Measurement invariance of the Positive Gains Scale in families of children with and without disabilities

Mikeda Jessa, Tom Bailey, Ineke M. Pit-ten Cate, Vasiliki Totsika, Richard P. Hastings

Centre for Educational Development, Appraisal, and Research (CEDAR), University of Warwick, UK
Luxembourg Centre for Educational Testing (LUCET), University of Luxembourg, Luxembourg
Division of Psychiatry, University College London, UK
Centre for Developmental Psychiatry and Psychology, Department of Psychiatry, Monash University, Australia

ARTICLE INFO

Number of reviews completed is 2
Keywords:
- Developmental disability
- Measurement invariance
- Case-control studies
- Positive gain

ABSTRACT

Background: Despite the high frequency of case-control studies in the developmental disability literature, there is a paucity of research establishing the measurement equivalence of instruments used, and particularly those relating to positive perceptions and experiences in family disability research.

Aims: The present study sought to establish measurement invariance for the Positive Gains Scale (PGS) across 1219 mothers of children with developmental disabilities, 234 mothers of children with spina bifida/hydrocephalus, and 157 mothers of children without disabilities.

Methods and procedures: A three-step test for measurement invariance across the three groups was conducted using Multigroup Confirmatory Factor Analysis.

Outcomes and results: Loadings between the three groups were invariant, suggesting the criteria to assume metric invariance was met. However, the assumption of scalar invariance was not met, suggesting that item intercepts differed between the three groups.

Conclusions and implications: Our findings suggest that the PGS cannot be meaningfully used to compare outcomes between mothers of children with developmental disabilities and other mothers. These findings may have wider implications for research utilising well-being measures to make comparisons with carers of children with developmental disabilities.

What this paper adds?

The present paper calls into doubt the fundamental assumptions behind case-control studies within the developmental disabilities family research literature. Such studies aim to identify differences and similarities between groups with developmental disabilities and those without using instruments hypothesised to measure the same construct in the same way across the different groups. Our findings suggest that this hypothesis cannot be supported in the case of the Positive Gain Scale (PGS; Pit-ten Cate, 2003) and hence scores cannot be meaningfully compared between mothers of children with developmental disabilities, those chronic physical health conditions, and mothers of children who are typically developing.
1. Introduction

One of the most common methodological paradigms for research that aims to describe the experiences of families raising a child with developmental disabilities is group comparison to families raising a child without disabilities. Findings from such “case-control” studies indicate that parents raising a child with developmental disabilities generally experience poorer mental health than parents raising a child without a disability (Estes et al., 2013; Totsika, Hastings, Emerson, Berridge, & Lancaster, 2011). Maternal psychological adjustment has also been compared among various developmental disability groups. For example, mothers of children with autism spectrum disorder (ASD) have reported lower psychological well-being compared with mothers of children with cerebral palsy, Fragile X syndrome, and Down syndrome (Abbeduto et al., 2004; Griffith et al., 2011; Pisula, 2007). Identifying differences between groups may support the development of targeted interventions and supports for families.

Whilst findings from studies that focus on negative impact of caring for a child with developmental disabilities are useful in that they highlight challenges for families, such studies perpetuate an unfavourable narrative and do not allow for a broader insight into the adjustment of these families (Hastings, 2016). Moving away from the focus on negative aspects associated with disability, there is a growing interest within disability family research on positive perceptions and experiences. In an earlier review, Hastings and Taunt (2002) highlighted that although parental positive perceptions appeared to exist alongside negative experiences, measurement of positive perceptions was largely neglected within disability family research. Hastings and Taunt hypothesised a number of different functions for parents’ positive perceptions. However, to examine the development and functions of parental positivity researchers must be able to measure positive perceptions and experiences explicitly.

The Positive Gain Scale (PGS; Pit-ten Cate, 2003) is a seven-item measure originally developed to assess parental perceptions of positive aspects of raising a child with a disability. Since its development, the PGS has been used in research of families who have children with developmental disabilities (MacDonald, Hastings, & Fitzsimons, 2010; Minnes, Perry, & Weiss, 2015; Weiss, MacMullin, & Lunska, 2015). The PGS has been reported to have a high level of content and construct validity in families of children with spina bifida or hydrocephalus (Pit-ten Cate, 2003) and high levels of internal consistency in families of children with asthma (Pit-ten Cate, 2003), ASD, intellectual disability, and a variety of rare genetic syndromes associated with intellectual disability (Adams et al., 2018; Griffith et al., 2011; Halstead, Ekas, Hastings, & Griffith, 2018; Jones, Hastings, Totsika, Keane, & Rhule, 2014; MacDonald et al., 2010; Weiss et al., 2015). However, the scale has not yet been subjected to more rigorous psychometric evaluation. Specifically, if the PGS is to be used in group comparison studies including families of children with or without various disabilities, similar measurement properties of the scale must be present for these different groups of parents to enable valid statistical inferences (Rasmussen, Verkuilen, Ho, & Fan, 2015). The requirement for such consistency in measurement properties of the PGS across groups necessitates that the measurement invariance (also termed measurement equivalence; cf. Van de Schoot, Lugtig, & Hox, 2012) of the PGS is established.

In the field of developmental disabilities research, case-control designs are often employed to gauge the level of differences between individuals with developmental disabilities (or their families) and those with other disabilities or without disability (typically developing). This usually entails using measures that were not originally developed for individuals with developmental disabilities. Despite the frequent application of this methodological paradigm, research that tested measurement invariance of measures in developmental disabilities research is limited. For example, research has established measurement invariance for a measure of self-concept across adolescents with either developmental disability or physical illness and typically developing adolescents (Ferro & Boyle, 2013), partial invariance for the Physical Self-Inventory (Maïano, Morin, Bégarie, & Ninot, 2011), and lack of invariance for Wechsler scales (MacLean, McKenzie, Kidd, Murray, & Schwannauer, 2011) between individuals with and without developmental disabilities. Such mixed findings highlight the importance of testing whether measures operate in a similar manner across different groups of participants. Therefore, the aim of the current study was to assess the measurement invariance for the PGS across three groups: mothers of children with developmental disabilities, mothers of children with a spina bifida/hydrocephalus, and mothers of children without disabilities (typically developing). These three conditions were selected to reflect the groups of children most often included in case-control studies that collect PGS data. It was decided to examine data from mothers as they are typically the primary carer for the child.

2. Methods

2.1. Participants

A total of 1610 mothers provided PGS data: 1219 (75.7%) were mothers of children with developmental disabilities, 234 (14.5%) were mothers of children with spina bifida/hydrocephalus and there were 157 (9.8%) mothers of typically developing children.

In the typically development group, participants were mothers of 70 girls (44.6%) and 87 boys (55.4%) ranging from 5 to 12 years old with a mean age of 9.24 years ($SD = 3.08$). The spina bifida/hydrocephalus group included mothers of 103 girls (44.0%) and 131 boys (56.0%) with an age range of 6 to 13 years and a mean age of 9.26 years ($SD = 2.08$). Thirty-nine (16.7%) of the children had spina bifida, 137 (58.5%) had hydrocephalus, and 58 (24.8%) children were diagnosed with spina bifida and hydrocephalus. In the developmental disabilities group (children with intellectual disability or ASD), there were mothers of 879 boys (72.1%) and 337 girls (27.6%) ranging from 3 to 17 years old with a mean age of 8.71 years ($SD = 1.81$). 603 (49.5%) mothers of children with developmental disabilities were in paid employment, as were 120 (51.3%) of the mothers of children with spina bifida/hydrocephalus, and 129 (82.2%) of the mothers of the typically developing children.

A one way analysis of variance showed that the three groups did not significantly differ statistically with respect to age, $F(2,1545)$
Through seven items is valid across the three groups. PGS is equal across groups. In the case of the PGS, this is simply a test that the assumption of a single factor measuring positive gain (McDonald, 1999) in the current study was .88 for the developmental disabilities group, .88 for the spina bifida/hydrocephalus group and .86 for the typical development group, thus indicating excellent internal consistency across groups.

2.2. Positive Gains Scale

Items on the PGS are rated on a five-point scale (1 = strongly disagree, 5 = strongly agree). Five items reflect the perceived benefits for the individual parent (e.g., “since having this child I have grown as a person”), and two reflect positive gains for the family (e.g., “since having this child, my family has become closer to one another”); although the seven items are posited to represent a global positive gains factor (Pit-ten Cate, 2003). Items are recoded so that higher scores indicate greater positive gain. McDonald’s Omega (McDonald, 1999) in the current study was .88 for the developmental disabilities group, .88 for the spina bifida/hydrocephalus group and .86 for the typical development group, thus indicating excellent internal consistency across groups.

2.3. Procedure

Participant data were extracted from multiple UK-based studies. All studies had received ethical approval. The typical development group was recruited from mainstream schools in South-East England. Schools were asked to distribute questionnaires amongst parents of children aged 6–12 years. As some parents had more than one child attending the same school, the mother was asked to complete the questionnaire for their oldest child at that school (Bayless, Pit-ten Cate, & Stevenson, 2008).

Most mothers of children with spina bifida/hydrocephalus completed a follow up postal questionnaire after initially taking part in a comprehensive study concerning the developmental, behavioural and educational characteristics of children with these conditions (Stevenson & Pit-ten Cate, 2003). The sample was recruited through the register of the Association for Spina Bifida and Hydrocephalus (ASBAH; recently rebranded as “Shine”). Families were entered on this register when they contacted ASBAH for information and/or support.

In the developmental disabilities sample, 1018 mothers were drawn from the 1000 Families Study, which is an ongoing UK-wide survey of families with a child with intellectual disability (aged between 5 and 15 years; Hastings et al., 2020). An additional 201 mothers were from a UK-wide study of families who had a child with ASD aged between 5 and 17 years (Petals et al., 2012). In both studies of families with children with developmental disabilities, participants were recruited through multiple routes: special schools, social media advertising, and advertising via disability charities. Survey packs distributed directly to parents (e.g., via the child’s school) included an information sheet, consent form, the survey questionnaire and a prepaid return envelope. Participants could also request a survey pack to be sent to their home. In addition, participants had the option to complete the survey online.

2.4. Approach to statistical analysis

All analyses were conducted in MPlus v 7.1 (Muthén & Muthén, 2017). To evaluate the dimensional structure of the single PGS factor, we first performed a confirmatory factor analysis (CFA) using the data from each of the three groups separately. Then, to ascertain the extent to which the PGS items were similarly related to the latent construct across the three groups, a three step test for measurement invariance was conducted using a Multigroup Confirmatory Factor Analysis (MGCFA). As outlined below, testing for measurement invariance involves increasingly strict equality constraints being imposed on the model items at each step of analysis (Byrne, 2009). Whilst Mardia’s (1970) tests of skewness and kurtosis suggested no statistically significant deviations from multivariate normality; all models were estimated using the robust maximum likelihood estimation (MLR); which provides item standard errors and a chi-square model fit index that is robust to non-normality of observations (cf. Yuan & Bentler, 2000).

Step 1 tested whether participants from each group conceptualised the PGS construct in the same way, and is the baseline model to which subsequent models are compared. Known as configural invariance, this initial step tests whether the factor structure of the PGS is equal across groups. In the case of the PGS, this is simply a test of whether a single factor measuring positive gain through seven items is valid across the three groups.

To assess the model fit of the configural model (and all other models) a number of fit indices were consulted. In large samples, the risk of a Type I error is present if conclusions are made on the basis of the chi-square test only (Hoyle & Panter, 1995). Therefore, we also examined the Root Mean Square Error of Approximation (RMSEA), with values under .08 preferred (Browne & Cudeck, 1993), and the Comparative Fit Index (CFI), with values above .95 indicating good fit (Hu & Bentler, 1999). To allow accurate comparison of factor loadings and intercepts, models were operationalised such that the mean and variance of the latent factor were constrained at 0 and 1 respectively (Van de Schoot et al., 2012).

Step 2 tested whether the relationship between observed PGS items and the latent construct of positive gain (which is operationalised by the factor loadings) is similar across groups. Termed metric invariance, this step constrains that factor loadings of PGS items to be equal across the three groups. If any discrepancies between this model and the baseline configural model arise (in terms of model fit), this suggests that the assumption of equal factor loadings may not be tenable. Specifically, our criteria for testing whether the assumption of metric invariance holds were a non-significant change in (Δ)χ². However, given the sensitivity of this index to minor differences between models, we sought a decrease in CFI equal to or less than 0.01 (Cheung & Rensvold, 2002) and an increase in RMSEA of less than .015 (Rutkowski & Svetina, 2014).

Step 3 tested whether the PGS measure is used in the same way by participants across the three groups (i.e., that participants who obtain the same PGS score would obtain the same score on the items of the PGS irrespective of which group they belonged). If this
Table 1
Standardised Item Loadings for the Positive Gains Scale within the three comparison groups.

<table>
<thead>
<tr>
<th>PGS Items</th>
<th>Developmental Disability</th>
<th>Chronic Physical Health Conditions</th>
<th>Typically Developing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Since having this child I feel I have grown as a person</td>
<td>.77</td>
<td>.60</td>
<td>.66</td>
</tr>
<tr>
<td>2. Having this child has helped me learn new things/skills</td>
<td>.74</td>
<td>.53</td>
<td>.64</td>
</tr>
<tr>
<td>3. Raising this child helps putting life into perspective</td>
<td>.75</td>
<td>.57</td>
<td>.46</td>
</tr>
<tr>
<td>4. Since having this child, my family has become closer to one another</td>
<td>.42</td>
<td>.52</td>
<td>.50</td>
</tr>
<tr>
<td>5. Since having this child, my family has become more tolerant and accepting</td>
<td>.37</td>
<td>.55</td>
<td>.53</td>
</tr>
<tr>
<td>6. Since having this child I have become more determined to face up to colleagues</td>
<td>.65</td>
<td>.67</td>
<td>.59</td>
</tr>
<tr>
<td>7. Since having this child I have a greater understanding of other people</td>
<td>.59</td>
<td>.73</td>
<td>.67</td>
</tr>
</tbody>
</table>

Note. All p’s < .05 for factor loadings.

If full invariance cannot be demonstrated at the metric, scalar, or full invariance stage, the next step involves testing whether partial invariance is a plausible assumption (Byrne, Shavelson, & Muthen, 1989). Partial invariance is a more relaxed assumption than full invariance, as some of the factor loadings (partial metric invariance) and intercepts (partial scalar invariance) are allowed to differ across groups. Byrne et al. (1989) argued that full metric invariance is not necessary to continue further tests of invariance provided that at least one item is metrically invariant. If partial metric invariance is achieved, partial scalar invariance can be tested for. It was further proposed that if there were at least two factor loadings and intercepts equal across groups, valid inferences regarding group mean differences can be made (Byrne et al., 1989).

3. Results

3.1. Dimensional structure of the PGS within each subgroup

Confirmatory factor analyses (CFA) were conducted separately for each of the three groups of mothers to establish a baseline model. As shown in Table 1, results from the CFA indicated adequate fit for the spina bifida/hydrocephalus group (χ²(14) = 23.59, p = .05; CFI = .965; RMSEAw = .054, 90% CI [.000, .091]) and typical development group (χ²(14) = 28.56, p = .01; CFI = .923; RMSEAw = .081, 90% CI [.037, .124]), but poor fit for the developmental disabilities group (χ²(14) = 302.89, p < .001; CFI = .839; RMSEAw = .133, 90% CI [.120, .146]). We examined whether model fit could improve by reviewing the CFA specification: we examined bivariate correlations between PGS items in each group separately to check whether any PGS items were significantly correlated with one another. Family-level PGS items 4 and 5 correlated strongly (r = .53) within each group, therefore their error terms were allowed to correlate in the baseline models. Correlating error terms is not recommended (Gerbing & Anderson, 1984), as it produces multidimensional factor scores that can be difficult to interpret. However, in this case, the content similarity of Items 4 and 5 indicated that the items represented an alternative to a very similar theme and thus the covariance was justifi ed. CFAs with the correlated error terms demonstrated an improved fit when compared to a model without this error covariance (Δ χ²(1) > 3.84, p < .05 for all 3 groups). When including the error covariance, improved fit for the developmental disabilities group (χ²(13) = 58.48, p < .001; CFI = .975; RMSEAw = .055, 90% CI [.041, .070]), spina bifida/hydrocephalus group (χ²(13) = 7.50, p = .87; CFI = 1.000; RMSEAw = .000, 90% CI [.000, .033]) and the typical development group (χ²(13) = 19.29, p = .11; CFI = .967; RMSEAw = .056, 90% CI [.000, .104]) was shown. Factor loadings across the three groups ranged from .38 to .75 (all p’s < .001) and are shown in Table 1; with the respective model fit indices for each group shown in Table 2.

3.2. Configural invariance testing

We then moved to multigroup CFA (MGCFA) to cross-validate the baseline model structure across the three groups simultaneously. Results indicated a good fit to data (χ²(39) = 88.63, p < .001; CFI = .979; RMSEAw = .050, 90% CI [.036, .063]). Thus, configural invariance was demonstrated, indicating that the factorial structure of the PGS was the same for all three groups (Table 3).

3.3. Metric invariance testing

When equality constraints were imposed on all factor loadings in the MGCFA, good fit was shown for the metric invariance model (χ²(53) = 127.57, p < .001; CFI = .969; RMSEAw = .052, 90% CI [.041, .064]), and there were no significant decreases in the CFI.
loadings for items 1 and 7; as suggested by the modi
be meaningfully compared between groups. As such, the stricter form of measurement invariance, full uniqueness invariance, was not
Therefore, scalar invariance was not demonstrated, and only full metric invariance achieved, indicating that PGS mean scores cannot
3.4. Scalar invariance testing

Our fully metric invariant model observed in the previous step was carried forward as our default model to test for scalar invariance. Initially, constraints were released on all intercepts and loadings. Comparison of model fit revealed a significant $\chi^2$ difference ($\Delta \chi^2(28) = 441.16; p < .001$), a $\Delta$ CFI greater than .01 ($\Delta$ CFI = -.169) and an increase in RMSEA ($\Delta$ RMSEA = .065), suggesting that the fit was significantly worse than that of the configural model. Subsequently, each item intercept was examined for group invariance. We found significant $\chi^2$ differences and a $\Delta$ CFI greater than .01 for all combinations of intercept constraints. Therefore, scalar invariance was not demonstrated, and only full metric invariance achieved, indicating that PGS mean scores cannot be meaningfully compared between groups. As such, the stricter form of measurement invariance, full uniqueness invariance, was not examined.

3.5. Post-hoc analysis of ASD and DD subgroups

Given differences in family functioning dependent on the presence of ASD (e.g., Griffith et al., 2010), sub-group analyses tested for measurement invariance between the PGS in the sample of mothers of children with ASD (i.e., the aforementioned Petalas et al., 2012 sample, combined with families identified from the 1000 Families Study (Hastings et al., 2020), resulting in an overall N of 702) as compared to the sample of mothers of children with intellectual disability without ASD (i.e., the remaining 1000 Families sample; N = 405).

The overall configural invariance model for the two groups (with error covariation between items 4 and 5 included), showed good fit to the data ($\chi^2(26) = 56.99, p < .001$; CFI = .982; RMSEA = .046, 90% CI [.030, .063]). When constraints were placed on factor loadings, model fit did not significantly worsen ($\Delta \chi^2(7) = 36.14, p < .001$; $\Delta$ CFI = -.002, $\Delta$ RMSEA = -.002), therefore indicating full metric invariance. However, when constraints were released on all intercepts and loadings (to test scalar invariance), model fit was significantly poorer ($\Delta \chi^2(14) = 202.17, p < .001$; $\Delta$ CFI = -.084, $\Delta$ RMSEA = .042). Releasing constraints on loadings for items 1 and 7; as suggested by the modification indices, also indicated model non-invariance relative to the configural invariance model ($\Delta \chi^2(12) = 171.94, p < .001$; $\Delta$ CFI = -.069, $\Delta$ RMSEA = .057). As such, only full metric invariance was achieved, and the strictest criteria of full uniqueness invariance were not examined. Accordingly, whilst item loadings for the two groups on the PGS measure appear comparable, the extent to which groups can be compared is limited by a lack of similarity in the groups’ item loadings.

4. Discussion

Our findings suggest that PGS works well to measure positive gain in single groups of parents of children with similar disabilities
or no disabilities. Indeed, excellent internal consistencies were found for the PGS measure in all three groups and results provide strong support for the proposed PGS factor structure across all three groups. However, the utility of the PGS in case-control studies is limited, because although some of its items functioned comparably across groups, PGS total scores did not. This was even the case within the group of mothers of children with developmental disabilities, as PGS scores varied between the group of participants with ASD (with or without intellectual disability) compared to those with intellectual disability without ASD. It should be noted that although PGS mean scores may not be validly compared across groups, comparing scores of equivalent items may still offer useful insight into potential group differences and could be adopted where item-level comparisons are of interest.

It is proposed that items may be differentially interpreted between the three groups because the child rearing experiences of the mothers in the three groups may differ in a way that is directly related to the nature of their child’s disability or lack of disability. These differing experiences may also be associated with different perceptions about positive gains. For example, in addition to gender differences (a high percentage of boys in the developmental disabilities group), children between the groups are likely to have different communication skills, which have been shown to relate to maternal positive perceptions (Griffith, Hastings, Nash, & Hill, 2010). Further, the concept of positive gain may be related to the differential experiences of raising a child with specific needs, which could pose a challenge or threat to parents, whereby parents experiencing stressful caregiving situations may appraise ordinary events as positive to offset the negative consequences of possible adverse events (e.g., Folkan & Moskowitz, 2000). Such experiences may be more alike within disability groups than between disability groups. To examine this further would require some detailed qualitative research to understand the dimensions of mothers’ positive perceptions around child rearing and their relationship with their child with or without a disability.

One potential caveat of our findings, is that differences were noted between the three groups in terms of gender balance. The percentage of males in the developmental disabilities group was somewhat higher than the other two groups; it is suggested that the group differences observed are in accordance with the higher rates of prevalence seen for developmental disabilities in males as opposed to females (e.g., Maenner et al., 2016).

Although this study focused on measurement invariance for the PGS, our findings may have wider implications pertaining to the validity of case-control studies in developmental disabilities research. Many case-control studies within the disability literature utilise instruments that measure parental well-being, making comparisons between those raising children with and without developmental disabilities (Eisenhower, Baker, & Blacher, 2005; Estes et al., 2013; Totsika et al., 2011). However, these comparisons are often made in the absence of a definitive knowledge as to whether that instrument is indeed invariant across these groups. Indeed, within the extant disability literature, support for measurement invariance between developmental disabilities and typically developing or other disability groups is not always found (Ferro & Boyle, 2013; MacLean et al., 2011; Maiano et al., 2011) thus suggesting that subsequent group comparisons are not meaningful (Guenole & Brown, 2014). Our findings also demonstrated that measurement invariance in relation to parents’ perceived positive gain from raising children with different disabilities (even different developmental disabilities) cannot be assumed, further highlighting the need to examine the invariance for any measure used in case-control studies.

Limitations of the current study should also be considered when interpreting the results. The PGS was originally developed for parents of children with physical disabilities. For this reason, trying to demonstrate the measure’s invariance in a group of mothers of children without disabilities might not be appropriate. The underlying construct of positive gain may in fact be very different for parents raising a child without any disability, and more similar for parents raising a child with different disabilities. In addition, the concept of positive gain may also be different among developmental disability diagnoses that were grouped together in this study (i.e., intellectual disability and ASD), and this was not tested. These limitations may lend some explanation as to why scalar invariance could not be established across all three groups. This is something to explore in future research.

In conclusion, the PGS has strong psychometric properties, as evidenced by the consistently high Cronbach’s alpha achieved across a range of populations, and further supported by the analyses in this study. However, it is recommended that researchers do not compare mean PGS scores between groups involving children with developmental disabilities and typical development groups, or children with physical disabilities without testing invariance within their own samples, as measurement invariance could not be demonstrated in the present study.

References


