Received: 27 April 2023 | Revised: 17 August 2023 | Accepted: 22 August 2023

DOI: 10.1002/bin.1978

RESEARCH ARTICLE

WILEY

Creating a reference range of common problem behaviors and replacement behaviors in neurotypical children

Jessica VanDevander[1] | **Allison Warner**[1] | **Ellie Kazemi**[2] | **Tara Fahmie**[3]

[1]Achieving True Self, Inc., Irwin, Pennsylvania, USA

[2]California State University, Northridge, Northridge, California, USA

[3]University of Nebraska Medical Center, Omaha, Nebraska, USA

Correspondence
Jessica VanDevander, Achieving True Self, Inc., Irwin, PA, USA.
Email: jvandevander@achievingtrueself.com

Abstract

There are limited resources outlining age-referenced rates of problem behavior displayed by neurotypical children. Such information is important for practitioners as a basis for social comparison when they are targeting behavior reduction goals for neurodivergent clients. We distributed a survey to parents of children aged 1–10 years without a developmental diagnosis in which parents reported frequency of five problem behaviors across a 24-h period, as well as commonly targeted replacement behaviors. Problem behavior was reported across all age groups to varying degrees based on topography and age. Replacement behaviors, such as waiting and tolerating denials generally improved as children increased in age but still largely remained below 80% of opportunities. The present study may serve as a reference for researchers and clinicians to set goals that are developmentally appropriate.

KEYWORDS
behavior analysis, neurotypical, problem behavior, rates of behavior, replacement behavior

1 | INTRODUCTION

Behavior analysts have several norm- and criterion-referenced assessments and resources available to assist with programming skill acquisition targets in neurodivergent children. When working with early learners, assessments such as the Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008), Assessment of Basic Language and Learning Skills—Revised (Partington, 2010), and Promoting the Emergence of Advanced

Knowledge (Dixon et al., 2014) are commonly used to identify areas of need and objectives to be targeted for acquisition. For example, the VB-MAPP evaluates 170 language and learning milestones demonstrated by children aged 0–4 years across three levels: 0–18 months, 18–30 months, and 30–48 months. The VB-MAPP provides age ranges for specific skills that are aligned with developmental milestones. In addition, it examines common learning and language acquisition barriers and the skills needed to transition to a least restrictive environment. Padilla (2020) surveyed 1500 behavior analyst professionals on their use of various assessments and found that an overwhelming 76% of them utilized the VB-MAPP, while 45% reported using ABLLS-R and 14% reported using PEAK.

Many of these assessments also include information related to problem behavior. For example, the barriers assessment of the VB-MAPP includes the following areas: problem behaviors, instructional control, self-stimulation, obsessive–compulsive behavior, hyperactive behavior. Similarly, the transitions assessment covers the following areas: negative behaviors, adaptability to change, and spontaneity. Additional assessments such as the Vineland-3 (Sparrow et al., 2016), Child Behavior Checklist (Achenbach, 1999), Social Skills Improvement System Rating Scales (Gresham & Elliott, 2008), and the Behavior Assessment System for Children (Merenda, 1996) also include a behavior section with norm-referenced data on problem behavior (e.g., "is very irritable or moody," "is physically aggressive," and "acts without thinking").

Despite the widespread use of assessments such as the VB-MAPP and Vineland-3, there remains a gap in resources available to identify specific amounts of behavior commonly observed across age ranges. The latter may be helpful for setting acquisition or reduction goals in interventions. That is, current assessments identify how often behaviors are observed using summary ratings such as "always," "sometimes," or "never", which provide information on general level of behavior. However, these assessments do not provide clear points of comparison with the behavioral measures (typically a rate, duration, or percentage of time samples) that are typically collected in clinical practice. Assessments that quantify and compare rates or durations of behavior across age ranges (e.g., two tantrums per day may be expected from a 2-year-old child but that same frequency of tantrums would be inappropriate for a 9-year-old) are relatively scarce.

MacDonald et al. (2007) conducted direct observations of stereotypic behavior in typically developing and autistic children between 2 and 4 years of age. Results of this study showed that autistic children engaged in somewhat higher levels of stereotypic behavior by the age of 2, but the gap widened at ages 3 and 4. Similarly, Bush et al. (2022) and Nuhu and Rapp (2020) examined eye contact in neurotypical individuals, which is a commonly targeted prosocial behavior in behavior analytic programming. Relatively more research on problem behavior prevalence, trajectories, and environmental correlates has been conducted using survey measures (e.g., Campbell et al., 2006; Harrison et al., 2012; Koot & Verhulst, 1991; McDermott & Schaefer, 1996; NICHD, 2004). For example, McMahon et al. (2022) provided surveys to pre-kindergarten through 12th-grade teachers. Respondents reported high levels of physical aggression in elementary school, high levels of verbal aggression and property aggression in middle school, and high levels of verbal aggression in high school.

Potegal et al. (2003) used parent narratives from 335 children aged 18 months–5 years of age with no documented diagnosis and found that 75% of tantrums lasted 5 min or less. Sisterhen and Wy (2022) evaluated the average rate and duration of temper tantrums in the same age range for children with diagnoses. Information on quantities (e.g., rates and durations) of problem behavior, across a wider range of topographies (e.g., self-injury and aggression), and across a wider range of ages would be useful.

Research is also needed to establish normal rates of prosocial behaviors (e.g., communication, compliance, delay, and tolerance) that are commonly used to replace problem behavior in behavioral intervention programs. Assessments of replacement behaviors (e.g., Cipani, 2008) and reviews of replacement behaviors (e.g., Matson et al., 2011) have not provided information on level of replacement behaviors displayed by neurotypical children. Thus, mastery criteria for these skills are often set by behavior analysts without reference to ranges based on the client's age. Richling et al. (2019) conducted a survey of practitioners and found that the most commonly used mastery criterion was 80% across 3 consecutive sessions. Skills are often targeted with this mastery criterion without an acknowledgment of whether that criterion is age appropriate. For example, a 2-year-old and 5-year-old can both be working on

tolerating "no" with the same mastery criterion (80% across 3 consecutive sessions) despite the appropriateness of this skill varying between ages.

Richling et al. (2019) and Fuller and Fienup (2018) conducted studies on different mastery criteria and their effects on maintenance and found that mastery criteria higher than 80% were necessary for maintenance. Citing a lack of empirical support for setting specific mastery criteria, Rapp et al. (2019) tried reducing mastery criteria for a child that was unable to maintain at 80%, which resulted in the client demonstrating maintenance at the modified criteria. These results suggest that the most used mastery criteria may be insufficient in producing maintenance in some cases and may need to be individualized. However, even the definition of maintenance should be informed by the percentage of opportunities with which typically developing children are engaging in those behaviors.

To ensure that behavior analysts are not holding their neurodivergent clients to an unrealistic standard, it is important to consider developmental norms related to problem and replacement behaviors. A resource for determining age-appropriate ranges of problem behavior would be useful for behavior analysts. These data could support more appropriate programming and support behavior analysts who might not have as much experience with typically developing children and age-appropriate levels of problem behavior. Additionally, the data could provide all practitioners with a reference range for setting evidence-based objectives related to decreasing problem behavior and increasing replacement behavior instead of choosing arbitrary targets. Thus, the purpose of this study was to survey caregivers of neurotypical children between the ages of 1 and 10, to gather reports of the amount of problem and adaptive replacement behaviors present at each age range.

2 | METHOD

2.1 | Participants

The study was conducted using an anonymous survey provided to parents/guardians of neurotypical children. The survey was distributed via an online link that was posted on the following social media pages: Instagram via a parent blog page, Facebook parenting groups (both local and general groups), and Facebook behavior analysis groups. Each post included a request for participation and distribution. Participants were the caregivers of children between the ages of 1 and 10 years old. We designated the first age range as 12–18 months and the second as 18 months–2 years. Additional age ranges consisted of children in each sequential year of life (2 years old, 3 years old, 4 years old, etc.). In total, 1746 responses were collected. Of those, 142 respondents indicated that their child had a developmental diagnosis (autism, ADHD, etc.) and, therefore, were excluded from the dataset, leaving 1604 total responses included in this study. Responses were recorded from a total of 37 countries and 40 different states within the United States. Responses per age group were not equally distributed with 83% of responses being children aged 1–4 years. Notably, 81% of participants identified their ethnicity as white (see Table 1).

2.2 | Materials

The survey was provided online and conducted anonymously by providing a link. A description of the purpose of the survey was provided along with a disclaimer that the survey should only be filled out by parents of children without a developmental diagnosis. The survey included 15 questions: age of child, ethnicity of child, confirmation of a lack of developmental disability/diagnosis, location (country and state if within US), if siblings are in the home, five questions regarding the frequency of problem behavior within a 24-h period, one question regarding the duration of tantrums, and four questions regarding adaptive skills frequently targeted as replacement behaviors.

Caregivers were asked about five problem behaviors: tantrums (frequency and duration), aggression (excluding biting), biting, throwing, and property destruction. An operational definition was provided for each behavior, with

TABLE 1 Responses per age group and ethnicity.

Age range	Total responses	Percent of responses
12–18 months	405	25%
18 months–2 years	221	14%
2 years	214	13%
3 years	200	12%
4 years	170	11%
5 years	121	8%
6 years	86	5%
7 years	65	4%
8 years	52	3%
9 years	31	2%
10 years	39	2%
Ethnicity	**Total responses**	**Percent of responses**
Asian	16	1%
Black	17	1%
White	1297	81%
Native Hawaiian or other Pacific Islander	2	0%
Hispanic	79	5%
American Indian or Alaska Native	4	0%
Mixed race	170	11%
Prefer not to answer	19	1%

examples and non-examples when relevant. Parents were asked how often their child engaged in each individual behavior within a 24-h period and were provided with the options of 0 times, 1 time, 2 times, 3 times, 4 times, 5 times, 6 times, 7 times, 8 times, 9 times, and 10+ times. For the purpose of averaging rates of behavior, scores of 10+ were scored as 10 instances per day.

Four replacement skills were assessed: relinquishing an item the first time asked, transitioning to a non-preferred task when asked for the first time, being denied access to preferred item in the absence of problem behavior, and waiting to access an item/activity in the absence of problem behavior. Questions were worded by asking the parent/guardian to describe the number of times their child engaged in a skill out of the last 10 times the child was presented with an opportunity. While acquisition targets within behavior analytic programming are often measured by trial, percent of opportunities was used as a metric for the current survey because this represented a more naturalistic estimate of behavior occurring in everyday routines that caregivers oversee. For instance, "Out of the last 10 times you asked your child to give up a preferred item they had, how many times did they give it up the first time they were asked?" That number was then converted to a percentage. For example, if a parent/guardian reported that their child was able to wait 3 times out of 10, data was recorded as 30%.

3 | RESULTS

Figure 1 displays the average rate of problem behavior per day reported across designated age ranges. Reports ranged from 0 instances per day to 2.92 instances per day. Tantrums were reported at the highest rate for all age ranges except 10-year-olds, whereas biting was reported as the lowest rate across all age ranges. Biting was the only behavior that reportedly did not occur in particular age ranges, and it occurred most frequently in the age ranges in which children may be teething (i.e., 12 months–2 years old).

Overall, with the exception of aggression, all problem behaviors were reportedly highest between 12 months and 2 years, thereafter showing a consistent downward trend. In addition, all behaviors except biting were reported

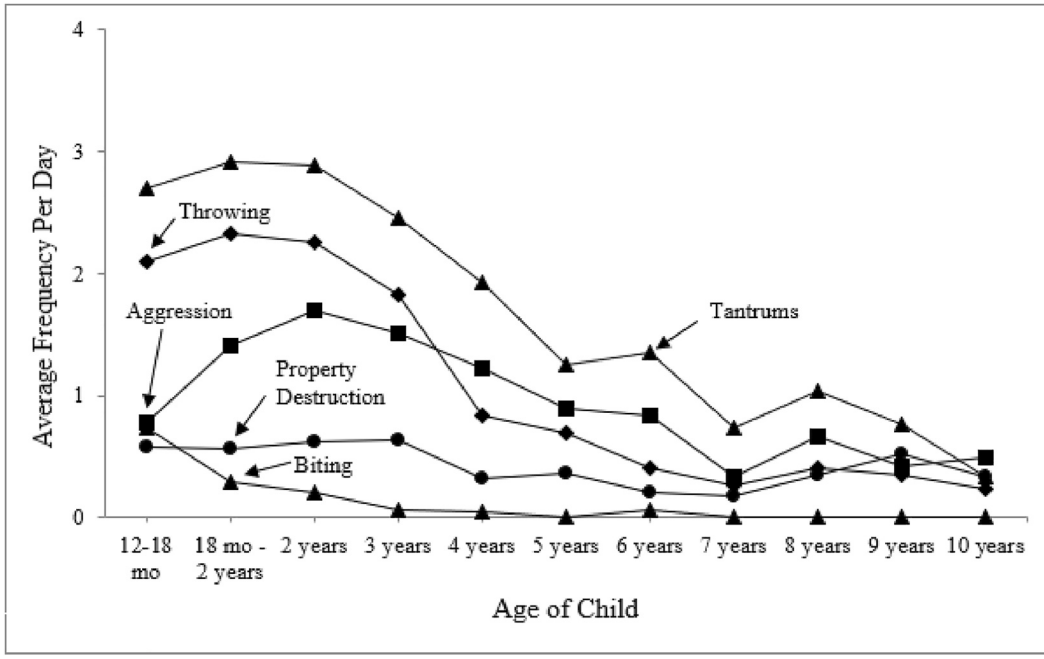

FIGURE 1 Average occurrence of problem behavior per day by age.

Tantrum Frequency	12-18 months	18 mo- 2 years	2 years	3 years	4 years	5 years	6 years	7 years	8 years	9 years	10 years
0	11%	7%	5%	10%	18%	34%	41%	54%	42%	58%	79%
1	22%	20%	21%	30%	28%	33%	23%	31%	31%	16%	15%
2	25%	23%	25%	21%	25%	21%	22%	12%	17%	23%	3%
3	16%	15%	21%	21%	15%	6%	3%	0%	4%	0%	0%
4	9%	17%	7%	5%	4%	3%	5%	0%	2%	0%	0%
5	7%	7%	11%	7%	6%	2%	2%	2%	4%	3%	3%
6	3%	3%	6%	1%	2%	0%	1%	0%	0%	0%	0%
7	1%	3%	0.9%	1.5%	1.2%	0%	1%	2%	0%	0%	0%
8	2%	3%	1%	2%	1%	1%	0%	0%	0%	0%	0%
9	0.5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
10+	3%	2%	2%	3%	0%	1%	1%	0%	0%	0%	0%

Tantrum Duration	12-18 months	18 mo-2 years	2 years	3 years	4 years	5 years	6 years	7 years	8 years	9 years	10 years
N/A	9%	7%	4%	7%	13%	31%	35%	46%	37%	55%	74%
30s	46%	33%	15%	14%	8%	4%	16%	11%	6%	6%	3%
1 min	25%	31%	27%	18%	14%	18%	16%	9%	19%	3%	3%
2-3 min	15%	22%	39%	38%	41%	26%	13%	14%	17%	16%	5%
4-5 min	2%	6%	13%	13%	18%	14%	17%	11%	12%	10%	5%
6-10 min	1%	1%	0.5%	8%	5%	4%	1%	3%	6%	6%	3%
10-15 min	0%	1%	1%	3%	1%	2%	1%	0%	4%	0%	3%
>15 min	0%	0%	0%	0.5%	1%	0%	0%	6%	0%	3%	5%

Aggression Frequency	12-18 months	18 mo-2 years	2 years	3 years	4 years	5 years	6 years	7 years	8 years	9 years	10 years
0	64%	43%	33%	38%	48%	55%	69%	77%	62%	71%	67%
1	15%	27%	26%	31%	24%	26%	13%	17%	19%	19%	26%
2	11%	12%	20%	11%	12%	12%	8%	3%	12%	6%	5%
3	3%	7%	7%	9%	6%	3%	0%	2%	6%	3%	0%
4	3%	4%	5%	0.5%	4%	2%	7%	2%	2%	0%	0%
5	2%	3%	4%	8%	4%	2%	0%	0%	0%	0%	3%
6	0.2%	0.9%	3%	0%	0.6%	0%	1%	0%	0%	0%	0%
7	0%	0%	0.5%	0.5%	0.6%	0%	0%	0%	0%	0%	0%
8	0.2%	1%	2%	1%	1%	0%	2%	0%	0%	0%	0%
9	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
10+	0.2%	3%	1%	2%	0.6%	2%	0%	0%	0%	0%	0%

FIGURE 2 Percentage of responses per age group by frequency (tantrum and aggression) and duration (tantrum). The gradient shading of the cell gets progressively darker, with the darker shading of the cell indicating that a higher proportion of respondents endorsed that category.

to increase slightly in rate at age 8. Property destruction and biting had the least amount of variability between their highest and lowest reported rates, 0.46 (ages 3–7) and 0.74 (ages 12 months–5 and 7–10 years) respectively. Tantrums had the greatest difference in their highest and lowest reported rates, with a difference of 2.59 instances per day between ages 18 months–2 years and 10 years. Apart from property destruction, all behaviors had the most significant drops in reported rates between 1 and 5 years old, with still decreasing but more stable trends from 5 to 10 years old.

Figures 2 and 3 display the same survey data across rating options in each category to provide a clearer picture of response distributions across the sample. Additionally, reported average duration of tantrums is included. In this figure, the darker shading of the cell indicates that a higher proportion of respondents endorsed that category. Across all problem behaviors, the proportion of responses was higher in 0–2 instances per-day category than in any other category. Fewer respondents endorsed a rate of 4 to 10+ instances per day. Biting shows the least variability in

Biting Frequency	12-18 months	18 mo-2 years	2 years	3 years	4 years	5 years	6 years	7 years	8 years	9 years	10 years
0	71%	83%	85%	95%	96%	100%	98%	100%	100%	100%	100%
1	13%	10%	11%	5%	3%	0%	0%	0%	0%	0%	0%
2	6%	5%	4%	0.5%	0%	0%	1%	0%	0%	0%	0%
3	5%	1%	0.5%	0%	0%	0%	1%	0%	0%	0%	0%
4	2%	0%	0%	0%	0.6%	0%	0%	0%	0%	0%	0%
5	2%	0.9%	0%	0%	0%	0%	0%	0%	0%	0%	0%
6	0%	0.5%	0%	0%	0%	0%	0%	0%	0%	0%	0%
7	0.2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
8	0.2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
9	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
10+	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Throwing Frequency	12-18 months	18 mo-2 years	2 years	3 years	4 years	5 years	6 years	7 years	8 years	9 years	10 years
0	43%	28%	21%	29%	54%	63%	73%	85%	75%	77%	82%
1	13%	21%	26%	31%	26%	18%	16%	11%	13%	13%	13%
2	12%	19%	24%	19%	12%	12%	7%	2%	10%	6%	5%
3	8%	10%	10%	8%	4%	3%	3%	2%	0%	3%	0%
4	5%	6%	5%	3%	2%	2%	0%	0%	2%	0%	0%
5	7%	4%	7%	5%	1%	3%	0%	2%	0%	0%	0%
6	3%	4%	2%	1%	0%	0%	0%	0%	0%	0%	0%
7	0.7%	0%	0.9%	1%	0%	0%	0%	0%	0%	0%	0%
8	0.7%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%
9	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%
10+	6%	6%	3%	5%	0.6%	0%	0%	0%	0%	0%	0%

Property Destruction Frequency	12-18 months	18 mo-2 years	2 years	3 years	4 years	5 years	6 years	7 years	8 years	9 years	10 years
0	73%	74%	70%	66%	81%	79%	86%	88%	79%	71%	79%
1	14%	14%	17%	21%	11%	13%	10%	11%	10%	16%	18%
2	7%	5%	7%	7%	6%	2%	1%	0%	10%	10%	0%
3	2%	4%	2%	5%	0.6%	6%	1%	0%	2%	0%	0%
4	1%	0%	0.5%	2%	0%	0%	1%	0%	0%	0%	0%
5	1%	1%	0.5%	1%	1%	0%	0%	2%	0%	3%	0%
6	0.2%	1.4%	1%	0%	0.6%	0%	0%	0%	0%	0%	3%
7	0.5%	0.5%	0.5%	0.5%	0%	0%	0%	0%	0%	0%	0%
8	0%	0%	0.5%	0%	0%	0%	0%	0%	0%	0%	0%
9	0%	0%	0.5%	0%	0%	0%	0%	0%	0%	0%	0%
10+	0.5%	0%	0.5%	0.5%	0%	0%	0%	0%	0%	0%	0%

FIGURE 3 Percentage of responses per age group by frequency (biting, throwing, and property destruction). The gradient shading of the cell gets progressively darker, with the darker shading of the cell indicating that a higher proportion of respondents endorsed that category.

distribution and tantrums show the highest. Average tantrum duration was reported mostly between 30s and 3 min until age 6 when the highest distribution of responses decreased to 0.

Figure 4 displays the average percentage that each age range was reported to demonstrate skills that are commonly targeted as replacement behaviors. Waiting was the highest reported skill across all ages except those under 2 years old. It is also the only skill that had any reports at or above 80%, with 7-year-olds reportedly demonstrating it on 81.8% and 10-year-olds on 80% of the opportunities. The lowest reported skill was engaging in a non-preferred activity the first time asked, with 2-year-olds reportedly demonstrating it 41.9% of the time (on the low range) and 10-year-olds reportedly demonstrating it 60.3% of the time (on the high range). All replacement skills,

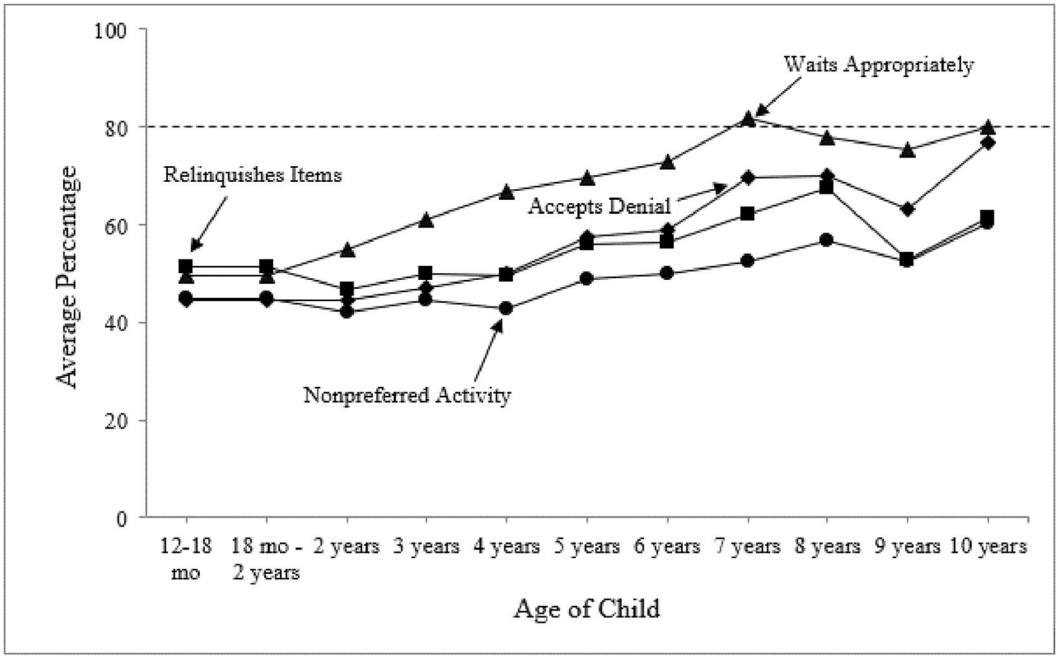

FIGURE 4 Average percentage of occurrence per opportunity of replacement behavior by age.

with the exception of waiting, had a stable trend with little improvement until age 4. All skills showed slight increases in percentage between 4 and 7 years old; some skills showed a decrease between ages 8 and 9 years old, with a consistent increase at 10 years old.

4 | DISCUSSION

The present study identified rates of problem and replacement behaviors reported by caregivers of neurotypical children aged 1–10 years old. Current assessments and resources available such as the VB-MAPP (Sundberg, 2008) and the Vineland-3 (Sparrow et al., 2016) provide generalized age-referenced guidance for practitioners when programming for skill acquisition and reduction of problem behaviors. However, these assessments do not provide quantifiable rates per age that can be used while developing behavior acquisition or reduction goals. Our study provides data, both averages and distribution ranges, which may be used while setting goals in clinical programs that are better aligned with developmental norms. For example, if a practitioner has a student who is 3 years old and engages in five tantrums per day, they can reference these data to determine that, on average, the 3-year-olds in this study engaged in 2.46 tantrums per day. In this case, the practitioner should not make a reduction goal below 2 tantrums per day. Practitioners would also see that 7% of parents surveyed reported that their 3-year-old child engages in 5 tantrums per day and that 1– to 3 tantrums per day were common.

The data provided in this study are a starting point for further research to determine levels of behavior demonstrated by neurotypical children. However, several limitations must be acknowledged. Foremost, the extent to which the survey can be generalized to the population at large is unknown, given uneven sampling across age ranges and ethnicities. The number of parents with children aged 12–18 months who responded to the survey was almost twice as large as the next age range with the highest responses, aged 18 months–2 years. We collected markedly fewer responses from parents of children aged 6–10. Furthermore, there was a significant lack of representation of ethnicities outside of white respondents. Responses were collected from 37 different countries, which likely varied substantially on cultural, financial, and social variables and influenced the data to an unknown degree. Attempts were

made to improve the representation of ages and ethnicities. The survey was distributed multiple times and this was successful in increasing the number of participants in the age ranges of 5–10; however, a significant skew remained with respect to both age and ethnicity. Additionally, the researchers reached out to organizations specific to behavior analysis in different cultures in an attempt to further distribute the survey to underrepresented populations, although this did not produce an outcome that resulted in reducing the gap between white participants and other ethnicities. Future research should ensure more diverse and representative sampling methods.

Next, the data were collected via an anonymous survey completed by caregivers and were likely educated guesses. Further analysis of objective measures of behavior collected via direct observation would provide more accurate data that can be relied on for programming. Additionally, behavior analytic goals are often written using a percent of trials measure, which differed from the frequency and percent of opportunities measures requested by respondents to our survey. Although the measures we requested were intended to reflect natural everyday interactions with caregivers, this feature of our survey limits the generality of our results to goal setting in discrete-trial instructional situations.

Another limitation was the way we summarized reports of 10+ occurrences per day. These reports were considered as only 10 occurrences per day to allow for mean and range calculations. Although a small percentage of our sample (6.7%) reported rates of 10+ responses per day, we caution that this calculation method biased our data toward lower averages than would have otherwise been obtained without an upper cap. Future studies should provide additional options for frequencies of behavior above 10 and should further investigate the intensities of reported problem behavior. For example, a 2-year-old hitting someone's arm frequently with level 2 intensity on a 10-point Likert scale would differ greatly from a 9-year-old hitting someone in the face infrequently with level 8 intensity on the same scale.

The final limitation of the survey being completed by caregivers was a lack of information relevant to distinct settings in which problem and replacement behaviors occurred. Caregivers completing the survey may have been reporting on the frequency of behaviors in the home setting, which may differ from other settings of interest to behavior analysts (e.g., school, community, and clinic). Future research should determine whether our results are setting-specific. A final limitation is that we sampled only a range of problem and replacement behaviors and did not gather information on some skills commonly targeted by behavior analysts (e.g., functional communication skills).

Problem behaviors were observed across all age ranges, with the highest levels being reported in ages 1–2 and remaining high until age 4, when a more significant downward trend was noted. Conversely, the replacement skills showed stable lower rates until age 4 and then more notable increasing trends. Further research would be beneficial in determining a potential correlation between the ages at which adaptive behaviors are acquired and their impact on the emergence or prevention of problem behavior. Additionally, further research should be conducted to determine the function of biting across all age ranges to determine potential medical causes during teething periods.

There is an abundance of available research on teaching prosocial skills to decrease problem behavior (e.g., McKenna et al., 2017); however, we know relatively little about the preventive effects of acquiring specific prosocial skills (Fahmie et al., 2020). The lack of empirical evidence on how consistently children demonstrate various replacement skills limits practitioners when developing treatment goals and selecting mastery criteria that are attainable and appropriate for their client. Thus, future research to guide practitioners in everyday programming is needed.

These data can help practitioners avoid holding neurodivergent clients to a standard that does not reflect neurotypical development. However, this does not mean that we should be setting targets for neurodivergent individuals solely based on the rates of behavior demonstrated by neurotypical peers. Researchers have acknowledged that social norms and comparison standards should be secondary to data on the preferences of neurodivergent individuals and the impact of those standards on the success of individuals from that population (Hood et al., 2021). Thus, research on the validity and potential impacts of using data from neurotypical individuals to formulate goals for the neurodivergent population is needed.

Interestingly, reports of replacement behavior reached 80% or higher on only two occasions (waiting appropriately at 7 and 10 years old). Richling et al. (2019) indicated that 80% across 3 consecutive days is the most used mastery criteria in behavior analytic programming. Richling et al. (2019) and Fuller and Fienup (2018) both presented data suggesting that mastery criteria of 90% should be used to facilitate higher accuracy of responding during

maintenance. However, the data presented within this study indicate that higher requirements for mastery and maintenance might lead practitioners to target developmentally inappropriate and unattainable criteria for skills depending on the client's age. In fact, over half of the data points across all age ranges and skills were below 60%, once again bearing the question on whether we hold our neurodivergent clients to a higher standard than their neurotypical peers. Additionally, most studies on mastery criteria have evaluated criteria related to rote and discretely taught skills as opposed to prosocial skills demonstrated in the natural environment.

More data need to be collected on a wider variety of adaptive skills and their occurrence across settings within each age group. When targeting skills like those reported in this study, a child's age should be taken into consideration when determining mastery criteria, and that criteria should be individualized for each child and skill. Further research is needed to determine when these skills are developed and how consistently they should be demonstrated across each age range to promote the child's autonomy, happiness, and success. Additionally, research into the social validity of mastery criteria and targeted rates of behavior would allow practitioners to assess caretakers' satisfaction with programming and whether they feel their child's goals are a pathway to autonomy, happiness, and success.

In summary, the present study may be a starting point for research on socially significant programming for the neurodivergent population. Despite vast research in the areas of problem behavior and replacement behavior, there remains a significant research gap with respect to measurable dimensions of behavior across different age ranges. Despite its limitations, this study can serve as a springboard to future research and a resource for practicing behavior analysts. Ultimately, the continuation of this line of research is needed to ensure that our interventions are attainable, developmentally appropriate, ethical, and socially significant.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Jessica VanDevander https://orcid.org/0009-0003-1296-9871

REFERENCES

Achenbach, T. M. (1999). The child behavior checklist and related instruments. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (pp. 429–466). Lawrence Erlbaum Associates Publishers.

Bush, H., Rapp, J. T., & Nuhu, N. (2022). A pilot study exploring social validity ratings of social profiles for college students. *Developmental Neurorehabilitation*, *25*(4), 263–273. https://doi.org/10.1080/17518423.2021.2011460

Campbell, S. B., Spieker, S., Burchinal, M., & Poe, M. D. (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry*, *47*(8), 791–800. https://doi.org/10.1111/j.1469-7610.2006.01636.x

Cipani, E. (2008). Analogue assessment of the replacement behavior. *International Journal of Behavioral Consultation and Therapy*, *4*(4), 374–379. https://doi.org/10.1037/h0100868

Dixon, M. R., Carman, J., Tyler, P. A., Whiting, S. W., Enoch, M. R., & Daar, J. H. (2014). PEAK relational training system for children with autism and developmental disabilities: Correlations with peabody picture vocabulary test and assessment reliability. *Journal of Developmental and Physical Disabilities*, *26*(5), 603–614. https://doi.org/10.1007/s10882-014-9384-2

Fahmie, T. A., Garcia, A. R., Poetry, J. H., Tierman, E. M., Hamawe, R., Marks, S. T., & Jin, S. (2020). Topographies and functions of emerging problem behavior and appropriate requests in neurotypical preschoolers. *Journal of Applied Behavior Analysis*, *53*(4), 2199–2214. https://doi.org/10.1002/jaba.741

Fuller, J. L., & Fienup, D. M. (2018). A preliminary analysis of mastery criterion level: Effects on response maintenance. *Behavior Analysis in Practice*, *11*(1), 1–8. https://doi.org/10.1007/s40617-017-0201-0

Gresham, F. M., & Elliott, S. N. (2008). *Social skills improvement system: Rating scales*. Pearson Assessments.

Harrison, J. R., Vannest, K., Davis, J., & Reynolds, C. (2012). Common problem behaviors of children and adolescents in general education classrooms in the United States. *Journal of Emotional and Behavioral Disorders*, *20*(1), 55–64. https://doi.org/10.1177/1063426611421157

Hood, S. A., Beauchesne, B. M., Fahmie, T. A., & Go, A. (2021). Descriptive assessment of conversational skills: Towards benchmarks for young adults with social deficits. *Journal of Applied Behavior Analysis*, *54*(3), 1075–1094. https://doi.org/10.1002/jaba.831

Koot, H. M., & Verhulst, F. C. (1991). Prevalence of problem behavior in Dutch children aged 2-3. *Acta Psychiatrica Scandinavica*, *83*(s367), 1–37. https://doi.org/10.1111/j.1600-0447.1991.tb10574.x

MacDonald, R., Green, G., Mansfield, R., Geckeler, A., Gardenier, N., Anderson, J., Holcomb, W., & Sanchez, J. (2007). Stereotypy in young children with autism and typically developing children. *Research in Developmental Disabilities*, *28*(3), 266–277. https://doi.org/10.1016/j.ridd.2006.01.004

Matson, J. L., Shoemaker, M. E., Sipes, M., Horovitz, M., Worley, J. A., & Kozlowski, A. M. (2011). Replacement behaviors for identified functions of challenging behaviors. *Research in Developmental Disabilities*, *32*(2), 681–684. https://doi.org/10.1016/j.ridd.2010.11.014

McDermott, P. A., & Schaefer, B. A. (1996). A demographic survey of rare and common problem behaviors among American students. *Journal of Clinical Child Psychology*, *25*(3), 352–362. https://doi.org/10.1207/s15374424jccp2503_11

McKenna, J. W., Flower, A., Falcomata, T., & Adamson, R. M. (2017). Function-based replacement behavior interventions for students with challenging behavior. *Behavioral Interventions*, *32*(4), 379–398. https://doi.org/10.1002/bin.1484

McMahon, S. D., Cafaro, C. L., Bare, K., Zinter, K. E., Murillo, Y. G., Lynch, G., Anderman, E. M., Espelage, D. L., Reddy, L. A., & Subotnik, R. (2022). Rates and types of student aggression against teachers: A comparative analysis of U.S. elementary, middle, and high schools. *Social Psychology of Education*, *25*(4), 767–792. https://doi.org/10.1007/s11218-022-09706-6

Merenda, P. F. (1996). BASC: Behavior assessment system for children. *Measurement and Evaluation in Counseling and Development*, *28*(4), 229–232.

National Institute of Mental Health. (2004). *Child development and behavior branch*. Report to the NACHHD Council.

Nuhu, N. N., & Rapp, J. T. (2020). Descriptive analyses of behaviors during social interactions. *Developmental Neurorehabilitation*, *23*(5), 302–308. https://doi.org/10.1080/17518423.2019.1687597

Padilla, K. L. (2020). Global assessment use and practices in applied behavior analysis: Surveying the field. *Research in Autism Spectrum Disorders*, *79*, 101676. https://doi.org/10.1016/j.rasd.2020.101676

Partington, J. W. (2010). *The ABLLS-R—the assessment of basic language and learning skills-revised*. Behavior Analysts, Inc.

Potegal, M., Kosorok, M. R., & Davidson, R. J. (2003). Temper tantrums in young children: 2. Tantrum duration and temporal organization. *Journal of Developmental and Behavioral Pediatrics*, *24*(3), 148–154. https://doi.org/10.1097/00004703-200306000-00003

Rapp, J. T., Cook, J. L., Nuta, R., Balagot, C., Crouchman, K., Jenkins, C., Karim, S., & Watters-Wybrow, C. (2019). Further evaluation of a practitioner model for increasing eye contact in children with autism. *Behavior Modification*, *43*(3), 389–412. https://doi.org/10.1177/0145445518758595

Richling, S. M., Williams, W. L., & Carr, J. E. (2019). The effects of different mastery criteria on the skill maintenance of children with developmental disabilities. *Journal of Applied Behavior Analysis*, *52*(3), 701–717. https://doi.org/10.1002/jaba.580

Sisterhen, L. L., & Wy, P. A. W. (2022). *Temper tantrums*. StatPearls. StatPearls Publishing.

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland adaptive behavior scales* (3rd ed.). NCS Pearson. Vineland-3.

Sundberg, M. L. (2008). *VB-MAPP verbal behavior milestones assessment and placement program: A language and social skills assessment program for children with autism or other developmental disabilities*. AVB Press.

How to cite this article: VanDevander, J., Warner, A., Kazemi, E., & Fahmie, T. (2024). Creating a reference range of common problem behaviors and replacement behaviors in neurotypical children. *Behavioral Interventions*, *39*(1), e1978. https://doi.org/10.1002/bin.1978