information in preschoolers with autism. *Research in Autism Spectrum Disorders, 1*, 210–217. <https://doi.org/10.1016/j.rasd.2006.10.003>
- Fryling, M. J. (2017). The functional independence of skinner's verbal operants: Conceptual and applied implications. *Behavioral Interventions, 32*, 70–78. <https://doi.org/10.1002/bin.1462>

- Gamba, J., Goyos, C., & Petursdottir, A. I. (2015). The functional independence of mands and tacts: Has it been demonstrated empirically? *The Analysis of Verbal Behavior*, *31*, 10–38. <https://doi.org/10.1007/s40616014-0026-7>
- Geiger, K. B., LeBlanc, L. A., Dillon, C. M., & Bates, S. L. (2010). An evaluation of preference for video and in vivo modeling. *Journal of Applied Behavior Analysis*, *43*, 279–283. <https://doi.org/10.1901/jaba.2010.43-279>
- Grannan, L., & Rehfeldt, R. A. (2012). Emergent intraverbal responses via tact and match-to-sample instruction. *Journal of Applied Behavior Analysis*, *45*, 601–605. <https://doi.org/10.1901/jaba.2012.45-601>
- Grow, L. L., & Kodak, T. (2010). Recent research on emergent verbal behavior: Clinical applications and future directions. *Journal of Applied Behavior Analysis*, *43*, 775–778. <https://doi.org/10.1901/jaba.2010.43-775>
- Hayes, S. C. (1996). Developing a theory of derived stimulus relations. *Journal of the Experimental Analysis of Behavior*, *65*, 309–311. <https://doi.org/10.1901/jeab.1996.65-309>
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York, NY: Springer Science & Business Media.
- Howlett, M. A., Sidener, T. M., Progar, P. R., & Sidener, D. W. (2011). Manipulation of motivating operations and use of a script-fading procedure to teach mands for location to children with language delays. *Journal of Applied Behavior Analysis*, *44*, 943–947. <https://doi.org/10.1901/jaba.2011.44-943>
- Kooistra, E. T., Buchmeier, A. L., & Klatt, K. P. (2012). The effect of motivating operations on the transfer from tacts to mands for children diagnosed with autism. *Research in Autism Spectrum Disorders*, *6*, 109–114. <https://doi.org/10.1016/j.rasd.2011.03.010>
- Lipkens, R., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology*, *56*, 201–239. <https://doi.org/10.1006/jecp.1993.1032>
- Malkin, A., Dixon, M. R., Speelman, R. C., & Luke, N. (2017). Evaluating the relationships between the peak relational training system-direct training module, assessment of basic language and learning skills-revised, and the vineland adaptive behavior scales–II. *Journal of Developmental and Physical Disabilities*, *29*, 341–351. <https://doi.org/10.1007/s10882-016-9527-8>
- Matson, J. L., & Vollmer, T. R. (1995). *User's guide: Questions about behavior function (QABF)*. Baton Rouge, LA: Scientific Publishers.
- May, R. J., Hawkins, E., & Dymond, S. (2013). Brief report: Effects of tact training on emergent intraverbal vocal responses in adolescents with autism. *Journal of Autism and Developmental Disorders*, *43*, 996–1004. <https://doi.org/10.1007/s10803-012-1632-7>
- McKeel, A. N., Rowsey, K. E., Belisle, J., Dixon, M. R., & Szekely, S. (2015). Teaching complex verbal operants with the PEAK relational training system. *Behavior Analysis in Practice*, *8*, 241–244. <https://doi.org/10.1007/s40617-015-0067-y>
- Nicholson, J., Konstantinidi, E., & Furniss, F. (2006). On some psychometric properties of the questions about behavior function (QABF). *Research in Developmental Disabilities*, *3*, 337–352. <https://doi.org/10.1016/j.ridd.2005.04.001>
- Paclawskyj, T. R., Matson, J. L., Rush, K. S., Smalls, Y., & Vollmer, T. R. (2000). Questions about behavioral function (QABF): A behavioral checklist for functional assessment of aberrant behavior. *Research in Developmental Disabilities*, *21*, 223–229. [https://doi.org/10.1016/S0891-4222\(00\)00036-6](https://doi.org/10.1016/S0891-4222(00)00036-6)
- Partington, J. W. (2008). *The assessment of basic language and learning skills-revised (the ABLLS-R)*. Pleasant Hill, CA: Behavior Analysts Inc.
- Partington, J. W., & Bailey, J. S. (1993). Teaching intraverbal behavior to preschool children. *The Analysis of Verbal Behavior*, *11*, 9–18. <https://doi.org/10.1007/BF03392883>
- Ribeiro, D. M., Miguel, C. F., & Goyos, C. (2015). The effects of listener training on discriminative control by elements of compound stimuli in children with disabilities. *Journal of the Experimental Analysis of Behavior*, *104*, 48–62. <https://doi.org/10.1002/jeab.161>
- Rowsey, K. E., Belisle, J., & Dixon, M. R. (2015). Principal component analysis of the PEAK relational training system. *Journal of Developmental and Physical Disabilities*, *27*, 15–23. <https://doi.org/10.1007/s10882-014-9398-9>
- Rowsey, K. E., Belisle, J., Stanley, C. R., Daar, J. H., & Dixon, M. R. (2017). Principal component analysis of the PEAK Generalization Module. *Journal of Developmental and Physical Disabilities*, *29*, 489–501. <https://doi.org/10.1007/s10882-017-9539-z>

- Shillingsburg, M. A., Kelley, M. E., Roane, H. S., Kisamore, A., & Brown, M. R. (2009). Evaluation and training of yes-no responding across verbal operants. *Journal of Applied Behavior Analysis, 42*, 209–223. <https://doi.org/10.1901/jaba.2009.42-209>
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech, Language, and Hearing Research, 14*, 5–13. <https://doi.org/10.1044/jshr.1401.05>
- Sidman, M., & Tailby, W. (1982). Conditional discrimination versus matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*, 5–22. <https://doi.org/10.1901/jeab.1982.37-5>
- Skinner, B. F. (1953). *Science and human behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Skinner, B. F. (1957). *Verbal behavior*. New York, NY: The Free Press.
- Sundberg, M. L. (2008). *VB-MAPP: Verbal behavior milestones assessment and placement program*. Concord: AVB Press.
- Sundberg, M. L., & Michael, J. (2001). The benefits of skinner's analysis of verbal behavior for children with autism. *Behavior Modification, 25*, 698–724.
- Wallace, M. D., Iwata, B. A., & Hanley, G. P. (2006). Establishment of mands following tact training as a function of reinforcer strength. *Journal of Applied Behavior Analysis, 39*, 17–24. <https://doi.org/10.1901/jaba.2006.119-04>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.