Toilet Training Individuals with Developmental Delays: A Comprehensive Review

Abstract

Acquisition of toileting skills is of vital significance for children with developmental disabilities (DD) because individuals with DD are often delayed in or fail acquiring these skills. The purpose of the study is to comprehensively review the studies demonstrating the effectiveness of toilet training for children with DD between 2009 and 2019. In this literature search, four databases (ScienceDirect, EBSCOhost, Wiley Online Library and SpringerLink) were searched by using eight keywords and a total of 1360 studies were identified. The studies were evaluated in terms of eligibility criteria and 23 studies were included in the review. Authors examined the studies by the demographic, methodological and outcome characteristics. The review highlighted that toilet training programs, packages or protocols are developed according to researchers' preferences and participant characteristics, as such, there is still no agreement on a common toileting program in literature. Directions for future research are also discussed.

Keywords: Toilet Training Individuals with Developmental Delays

Introduction

Toileting refers to successful and unprompted accomplishment of recognizing the need of visiting the toilet, waiting before elimination, going to toilet, in-toilet elimination, having less or no toileting accidents and consistent dryness (Kurniawan, Purnamasari, Rakhmawati, & Jalaputra 2018). In addition to the chain of these steps, toileting contains such skills as pulling pants up/down, sitting on the toilet, appropriate use of toilet paper, flushing and washing hands (Suppo & Mayton 2012). In the population with typical development, these skills are commonly acquired between the ages of 3 and 5 (Dalrymple & Ruble 1992). However, many children with developmental disabilities (DD) such as autism spectrum disorder (ASD) and intellectual disability (ID) are often delayed in or fail acquiring these skills (Keen, Brannigan, & Cuskelly 2007; Levato et al., 2016). In a study by Szyndler (1996), 82% of parents of children with DD indicated their children were affected by toileting difficulties. In addition, Matson, Horovitz and Sipes (2011) found that more than a half of 153 participants with DD had frequent toileting problems. It is also notable that prevalence rate of encopresis is 1.6% for typically developing individuals between 10-13 (Van Der Wal, Benniga, & Hirasing 2005), while it is 11.1% for the same age group with DD (Simonoff et al., 2008). These studies suggest that children with DD are more likely to have toileting problems more so than their peers with typical development.
due to emotional factors, cognitive and communication problems (Leader, Mannion & Chen 2018).

Acquisition of toileting skills is a milestone for the population with DD (Cicero & Pfadt 2002). Problematic toileting skills place barrier for this population in socialization or vocational and residential placements (Kroeger & Sorensen-Burnworth 2009). Furthermore, toileting accidents mainly cause hygiene, stigmatism, sleep problems and derision, thus decreasing the life quality of individuals with DD (Lott & Kroeger 2004). Children having frequent toileting accidents not only show anxiety disorder and depression symptoms, but have difficulties in gathering and maintaining attention and have poorer academic performance than their normally developing peers (Cox, Morris, Borowitz & Sutphen 2002). All these difficulties and problems diminish their self-confidence, reduce access to daily life activities and social acceptance and impede with their emotional, behavioral and academic skill development (Joinson et al., 2006). Therefore, competent toileting is the main self-care skill that should be taught to the individuals with DD (Call, Mevers, McElhanon, & Scheithauer 2017; Kircaali-Iftar, Ulke-Kurkcuoglu, Cetin, & Unlu 2009).

The most comprehensive toilet training program, Rapid Toilet Training (RTT) method, was developed and published in 1971 by Azrin and Foxx in order to toilet train individuals with learning disabilities rapidly. RTT method that heavily draws on operant conditioning technique includes such components as positive reinforcement, positive punishment, hydration, and scheduled sittings. The method has demonstrated effectiveness in toilet training individuals with DD in wide variety of studies (Foxx & Azrin, 1973). Despite its success, the procedure is often broken down into its components, which are used as toilet training protocols in or/and of themselves (e.g., Cicero & Pfadt, 2002). Moreover, a lot of studies which modified and shortened RTT have been used ever since and have also shown to be effective in teaching these skills (e.g., Rinald & Mirenda, 2012).

Although literature suggests that teaching toileting skills should be one of the primary research topics, only a handful of studies has been conducted on these skills (Francis, Mannion, & Leader 2017; Mannion & Leader 2013). In fact, more toilet training interventions are reported from diverse global settings such as the USA (e.g., Kroeger & Sorensen, 2010), the Netherlands (e.g., Van Oorsouw, Duker, & Averink 2009), and Taiwan (e.g., Chang, Lee, Chou, Chen, & Chen, 201) as compared to developing countries such as Turkey, where only a limited number of studies have been published (e.g., Unlu, 2019). On the other hand, only two studies reviewed the literature on toilet training (Francis et al., 2017; Kroeger & Sorensen-Burnworth 2009). Kroeger & Sorensen-Burnworth (2009) reviewed data-based studies published in peer-reviewed journals until 2008 and included a total of 28 studies. The studies were reviewed for participant characteristics, training program components, setting, length of time and results. Francis et al. (2017) synthesized the literature on toilet training interventions and reviewed 15 data-based and peer-reviewed journal submissions between 2009 and July 2016 for the same variables with the previous literature review study. To our knowledge, there is not a recent comprehensive review in literature on the topic since then. Additionally, these two review studies mainly focused on description of toilet training programs. Therefore, a comprehensive review of the literature on toilet training for individuals with DD is further necessary to highlight demographic, methodological and outcomes characteristics of the studies included. Furthermore, this review study may assist teachers in identifying components of the procedures as they are used as toilet training programs in and of themselves. This study also serves as an agenda for future research, which guides researchers and the academia as by identifying research voids in extant studies. As a result, the purpose of this study is to conduct a comprehensive review of the literature on toilet training to the individuals with DD and to extend the previous descriptive analysis studies by evaluating demographic, methodological and outcomes characteristics of the studies published in the last 10 years (2009-2019). Specifically, the following research questions guided the study: (a) What are the demographic characteristics (age, gender and diagnosis) of the individuals profiled in the studies?; (b) What are the primary training
settings for toilet training programs?; (c) What are the materials utilized in the studies?; (d) Which toilet training program components are included in the toilet training packages?; (e) What are the characteristics of the implementers of toilet training programs?; (f) What is the total length of training in the studies?; (g) What are the experimental designs used in the studies?; (h) Do the studies include information on maintenance, generalization, interobserver agreement and procedural fidelity data? What are the results of them?; (i) Do the studies include information on social validity?, Which method are they collected by?, Which dimensions of social validity are evaluated?, and What are the results?; and (j) What are the effectiveness results of toilet training programs?

Method

Literature Search

This study conducted a descriptive analysis of the papers on investigating the effectiveness of toilet training programs for individuals with DD published among 2009-2019. For this purpose, the studies were identified through systematic databases search. The search entailed two steps. First, ScienceDirect, EBSCOhost, Wiley Online Library and SpringerLink electronic databases were searched using the keywords autism, intellectual disability, mental retardation, developmental disability, toilet (training), and (in)continence. These keywords were searched in abstracts, methods, and results sections of the articles. The literature search produced 497, 162, 137 and 564 publications from the databases respectively, which resulted in a total of 1360 studies. Second, studies were selected if they met eligibility criteria.

Eligibility Criteria

The following inclusion criteria were determined to include relevant studies in the review: (a) publication between the years of 2009-2019, (b) reporting of results for at least one participant diagnosed with DD, (c) designed with a single-subject research model, (d) publication in Turkish or English peer-reviewed journal. On the other hand, the studies were excluded if they (a) included participants all of whom were with no diagnosis, (b) did not use a single-subject experimental design, (c) published in a non-peer-reviewed journals (d) did not investigate the effectiveness of a toilet training component, package or program. On the basis of the eligibility criteria, 23 articles of 1360 were identified and included in the study. Common reasons for excluding articles were article duplications, focus on intervention not investigating the effects of a toilet training program and using a non-single-subject design. The process of identifying eligible studies are depicted in Fig. 1.

Intercoder Agreement

An independent coder who was a Ph.D. candidate and a research assistant in Special Education department analyzed and coded %30.43 (n=7) of the included articles. All of the demographic, methodological and outcome characteristics were coded for the randomly selected articles. The formula of \[
\frac{\text{Total number of intercoder agreements} - \text{Total number of intercoder agreements and disagreements}}{\text{Total number of intercoder agreements and disagreements}} \times 100
\] was used to calculate the intercoder agreement. The first author and the independent coder came together and examined the content in case of any disagreements. However, there were no observed discrepancies between the observers and the intercoder agreement was calculated as 100%.

Results

Descriptive analysis of demographic, methodological and outcome characteristics of the reviewed studies on toilet training to the individuals with DD are listed and summarized in Table 1. The following sections include

Demographic Characteristics

Participants

The reviewed studies included a total of 117 individuals. Demographic characteristics of the participants in the studies were evaluated by (a) identified disability, (b) gender and (c) age.
Age. The majority of the studies provided data on the age of a total of 49 participants (n=21; e.g., Ozcan & Cavkaytar 2009). However, two studies only reported the mean and range years of the participants’ age that were 14.4 (range: 7.6-14.4) and 2.1 (range: 1.5-3.25) (Greer, Neidert, & Dozier 2016; Van Oorsouw et al., 2009). Therefore, they could not be included in classification of age groups. Of the studies that reported age, 21 were between 1-4 years, 21 were 5-10, and 7 were 11 and over.

Gender. Where gender was reported, 86 participants were male and 29 were female. In one study (Greer et al., 2016), the gender of 18 participants were derived from pseudo names given to them. Since two names were gender-neutral, they could not be included in gender classification.

Identified disability. Almost half of the participants (n=55) had a diagnosis of ID. Additional participant diagnoses consisted of ASD (n=24; e.g., Henriksen & Peterson 2013), multiple disability (n=7; e.g., Ozcan & Cavkaytar 2009), developmental delay (n=5; e.g., Brown & Peace 2011), Down syndrome (n=4; e.g., Unlu, 2019), pervasive developmental disorder (n=2; e.g., Ardic & Cavkaytar 2014) and psychomotor developmental disability (n=1; Ardic & Cavkaytar 2014). Furthermore, a total of 19 participants with typical development participated in the studies (e.g., Greer et al., 2016).
### Table 1.

**Methodological Characteristics**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Age-Gender</th>
<th>Diag. Setting</th>
<th>Mat-</th>
<th>Ind.</th>
<th>Design</th>
<th>Ma.</th>
<th>Gen.</th>
<th>IOA and PF</th>
<th>S.V.</th>
<th>Outcomes (Positive Effects/Total Ps)</th>
<th>Implementer</th>
<th>Length of Train.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ozcan &amp; Cavkaytar (2009)</td>
<td>5-M</td>
<td>MD</td>
<td>MD</td>
<td>FTP</td>
<td>MPD</td>
<td>4 w</td>
<td>-</td>
<td>IOA: %100</td>
<td>-</td>
<td>Acq.: + (3/3) Ma.: + (3/3)</td>
<td>Parent</td>
<td>120 d</td>
</tr>
<tr>
<td></td>
<td>4-M</td>
<td>MD</td>
<td>ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5-M</td>
<td>Home</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Oorsouw et al. (2009)</td>
<td>Mean: 14.4</td>
<td>ID</td>
<td>Home</td>
<td>Diapers</td>
<td>H, PR,</td>
<td>4 w</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>Caretaker</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.6-14.4)</td>
<td>F(n=15)</td>
<td></td>
<td></td>
<td>RoP, SS,</td>
<td>24 w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>M(n=33)</td>
<td></td>
<td></td>
<td>GG, P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kroeger &amp; Sorensen (2010)</td>
<td>4-M</td>
<td>ASD</td>
<td>Home</td>
<td>Bever-</td>
<td>H, SS, Paper, Chair</td>
<td>2 w</td>
<td>Home</td>
<td>IOA: %100</td>
<td>+</td>
<td>Parent-Scale Acceptability and Effectiveness</td>
<td>Parent</td>
<td>4-5 d</td>
</tr>
<tr>
<td></td>
<td>6-M</td>
<td>ASD</td>
<td></td>
<td>age, Toys, Edibles, Chair</td>
<td>24 w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Home</td>
<td></td>
<td></td>
<td>ABA</td>
<td>3 y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang et al. (2011)</td>
<td>9-M</td>
<td>MD</td>
<td>School</td>
<td>Computer, Diapers</td>
<td>ABAB</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>Acq.: + (1/1)</td>
<td>Researcher</td>
<td>30 d</td>
</tr>
<tr>
<td>Sonmez &amp; Aykut (2011)</td>
<td>5-M</td>
<td>Down</td>
<td>Home</td>
<td>Potty, Edibles</td>
<td>FTP, SP, PR, PNA</td>
<td>5 d</td>
<td>Bath-</td>
<td>-</td>
<td>-</td>
<td>Acq.: + (1/1) Ma.: + (1/1) Gen.: + (1/1)</td>
<td>Parent</td>
<td>7 d</td>
</tr>
</tbody>
</table>
Toilet Training Individuals with Developmental Delays

<table>
<thead>
<tr>
<th>Authors</th>
<th>Age-Gender</th>
<th>Diagnosis</th>
<th>Setting</th>
<th>Materials</th>
<th>Ind. Variables</th>
<th>Design</th>
<th>M a.</th>
<th>Gen.</th>
<th>IOA and PF</th>
<th>S.V.</th>
<th>Outcomes (Positive Effect/Total Ps)</th>
<th>Implementer</th>
<th>Lengh of Train.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Henriksen &amp; Peterson (2013)</td>
<td>12-F</td>
<td>ASD</td>
<td>Home</td>
<td>Edibles</td>
<td>PR, PNA, EA</td>
<td>AB</td>
<td>3</td>
<td>Any other house</td>
<td>-</td>
<td>-</td>
<td>Acq.:+(1/1) Ma.:+(1/1) Gen.:+(1/1) Parent</td>
<td>35 d</td>
<td></td>
</tr>
<tr>
<td>Ozkubat &amp; Toret (2014)</td>
<td>7,5-M</td>
<td>ID</td>
<td>School</td>
<td>Timer</td>
<td>SCT, ES</td>
<td>MBD</td>
<td>2</td>
<td>w</td>
<td>-</td>
<td>Parent-Interview-Acceptability and Effectiveness</td>
<td>Acq.:+(3/3) Ma.:+(3/3) Gen.:+(3/3) S.V.: (3/3) Parent</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Lee et al. (2014)</td>
<td>4-M</td>
<td>ASD</td>
<td>Home</td>
<td>Picture cards, Toys, Diapers</td>
<td>ES, VM, SS, GG, EA, H, PR, CT</td>
<td>CCD</td>
<td>5</td>
<td>d</td>
<td>Sch ool</td>
<td>-</td>
<td>Parent, Teacher-Scale, Interview-Acceptability, Understandability</td>
<td>Acq.:+(1/1) Ma.: + (1/1) Gen.: + (1/1) S.V.: + (2/2) Researcher</td>
<td>-</td>
</tr>
<tr>
<td>Authors</td>
<td>Age-Gen-der</td>
<td>Diag.</td>
<td>Setting</td>
<td>Materials</td>
<td>Ind. Variable</td>
<td>Design</td>
<td>Ma.</td>
<td>Gen.</td>
<td>IOA and PF</td>
<td>S.V.</td>
<td>Outcomes (Positive Effect/Total P)</td>
<td>Implementer</td>
<td>Length of Train</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------</td>
<td>-------</td>
<td>---------</td>
<td>--------------------------------</td>
<td>---------------</td>
<td>--------</td>
<td>-----</td>
<td>------</td>
<td>------------</td>
<td>------</td>
<td>--------------------------------</td>
<td>----------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Drysdale et al. (2015)</td>
<td>4.1-M 5-M</td>
<td>ASD</td>
<td>Home</td>
<td>Tablet PC, Picture Cards, Diapers</td>
<td>VM, CT, SS, ES, PR, GG, RsP</td>
<td>MBD</td>
<td>4 h</td>
<td>School</td>
<td>IOA: %99</td>
<td>+ Parent-Scale-Acceptability, Understandability</td>
<td>Parent and Researcher</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>McLay et al. (2015)</td>
<td>8.1-M 7.2-M</td>
<td>ASD</td>
<td>Home</td>
<td>Edibles, Tablet PC</td>
<td>CT, VM, PR, ES, GG, DC, RsP</td>
<td>MBD</td>
<td>12 w</td>
<td>School</td>
<td>IOA: %96.3</td>
<td>+ Parent-Scale-Acceptability, Understandability</td>
<td>Parent</td>
<td>12-29 d</td>
<td></td>
</tr>
<tr>
<td>Ohtake et al. (2015)</td>
<td>12-M</td>
<td>ASD</td>
<td>Home</td>
<td>DVD Player</td>
<td>VM, PR, GG, EA</td>
<td>MBD</td>
<td>8 w</td>
<td>-</td>
<td>IOA: %93</td>
<td>+ Teacher-Interview-Process</td>
<td>Teacher</td>
<td>21-28 d</td>
<td></td>
</tr>
<tr>
<td>Axelrod et al. (2016)</td>
<td>13-M 14-M</td>
<td>MD</td>
<td>Home and School</td>
<td>Medications Timer, Toys, Stepstool</td>
<td>SS, PR, P</td>
<td>MBD</td>
<td>2 w</td>
<td>3 w</td>
<td>IOA: %100</td>
<td>+ Acceptability, Process, Effectiveness</td>
<td>Parent and school staff</td>
<td>63-70 d</td>
<td></td>
</tr>
<tr>
<td>Doan &amp; Toussaint (2016)</td>
<td>2.6-M 5.10-M 4.9-M</td>
<td>ASD</td>
<td>Home and Clinical Setting</td>
<td>Toys, Edibles, Diapers</td>
<td>PR, SS, H, P</td>
<td>MBD</td>
<td>1 w</td>
<td>3 w</td>
<td>IOA: %100</td>
<td>+ Parent-Scale-Acceptability, Process, Effectiveness</td>
<td>Parent</td>
<td>13-29 d</td>
<td></td>
</tr>
<tr>
<td>Greer et al. (2016)</td>
<td>Mean: 2.1 (1.5-3.25)</td>
<td>ASD TD(n=19)</td>
<td>Clinical Setting</td>
<td>Underwear, Potty</td>
<td>PNA, PR, SS</td>
<td>MBD</td>
<td>-</td>
<td>-</td>
<td>IOA: %93</td>
<td>+</td>
<td>Teacher</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
Toilet Training Individuals with Developmental Delays

<table>
<thead>
<tr>
<th>Authors</th>
<th>Age-Gender</th>
<th>Diag.</th>
<th>Setting</th>
<th>Materials</th>
<th>Ind. Variable</th>
<th>Design</th>
<th>Ma.</th>
<th>Gen.</th>
<th>IOA and PF</th>
<th>S.V.</th>
<th>Outcomes (Positive Effect/Total Ps)</th>
<th>Implementer</th>
<th>Length of Train</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mruzek et al. (2016)</td>
<td>15.8-M</td>
<td>ID</td>
<td>School</td>
<td>Edibles, Beverages</td>
<td>H, PR, SS, SCT, EA</td>
<td>MBD</td>
<td>-</td>
<td>-</td>
<td>PF: %93</td>
<td></td>
<td>Acq.: 2/3 S.V.: 7/7</td>
<td>Teacher and paraprofessional</td>
<td>37-38 d</td>
</tr>
<tr>
<td>Call et al. (2017)</td>
<td>8-M</td>
<td>MD</td>
<td>Clinical Setting</td>
<td>Potty, Medications, Edibles</td>
<td>PR, SS</td>
<td>MBD</td>
<td>4 w</td>
<td>-</td>
<td>IOA: %100</td>
<td></td>
<td>Acq.: 3/3 Ma.: 3/3</td>
<td>Therapist and Parent</td>
<td>13-21 d</td>
</tr>
<tr>
<td>Call et al. (2017)</td>
<td>8-M</td>
<td>MD</td>
<td>Clinical Setting</td>
<td>Potty, Medications, Edibles</td>
<td>PR, SS</td>
<td>MBD</td>
<td>4 w</td>
<td>-</td>
<td>IOA: %99</td>
<td></td>
<td>Acq.: 1/1 Ma.: 1/1</td>
<td>Researcher and Parent</td>
<td>39 d</td>
</tr>
<tr>
<td>Byra et al. (2018)</td>
<td>5.9-M</td>
<td>ASD</td>
<td>Clinical Setting</td>
<td>Table, Chair, Picture Card, Potty, Doll, Edibles</td>
<td>VA, SLP, PR</td>
<td>AB</td>
<td>24 w</td>
<td>Hom. e</td>
<td>IOA: %100</td>
<td></td>
<td>Acq.: 2/2 Ma.: 2/2 Gen.: 2/2</td>
<td>Researcher</td>
<td></td>
</tr>
<tr>
<td>Unlu (2019)</td>
<td>2.5-F</td>
<td>Dow n</td>
<td>Home</td>
<td>Edibles, Diapers</td>
<td>DC, SS, PR, H, FTP,</td>
<td>AB</td>
<td>1 w</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Acq.: 3/3 Ma.: 3/3</td>
<td>Mother</td>
<td>20 d</td>
</tr>
</tbody>
</table>
Methodological Characteristics
Methodological characteristics of the reviewed studies were evaluated by (a) training setting; (b) materials; (c) independent variable; (d) experimental design; (e) implementer; (f) length of training; (g) maintenance; (h) generalization; (i) reliability, and (j) social validity.

Training setting
The most common setting where the toilet training occurred was participants’ own homes (n=12; e.g., Kroeger & Sorensen 2010) followed by schools where the participants were getting education (n=5; Brown & Peace 2011), clinical setting such as rehabilitation center (n=4; e.g., Doan & Toussaint 2016), both home and school (n=1; Axelrod, Tornehl, & Fontanini-Axelrod 2016).

Materials
The training materials across studies varied greatly. Nearly half of the studies included edibles (n=13; e.g., Cocchiola, Martino, Dwyer, & Demezzo 2012) and eight used diapers (e.g., Lee, Anderson & Moore 2014). Toys were employed in six studies (e.g., Byra, White, Temple, & Cameron 2018), potty in five (e.g., Sonmez & Aykut 2011), timer and beverages in four studies (e.g., Axelrod et al., 2016; Ozkubat & Toret 2014). Additionally, picture cards (e.g., Byra et al., 2018), technological devices such as DVD player and computer (e.g., Chang et al., 2011) were utilized in three studies, medications (Axelrod et al., 2016; Call, Mevers, McElhanon, & Scheithauer 2017), stepstools (e.g., Rinald & Mirenda 2012) and chairs (e.g., Kroeger & Sorensen 2010) in two studies. One study included book as a training material (Rinald & Mirenda 2012). However, one study provided no information about the training materials (Ozcan & Cavkaytar 2009).

Independent variable
It is notable that all studies used toilet training programs and packages that mainly included applied behavior analysis strategies and techniques. Across 23 studies reviewed, the large majority (n=21) included positive reinforcement (PR) (e.g., Byra et al., 2018) followed by scheduled sitting (SS) (n=10; Cocchiola et al., 2012). Hydration (H) and graduated guidance (GG) from errorless teaching methods were used in 10 studies (e.g., Van Oorsouw et al., 2009), punishment procedures in six (e.g., Brown & Peace 2011) and diaper/pad removal (DPR) (i.e., Drysdale, Lee, Anderson & Moore 2015) and urine alarm (i.e., Ohtake, Takahashi, & Watanabe 2015) in five studies. Toilet training protocols in four studies included elimination schedule (e.g., Drysdale et al., 2015), communication training (e.g., Brown & Peace 2011) and video modeling (VM) (e.g., Drysdale et al., 2015). Both studies used decreasing the intensity of interaction with the participant in case of an accident (e.g., Greer et al., 2016), dry checks (DC) (e.g., Ardic & Cavkaytar 2014), scheduled chair sittings (SCS) (e.g., Kroeger & Sorensen 2010) and transfer of stimulus control (e.g., Mruzek, McAleavey, Engel, & Smith 2016), simultaneous prompting (Sonmez & Aykut 2011) and system of least prompts (SLP) and visual aids (Byra et al., 2018) were included in one study. One study investigated the effectiveness of parent training program on the target skill, yet did not report which toilet training components were used (Ozcan & Cavkaytar 2009).

Research designs
In order to provide the methodological rigor of the studies, the experimental design of each study was recorded. It is notable that all studies utilized effectiveness designs to assess intervention effects (e.g., Kroeger & Sorensen 2010). In this regard, more than a half of the studies (n=12) employed multiple baseline studies (e.g., Cocchiola et al., 2012). Four of the studies targeting toileting skills evaluated the results with AB design (e.g., Unlu, 2019) while multiple probe design was used in two studies (Ardic & Cavkaytar 2014; Ozcan & Cavkaytar 2009). Kroeger & Sorensen (2010) employed ABA design, while Chang et al. (2011) and Lee et al. (2014) and Sutherland et al. (2017) used ABAB, ABCA and ABC design respectively.
Implementers
With respect to the implementers, the parents represented the majority of the implementers (n=10; e.g., Rinald & Mirenda 2012) followed by experimenters in four studies (e.g., Ardic & Cavkaytar 2014). In three studies, both parent and experimenter were implemented the procedure (e.g., Call et al., 2017). Teachers implemented toilet training programs in two studies (e.g., Greer et al., 2016), a caregiver in one (Van Oorsouw et al., 2009) and both a teacher and paraprofessional aides in one (Mruzek et al., 2016). Lastly, one study reported the procedure was implemented by the child’s grandmother and staff at school (Axelrod et al., 2016).

Length of training
This study also measured the total length of training time in order to determine participants’ mastery on the implementation of toilet training protocols as it is a significant measure for efficiency (Halcombe, Wolery, & Gast, 1994). In 16 of the studies, researchers reported the length of training. These studies are grouped by months. In fact, total training time took less than a month in half of the studies (n=9; e.g., Kroeger & Sorensen 2010), between 1-2 months in five (e.g., Brown & Peace 2011) and more than 2 months in two (e.g., Axelrod et al., 2016). Cocchiola et al. (2012) reported total training time varied between 32 and 88 days.

Maintenance
Across the 23 studies reviewed, maintenance measures were evaluated in large majority (n=19; e.g., Ozcan & Cavkaytar 2009) while the remaining did not provide information regarding this (e.g., Ardic & Cavkaytar 2014). In those studies, maintenance was evaluated between one and three probes. Almost half of the studies (n=11) evaluated maintenance with one probe between five days and six months after the intervention (e.g., Ozcan & Cavkaytar 2009); five studies with two probes between two weeks and six months (e.g., Van Oorsouw et al., 2009) and four studies with three probes between one week and up to three years (e.g., Brown & Peace 2011).

Generalization
Only generalization across settings was evaluated due to the nature of the skill. The generalization was coded if the target skill was measured in a context different from training setting. In this regard, only 10 studies presented data measuring generalization (e.g., Kroeger & Sorensen 2010). Six of those conducted generalization training in home setting (e.g., Brown & Peace 2011), three in participants’ school (e.g., Drysdale et al., 2015). In one study of Rinald & Mirenda (2012), a systematic generalization was not programmed; however, it was reported participants generalized the target skills according to parent report.

Reliability
Reliability was coded in the studies by two aspects: (a) IOA and (b) PF. 17 studies collected and analyzed interobserver agreement data on occurrence of target behaviors (e.g., Byra et al., 2018), which ranged from 92% to 100%. In Rinald & Mirenda (2012), authors calculated IOA by asking questions to participant’s parent implementer, yet did not report the results. On the other hand, a number of studies measured PF (n=12; e.g., Drysdale et al., 2015). When PF was reported, it ranged from 93% to 100%. In Drysdale et al. (2015), despite vaguely reported, PF calculation was stated as over 95%.

Social validity
Across 23 studies included in this review, studies rarely contained measurement of social validity. With respect to this, only 10 studies collected social validity data, and all used subjective evaluation strategy to analyze specific aspects of the intervention process (e.g., Rinald & Mirenda 2012). Seven of them collected social validity information from participants’ parents (e.g., Ozcan & Cavkaytar 2009), one from teachers (Ohtake et al., 2015), one from both parents and teachers (Lee et al., 2014) and one from teachers and paraprofessionals (Mruzek et al., 2016). The remaining of the studies did not provide information regarding social validity assessment (e.g., Brown & Peace 2011). With respect to data collection techniques, six studies collected social validity information through scale (e.g., Rinald & Mirenda 2012), three by interview (e.g., Ozkubat &
Toilet Training Individuals with Developmental Delays, 130

Toret 2014) and one by both (Lee et al., 2014) to evaluate the experimental effects of the independent variable(s). Those that used scale to collect this information analyzed specific aspects of the intervention process as acceptability, understandability of the intervention, training process, effectiveness of training program, behavior change and importance of the study. On the other hand, studies using interview asked individuals questions to analyze the aspects of training process, effectiveness of training program and its acceptability.

Outcome Characteristics
The studies were analyzed and reported by acquisition, maintenance, generalization and social validity.

Acquisition
Across 23 studies, toilet training programs were mostly effective on teaching target skill(s). Mruzek et al. (2016) reported two out of three children substantially progressed in acquisition but one child kept on toilet accidents despite some adaptations. Secondy, a study of Greer et al. (2016) reported eight children initially did not improve in toileting skills who then acquired the skills following implementation of additional components. In short, 99.14% of the participants in all studies included in the review acquired the target skills. However, it should be noted that Lee et al. (2014) stated the intervention was effective on dressing, sitting on toilet and flushing but not on acquisition of in-toilet voiding.

Maintenance
Across the studies reviewed, the skills acquired maintain for all participants after the termination of intervention process (e.g., Henriksen & Peterson 2013). However, Rinald and Mirenda (2012) reported they could not measure maintenance for one participant as he withdrew from the study after acquisition of target skills.

Generalization
Where generalization measured, eight studies reported the skills could generalize across a different setting. However, in a study of Sonmez and Aykut (2011), one out of three participants could not generalize the acquired skill.

Social Validity
The studies measuring social validity reported positive social validity outcomes (e.g., Kroeger & Sorensen 2010). However, Ardic and Cavkyatar (2014) reported one family stated their satisfaction with the study was not in a great scale as the intervention process did not include bowel control training. Furthermore, in a study of Rinald and Mirenda (2012) the authors could not measure social validity for one participant as he withdrew from the study.

Discussion
The current review evaluated studies on teaching toilet training skills to the individuals with DD. 23 studies that met inclusion criteria were included in the review. The overall results showed that almost all individuals with DD were trained on the implementation of toilet training programs, generalized the target skill(s) across other settings and maintained them after termination of the intervention. Furthermore, the current literature review indicated that most toilet training programs include components used in the earliest method, Rapid Toilet Training, developed by Azrin and Foxx (1971) and that there does not exist a standardized toilet training program, thus researchers compose toilet training intervention packages by components according to their preferences. In this regard, a number of prominent results that emerged in the current study may guide future research and practice for individuals with DD.

First of all, the majority of participants in the studies had a diagnosis of ASD and ID (n=81; e.g., Henriksen & Peterson 2013), which weaken generalizability of effectiveness results of the toilet training intervention packages across other disability groups. One possible reason for this could be that population with these diagnoses have more difficulties in acquisition of toileting skills than those with other disability types (Kircaali-Iftar, Ulke-Kurkuoglu, & Kurt 2014; Leader et al., 2018). Moreover, difficulty in learning toilet training skills for individuals with DD and in teaching due to length of time and effort can be deterrent and disincentive factors for educators and researchers in teaching toileting skills.
Considering the nature of the target skills and high preference of reinforcement-based strategies in the studies, inclusion of edibles and diapers in toilet training programs is meaningful. Interestingly, two recent studies have used medications for toilet training participants (Axelrod et al., 2016; Call et al., 2017). One possible reason for utilization of medications is because they might yield in faster acquisition for the individuals with DD. Nevertheless, future research of medication component is recommended to entail a clinical setting or involvement of medical staff for side effect analysis, approving administration of medications or consultancy.

Technology-based methods or components still appear to be rare but emerging in toilet training research. In fact, technology was utilized in 35% of the studies reviewed (e.g., Chang et al., 2011), which was 25% in the review study by Kroeger & Sorensen-Burnworth (2009). Considering the length of the interventions are over one month in half of the studies reviewed (e.g., Cocchiola et al., 2012), toilet training is quite burdensome interventions. Thus, including wireless (i.e., enuresis alarms) and digital technology (i.e., tablet computers) systems in protocols may provide implementers with easiness, thus streamlining the intensity and the length of the interventions. Moreover, advancements in technology may improve “cost-benefit” of the toilet training interventions in that such technological devices can be used alone in toilet training individuals with DD, which is an important topic for future investigations.

Within the current review, a prominent result was that independent variable in all studies was a training package composed of more than one method or toilet training components. Although toilet training programs are not alone evidence-based practices, they are comprised of such evidence-based practices as errorless teaching method (e.g., SLP and GG) and visual aids (e.g., VM). In that regard, it can be thought toilet training programs are not standardized training packages, as such, they are composed by researchers’ preferences, which means no precise agreement on toilet training components for a toilet training program still exists in literature. Even so, it is important to note that one toilet training component, PR, was included in the training packages in almost
all studies reviewed (n=21). Previous literature review studies outlined that the component had been commonly used in toilet training intervention programs since the development of the most cited and comprehensive toilet training protocol by Azrin and Foxx in 1971 (Francis et al., 2017; Kroeger & Sorensen-Burnworth 2009). Therefore, it can be proposed that PR is a standard component in a toilet training package. Pertaining to VM, our evaluation permitted the identification of interest in the method although Kroeger & Sorensen-Burnworth (2009) reported only two studies had utilized VM. In a study of Ohtake et al. (2015), for instance, the participant watched an animated hero in video with which the student was pre-occupied and improved his target behaviors. Considering that toileting is of private and personal, researchers can avoid violating both participants’ and families’ privacy and ethical issues through animated video modeling. Future studies can continue to further evaluate the effectiveness of videos in which animated cartoon characters are incorporated on teaching in-toilet elimination in addition to chains toileting skills such as dressing, flushing and washing hands.

No previous reviewed studies evaluated who implemented toilet training programs to participants with DD (Francis et al., 2017; Kroeger & Sorensen-Burnworth, 2009). Kroeger & Sorensen-Burnworth (2009) reported that the least used intervention was family training programs on teaching toileting skills. However, it should be noted that the current review found at least one implementer was participant’s parent in more than half of the studies reviewed most of which was conducted in the past five years (e.g., McLay et al., 2015) and that three studies investigated the effectiveness of family training programs on target skill(s) (Ozcan & Cavkaytar 2009; Rinald & Mirenda 2012; Sonmez & Aykut 2011). The trend of focus in training toward family member implementers should mean researchers are giving higher level of importance to recognition of family involvement. Considering family involvement in education of individuals with DD contributes to longer maintenance of acquired skills and helps them and their families fulfill their needs (Ozcan & Cavkaytar 2009), future research are warranted to investigate the effectiveness of intervention not only for participants’ performance, but also in terms of parental self-efficacy, success perceptions, and so on.

The majority of the studies reported length of training duration (n=17) which previous toilet training reviews also analyzed (Francis et al., 2017; Kroeger & Sorensen-Burnworth 2009). Within the current review, we found that the studies in which length of training was short (1-30 days) utilized the components of rapid toilet training method (e.g., PR, H, SCS and SS) developed by Azrin & Foxx (1971). However, those using the components of traditional toilet training method (e.g., DC and P) associated with bladder/bowel distension, reinforcement- and punishment-based strategies of behaviorist approach took longer (over 60 days). Therefore, future research should further evaluate which combinations of the components are more effective and efficient in order to help researchers and families implement a procedure to individuals with DD in the most effective and efficient manner possible.

No previous review studies on toilet training evaluated information in terms of experimental design (Francis et al. 2017; Kroeger & Sorensen-Burnworth 2009). However, we found that slightly more than half of the studies reviewed used multiple baseline design (n=12). This may be due to multiple baseline designs require repeated measurements during baseline, which increases credibility. Furthermore, a prominent finding of the current study is that only effectiveness of toilet training programs or components were investigated across all reviewed studies. In another words, there has been no comparative single-subject studies on examining efficacy of toilet training packages, programs or components for individuals with DD. Determining which toilet training components are efficient may be of significance to build easily implementable toilet training protocols that take a short time to complete, which should be the focus of future research endeavors.

The studies measuring data in maintenance represented the greater number (n=19), which previous toilet training reviews did not focus on. However, it should be noted that half of the studies measured maintenance with one probe and that 15 studies collected the maintenance data within one month following the termination of intervention. On the other
hand, a few studies collected long-term maintenance data which ranged from 2 and 3 years (Brown & Peace 2011; Kroeger & Sorensen 2010). Although these studies reported positive maintenance effect to a large extent, Hyams, McCoull, Smith and Tyrer (1992) reported 14 participants with ID could not maintain toileting skills 10 years after they acquired the skills. In this sense, future longitudinal research of larger sample sizes is needed to further measure long-term maintenance data of toileting skills since they are of great importance for participants’ with DD life quality, hygiene and social acceptance.

Nine of the studies included data on generalization of the target skill(s) and one study reported that they did not collect generalization data. However, participant’s parent informed researcher of positive effects for generalization of the skills across home setting (Rinald & Mirenda 2012). It is notable that none of the studies conducted in structured settings such as school or clinical settings (n=8) included information on generalization (e.g., Ardic & Cavkaytar 2014). If a person cannot apply a skill or concept across other situations (e.g., new people or settings), learning cannot be established since generalization is one of the stages of learning (Wolery, Ault, & Doyle 1992). Therefore, it is recommended future research be conducted for evaluating data in generalization whether the targeted skills will generalize across other settings.

Reporting critical treatment steps in studies may assist researchers develop a standardized and effective toilet training package. Slightly more than half of the studies reviewed (n=12) included data with respect to PF and IOA. Therefore, it is suggested that more studies be conducted for collecting both types of reliability data.

More than a half of the reviewed studies (n=13) did not appear to have assessed social validity. However, such information is necessary to measure stakeholder’s perceptions of social acceptability of the toilet training programs consisting of various components. Among the other studies that included social validity information, most of the studies (n=9) reported ratings of teachers and parents were high for training process, acceptability, effectiveness and understandability of the intervention (e.g., McLay et al., 2015). However, in a study by Ardic & Cavkaytar (2014) one participant’s mother stated she was not completely satisfied with the treatment as it did not include bowel control training. Among the studies that did measure social validity information, all collected data from indirect consumers (participants’ teachers and parents). The aim of measuring social validity is to identify the significance of behavior change and the acceptability of procedure which generates the change (Carr, Austin, Britton, Kellum, & Bailey 1999). In this sense, no studies specifically evaluated the consumers’ perceptions of toilet training components. On the other hand, evaluating the opinions, perceptions and suggestions of direct consumers (e.g., the participants with DD) toward treatment, intervention and toilet training components, in particular, may not only contribute to higher quality of scientific experiments, but also guide researchers to develop a standardized toilet training program or protocol in the future. Additionally, considering the sensitive and private nature of toileting skills, it is highly important for implementers to understand and protect participant’s privacy. Even so, of the ten studies from the current review that did report on social validity, none included questions or items pertaining to ethical rules or issues. Thus, including such questions in social validity measurement may guide researchers plan intervention procedure and help students protect their rights and privacy.

With respect to outcomes from the studies, a review study on toilet training by Kroeger and Sorensen-Burnworth (2009) evaluated 28 studies and reported positive outcomes for all children in 23 studies. Authors stated one participant in four studies could not improve the target toileting and that only 19 out of 40 participants improved the target skill in one study. Another review study by Francis et al. (2017) reported participants in one single-subject study could not acquire toileting skills. Furthermore, it should be noted that both review studies showed the evaluation of maintenance and generalization of the studies reviewed, yet did not report performances of participants in detail. However, as mentioned previously, the current study found the target skill(s) could effectively enhance, maintain and generalize across various settings in most of the studies included in the review for the individuals with DD. In fact, Mruzek et al. (2016) reported the
training package did not facilitate positive outcomes for one participant despite a number of adaptations made including differential reinforcement and increasing the intensity of reinforcers. Furthermore, Lee et al. (2014) indicated the intervention was not effective in teaching in-toilet elimination but dressing, sitting on toilet and flushing for the participant, although they made adaptations by adding in-vivo modeling component in intervention. In another study by Greer et al. (2016) on evaluation of toilet training components, it was reported dense SS and differential reinforcement components were not effective for eight children. However, they improved the target skills upon implementation of additional component, wearing underwear.

In conclusion, all of the toilet training programs in the studies included in the current study are derivatives and modified versions of original Azrin and Foxx study (1971). Furthermore, no standardized toiletting programs of fixed components for individuals with DD have existed ever since. With respect to this issue, it can be concluded that toilet training programs, packages or protocols are built according to researchers’ preferences and participant characteristics, as such, there is still no agreement on a common toiletting program in literature. Therefore, further studies are needed to teach toilet training skills to individuals with DD.

References


https://doi.org/10.1093/jpepsy/27.7.5


Sonmez, N., & Aykut, Ç. (2011). Using mother delivered simultaneous prompting for teaching independent

Toilet Training Individuals with Developmental Delays

Toileting skills to a child with developmental disability. *Journal of Human Sciences*, 8(1), 1151-1171.


