Intellectual and Developmental Disabilities

Evaluating Measurement Invariance Over Time of a Personal Opportunities Scale for People with IDD

--Manuscript Draft--

Manuscript Number:	IDD-D-23-00057R2
Article Type:	Research
Keywords:	Intellectual and Developmental Disabilities COVID-19 Personal Outcomes Measurement Invariance
Corresponding Author:	Sarah Nelson Lineberry, PhD Partnership for People with Disabilities Richmond, Virginia UNITED STATES
First Author:	Sarah Nelson Lineberry, PhD
Order of Authors:	Sarah Nelson Lineberry, PhD
	Michael Broda, PhD
	Parthenia Dinora, PhD
	Matthew Bogenschutz, PhD
	Seb Prohn, PhD
	Angela West, MS
Manuscript Region of Origin:	UNITED STATES
Abstract:	Personal opportunities refer to chances for people with intellectual and developmental disabilities (IDD) to take self-directed action based on their interests, strengths, and preferences. This study tested for measurement invariance across five years of data collection to determine whether the scale performed consistently over time. Analysis revealed significant differences in both the National Core Indicators In-Person Survey (NCI-IPS) outcomes and in the Personal Opportunities scale. Measurement invariance testing indicated partial threshold and loading invariance, but not intercept invariance, suggesting that the ways in which participants perceived or responded to scale items changed over time. We recommend that researchers utilizing scaled measures with longitudinal outcomes employ statistical checks, including measurement invariance, to ensure the scale performs consistently over time.

Evaluating Measurement Invariance Over Time of a Personal Opportunities Scale for People with IDD

Abstract

Personal opportunities refer to chances for people with intellectual and developmental disabilities (IDD) to take self-directed action based on their interests, strengths, and preferences**.** This study tested for measurement invariance across five years of cross-sectional data, including data collected during the COVID-19 pandemic, to determine whether the scale performed consistently over time. Analysis revealed significant differences in both the National Core Indicators In-Person Survey (NCI-IPS) outcomes and in the Personal Opportunities scale. Measurement invariance testing indicated partial threshold and loading invariance, but not intercept invariance, suggesting that the ways in which participants perceived or responded to scale items changed over time. We recommend that researchers utilizing scaled measures with longitudinal outcomes employ statistical checks, including measurement invariance, to ensure the scale performs consistently over time.

Keywords: Keywords: Intellectual and developmental disabilities, COVID-19, personal opportunities, measurement invariance

Introduction

The need for stable personal outcomes measures over time for people with intellectual and developmental disabilities (IDD), particularly those using Medicaid-funded services, predates the COVID-19 pandemic (Authors, 2020a; Authors 2020b; Shogren et al., 2015). Such data can be used to monitor service delivery, understand how investments in services affect life outcomes, and inform policy and regulatory decisions based on changes and trends over time (Authors, 2020a; Authors 2020b; Shogren et al., 2015). However, given the enormous impact of COVID-19 on people with IDD and the service delivery system, it is incumbent on researchers to investigate if and how the pandemic impacted Home and Community Based Services (HCBS) outcome measurement systems.

Significant historical events, like the COVID-19 pandemic, present a number of challenges to the internal and external validity of research (Butler et al., 2022; Mara & Peugh, 2020). For example, disruptions due to COVID-19 may have impacted people’s ability to participate in research, introducing the possibility of selection bias to the research design (Butler et al., 2022). The National Core Indicators In-Person Survey (described in more detail in the Methods section; NCI-IPS) shifted to virtual data collection due to COVID-19, potentially excluding people without reliable internet connections or support to use video conferencing platforms (Butler et al., 2022; Mara & Peugh, 2020; National Core Indicators, 2020). Additionally, collecting data during a significant historical event introduces new covariates that can independently affect outcomes, making it more difficult to identify relationships of interest (Butler et al., 2022). Butler and colleagues (2022) suggest that attending to these threats is particularly important when researching populations disproportionately impacted by the COVID-19 pandemic.

Measurement Invariance as a Tool to Assess Scale Performance

Testing for measurement invariance (MI) across time is necessary when looking at responses collected over periods of change, particularly in the context of major historical events like COVID-19. Establishing MI ensures that the measure maintains consistent meaning and interpretation over time, allowing researchers to assess changes accurately and make meaningful comparisons among different time points (Widaman et al., 2010). Failing to test for MI may lead to incorrect conclusions about the stability or change in outcomes over time, which could misinform policymakers and practitioners in their decision-making processes.

Furthermore, examining MI is necessary to determine whether the scale consistently measures the same underlying construct across time (Meredith, 1993). Demonstrating MI bolsters the scale's validity and strengthens the theoretical foundation for interpreting the findings (Vandenberg & Lance, 2000). In addition, investigating MI can reveal potential biases or differences in how individuals perceive or respond to the scale items over time, providing valuable insights into the scale's psychometric properties and informing future modifications or improvements.

Personal Opportunities

The Personal Opportunities scale was developed as a result of previous work by the authors (Authors, 2020b) using large administrative datasets merged at the individual level to better understand the relationship between IDD (HCBS) Medicaid expenditures, a person’s support needs, personal characteristics, and personal outcomes. This project began in response to an identified need in the IDD field for measures that accounted for differences in individual (support needs, disability, additional diagnoses, etc.) and systemic (services, residential setting, expenditures, etc.) characteristics (Kaye & Harrington, 2015).

One of the primary datasets used in these analyses, the NCI-IPS, contains a number of variables related to broad domains including community participation, rights, and choice. While these indicators are useful, prior measures developed using national data could not be replicated using state samples (Jones et al., 2018). Following recommendations from the National Quality Forum (2016) on the need for standardized scales using existing measures, authors (2020b) sought to develop new scales that could be validated with both state and national NCI-IPS data across multiple years of data collection.

Personal Opportunities refer to chances for people with IDD to take self-directed action based on their interests, strengths, and preferences (Authors, 2020a; Authors, 2022). This scale builds on previous measures using the NCI-IPS including Everyday Choice (Lakin et al., 2008), Social Participation and Relationships (Mehling & Tasse, 2014), and Home Privacy (Houseworth et al., 2019) as well as with subdomains identified by the National Quality Forum (2016). Based on these existing frameworks, Authors (2020a) used three years of NCI-IPS data (2017-2018, 2019-2019, and 2019-2020) from [state name redacted for peer review] to identify three domains of personal opportunities–rights, everyday choice, and community participation–and then validated the measure with the national NCI-IPS data from 2017-2018 (Authors, 2022). These domains, variables, and definitions from the NCI-IPS are presented in Table 1. In this study, we will use the Personal Opportunities scale to illustrate the importance of testing for MI over time, particularly in the context of the COVID-19 pandemic.

Impact of COVID on Personal Opportunities for People with IDD

While this study is predominantly concerned with how COVID-19 impacted the Personal Opportunities scale, it is also important to consider the experiences of people with IDD during the pandemic, who may have been particularly vulnerable to the COVID-19 pandemic and associated closures (Friedman, 2021; Lake et al., 2021; Pfeiffer et al., 2021; Rosencrans et al., 2021). In one web-based survey of people with IDD, 92.8% of respondents reported negative impacts from the COVID-19 pandemic (Fisher et al., 2021). People with IDD experienced changes to or loss of employment (Fisher et al., 2021) or disability services (Friedman, 2021; Linehan et al., 2022; Rosencrans et al., 2021), Linehan et al., 2022), and fewer social activities, particularly in person (Pfeiffer et al., 2022). These disruptions may have contributed to increased anxiety, worry, or mental health symptoms in people with IDD (Lake et al., 2021).

Importantly, IDD refers to a wide range of conditions and support needs (Schalock et al., 2019). Before the pandemic, authors (Authors, 2023) found that people with higher support needs had lower personal opportunities. Rosencrans and colleagues (2021) suggest that this pattern may hold for pandemic-related disruptions, with people who reported more functional limitations also reported experiencing more mental health symptoms since the start of COVID-19. Additionally, disruptions to daily routines and opportunities for community engagement may have been particularly challenging for people with more significant intellectual disabilities who had difficulties understanding public health interventions or their purpose (Lake et al., 2021)

These findings are also reflected in other secondary data analyses. In her analysis of the Personal Outcomes Measure (POM), Friedman (2021) found that respondents were less likely to participate in their community and have intimate relationships in 2020 compared to 2019. The POM also includes questions about participants' abilities to exercise their rights, make choices about their work and services, and have friendships, but these outcomes were not significantly different between 2019 and 2020 (Friedman, 2021).

Many of these changes may be closely related to the previously validated Personal Opportunities scale. However, before claims can be made about changes in personal opportunity outcomes during and after the COVID-19 pandemic, it is important to establish the internal validity of the scale itself across years. Specifically, are respondents interpreting items in the same way before, during, and after the pandemic?

Study Aims

This paper seeks to model the utility of testing for MI across time, particularly in the context of major historical events, to ensure accurate assessment, inform evidence-based policies and interventions, and enhance the scale's validity and reliability. Using the Personal Opportunities scale as a case example, we pursued the following research questions:

1) What is the optimal factor structure for a scale measuring personal opportunities for individuals with IDD using items from the NCI-IPS?

2) Given the optimal factor structure, does the Personal Opportunities scale perform similarly in theoretically similar cohorts sampled across five years from 2017-2022?

3) To what extent do responses collected in 2020-2021 function as an outlier in relation to the pattern of responses collected in the three years prior and one year after?

Methods

Procedures used to conduct this study were approved by the Institutional Review Board at the authors' university.

Data

Data for this study came from [state name redacted for peer review]'s National Core Indicators In-Person Survey (NCI-IPS). The NCI-IPS is a collaborative project that includes the Human Service Research Institute (HSRI), the National Association of State Directors of Developmental Disabilities Services (NASDDDS), and participating states and aims to collect data about public IDD systems and service users' outcomes. These data can be used to track performance and outcome measures over time, compare results across states, and establish national quality benchmarks (National Core Indicators, n.d.).

The NCI-IPS is an interview with adults (18 years or older) who use at least one state-funded disability service in addition to case management. Each year, participants are randomly selected from all HCBS users of the state's Developmental Disability (DD) waiver. For this study, we specifically used data from [state name redacted for peer review] from five years of data collection (2017-2018 through 2021-2022) for five cohorts of respondents. Data collection typically runs from July of one year until June of the following year; all data from 2019-2020 was collected prior to the start of the COVID-19 pandemic.

The NCI-IPS consists of three sections. The background section is completed using case files before the interview, usually by a case manager. It includes demographic information, details about the participant's disability diagnosis and/or health conditions, and basic information about their DD services. Section I contains subjective questions about the participant's life and services and may only be answered directly by the individual. Section II contains objective questions about outcomes in key life domains including community participation, rights, and personal choices which may be answered either by the individual or by someone who knows them well. Variables for this study were drawn from the background section and Section II of the NCI-IPS.

Variables

A description of the variables used in this analysis can be found in Table 1. A full description of the methods used to develop the Personal Opportunity measures can be found elsewhere (Authors, 2022). The measure was initially developed by (Authors, 2020a) to expand upon previously developed measures of Everyday Choice (Lakin et al., 2008), Social

Participation and Relationships (Mehling & Tasse, 2014), and Home Privacy (Houseworth et al., 2019) using four years of Virginia NCI-IPS data and was then validated by (Authors, 2022) on the national NCI-IPS data. The measure defines Personal Opportunities as “chances for self-directed action," specifically to make choices, participate in the community, and have privacy (Authors, 2020a). Authors (2022) included a fourth factor, expanded friendships. This factor could not be estimated in this study because the items were not collected in all five years of data collection.

Data Analysis

The study's primary research objectives were addressed through three sequential stages of data analysis. The first phase entailed a single-group confirmatory factor analysis (CFA) to assess the suitability of our previously developed personal opportunities scale on this larger sample. Next, a multigroup CFA model was utilized to explore MI between samples obtained across five years from 2017 to 2022. Given that all of our CFA model's indicators were categorical, we used the procedure for testing measurement equivalence/invariance described in Svetina et al. (2020). This included testing for MI using a series of nested models: 1) a multigroup model where all parameters are estimated separately in each group, 2) a model where thresholds are constrained to be equal across groups, 3) a model where thresholds and item loadings are constrained to be equal, and 4) a final model where thresholds, loadings, and intercepts are all constrained to be equal. In the final stage, we repeated the multigroup CFA, this time removing observations from 2020-2021 to test the premise that data collected amidst the widespread, acute impacts of COVID-19 differed from those collected before 2020 and from later stages of the pandemic, as restrictions loosened. The goal was to confirm whether measurement properties held for all years with the exception of 2020-2021 to guide future use of scales. For example, even though 2021-2022 was undoubtedly impacted by the longer-term effects of the pandemic, we were interested in whether the psychometric properties of our scale were closer to the patterns of responses observed pre-COVID, as the acute effects of the pandemic began to wane.

The statistical software Mplus, Version 7.11 (Muthén & Muthén, 2015) was used for data analyses. The robust maximum likelihood estimator (MLR) was employed to evaluate all the CFA models and conduct invariance testing. This estimation method performs the full information maximum likelihood (FIML) procedure when dealing with missing data, as well as providing robust standard errors.

In order to examine patterns of missing data, Little's (1988) MCAR test was used to determine if the data could be treated as missing completely at random (MCAR). The findings suggested that interpreting the data as MCAR would not be appropriate and more robust estimation methods for handling missing values, such as full information maximum likelihood (FIML) were required. Using FIML, the estimation process can incorporate all available data, yielding results comparable to those achieved through traditional missing data techniques like multiple imputation (Allison, 2001).

Data Assumptions and Model Fit/Parsimony

Before conducting the aforementioned data analysis, we assessed the assumption of multivariate normality via the “mvtest” function in Stata. The results of both Mardia's (1970) test and Doornik-Hansen's (2008) omnibus test yielded statistically significant outcomes, indicating that the data did not completely meet the multivariate normal assumption that is typically required in latent variable modeling. In order to address this potential violation, we implemented the robust maximum likelihood procedure (MLR in MPlus) and assessed model fit for all confirmatory factor analysis (CFA) and structural equation modeling (SEM) models using the Satorra-Bentler scaled chi-square (S-B $\chi 2$; Satorra & Bentler, 2001). These methods are recommended for situations where the data may not conform to the multivariate normal assumption, and have demonstrated robustness even in the presence of more substantial deviations from normality (Enders, 2001).

Moreover, we employed several other fit indices to evaluate the goodness-of-fit of our models, including the comparative fit index (CFI), standardized root-mean-square residual (SRMR), and root-mean-square error of approximation (RMSEA) with a 90% confidence interval. To appraise the adequacy of the model fit, we utilized the assessment criteria developed by Hu and Bentler (1999), which stipulate that an acceptable model fit entails a CFI value greater than or equal to .95, an SRMR value less than or equal to .08, and an RMSEA value less than or equal to .06. In cases where the CFI value fell short of the suggested threshold, we then consulted Hu and Bentler's (1999) recommendation to evaluate both SRMR and RMSEA simultaneously, with an RMSEA value less than or equal to .06 and an SRMR value less than or equal to .10 indicating an acceptable model fit.

Results

Participants

The sample size varied across the years, ranging from 512 in 2019-2020 to 842 in 2017-2018, with a total of 3,595 participants across the five years of data collection. Participants were mostly White (62.4%) and Black (31.5%) with smaller numbers of participants identifying as Latino, Asian, Pacific Islander, American Indian, or a different racial identity. Slightly more than half of participants identified as male (59.3%). “Other” was offered as an option for gender in some years of data collection, but was only selected by three participants. About half of the participants were their own guardian (51.43%) and most participants had a mild (30.44%) or moderate (38.58%) intellectual disability. The only significant differences across years were found among respondents with profound intellectual disability (X^2 (12) = 24.11, p = .20), probably due to the low level of respondents with more significant levels of intellectual disabilities (9.75% of the total sample).

Personal Opportunities

Table 2 presents a descriptive analysis of the items that comprise our Personal Opportunities scale measured across five years in a sample of individuals with IDD. The results indicate a shift in response patterns across time for several items. For instance, the proportion of participants who reported "yes/maybe" for having a key to their home increased from 50.7% in 2017-2018 to 69.5% in 2021-2022. Similarly, the percentage of participants reporting "yes/maybe" for having the ability to lock their bedroom increased from 58.5% in 2017-2018 to 74.1% in 2021-2022.

The frequency of engaging in various activities also changed over the years. For example, the percentage of participants who reported going shopping more than five times in a month decreased from 43.2% in 2017-2018 to 41.4% in 2021-2022. However, the proportion of participants reporting no entertainment experiences increased from 16.2% in 2017-2018 to 43.2% in 2021-2022. The results also showed variations in the frequency of eating out and running errands and changes in the extent to which individuals had control over their daily schedules, what to do in free time, and purchasing choices.

Confirming the Factor Structure of Personal Opportunities

After exploring descriptive patterns in the items over time, we used confirmatory factor analysis to investigate the underlying factor structure of the personal opportunities scale. Two separate confirmatory factor analysis models were tested to compare the relative fit of a one-factor versus three-factor model of personal opportunities. The three-factor model demonstrated the best model fit and parsimony according to the guidelines suggested by Hu and Bentler (1999), Satorra-Bentler χ^2 (24) = 183.74, p < .001, RMSEA = .046 (90% CI [.040-.053]), CFI = 0.979, and TLI = .968. The one-factor model demonstrated significantly poorer fit and parsimony, Satorra-Bentler χ^2 (27) = 2629.84, p < .001, RMSEA = .176 (90% CI [.170-.182]), CFI = 0.652, and TLI = .536. In both cases, the chi-square test of model fit was significant, though this test is highly sensitive to small model-data discrepancies in large samples (Kline, 2016). Therefore, the three-factor model for personal opportunities was a more accurate representation of the underlying construct than the one-factor model.

Assessing the Longitudinal Measurement Invariance of Personal Opportunities Scale

Having confirmed the appropriate structure for the Personal Opportunities Scale, we then conducted a series of measurement invariance tests to investigate the extent to which the scale operated equivalently across the five years of data collection reflected in the sample.

Table 3, top section, presents a comparison of four nested models testing different levels of measurement invariance in a longitudinal study. The models include the original 5-group model, and three more restrictive models with invariant thresholds (T), loadings (L), and intercepts (I). The original 5-group model shows an acceptable fit, with a significant Chi-square value, a CFI of 0.97, a TLI of 0.96, an RMSEA of 0.05 with a 90% confidence interval between 0.04 and 0.05, and an SRMR of 0.05. The threshold invariance, or “T Invariant” model exhibits a similar fit to the original model, with comparable CFI, TLI, RMSEA, and SRMR values. The threshold and loading, or “TL Invariant” model also demonstrates an acceptable fit, although with slightly lower CFI (0.96) and TLI (0.95) and a higher RMSEA (0.05) with a 90% confidence interval between 0.05 and 0.06. However, the threshold, loading, and intercept, or “TLI Invariant” model shows a poor fit, with a much lower CFI (0.59) and TLI (0.63), a substantially higher RMSEA (0.15), and a higher SRMR (0.07).

Table 3, bottom section, displays the differences in key fit statistics, making it easier to examine sequential differences in model fit moving from less to more restrictive models. Criteria from Chen (2007) and Cheung and Rensvold (2002) were used to evaluate changes in model fit, with changes in both RMSEA (ΔRMSEA $\leq$ .015) and CFI (ΔCFI $\leq$ .01) required to demonstrate full invariance. Applying these criteria, we would argue that the Personal Opportunities scale demonstrates partial threshold and loading invariance, but does not satisfy the requirements for threshold, loading, and intercept invariance. It is important to note that only partial invariance could be established, largely because models that attempted to constrain loadings and intercepts for the “key” and “lock” items would not converge. Once these two items were allowed to be freely estimated in the COVID-19 year (2020-2021), the models would successfully converge. The performance of the key to one’s home and ability to lock one’s bedroom door items in the NCI cohort most affected by COVID-19 disruptions (2020-2021) prompted further exploration, specifically, whether longitudinal invariance could be established for all years other than 2020-2021. Thus, we removed the data from 2020-2021 and proceeded with repeating the MI analysis as described above.

Assessing Longitudinal MI Without COVID-19 Year (2021)

Table 4, top section presents a comparison of four nested models testing different levels of MI, excluding the COVID-19 year (2020-2021). Just as seen in Table 3, the models included the original 5-group model, and three more restrictive models with invariant thresholds (T), loadings (L), and intercepts (I).

The original 5-group model showed a good fit, with high CFI and TLI (both 0.98), a small RMSEA (0.03) with 90% confidence interval between 0.02 and 0.04, and a low SRMR (0.04). The T Invariant and TL Invariant models also show good fit indices, with similar CFI, TLI, RMSEA, and SRMR values as the original model. Similar to the previous set of analyses, the TLI Invariant model demonstrates somewhat poorer fit, indicated by a lower CFI (0.89) and TLI (0.90), a higher RMSEA (0.07), and a slightly higher SRMR (0.05). Of note, however, the decrease in fit between the TL and TLI models was significantly less once the responses from the COVID-19 Year (2020-2021) were removed.

Table 4, bottom section presents the differences in fit statistics between the longitudinal measurement invariance models with No COVID-19 Year (2020-2021) included. Comparing the T Invariant model to the original model, there were no significant changes in any of the fit indices (CFI, TLI, RMSEA, 90% LB, 90% UB, and SRMR). Similarly, when comparing the TL Invariant model to the T Invariant model, no significant changes were observed across all the fit indices. However, when comparing the TLI Invariant model to the TL Invariant model, the fit indices indicated a considerable decrease in model fit with a ΔCFI of -0.09, ΔTLI of -0.08, and an increase in ΔRMSEA of 0.04 along with a change in the lower and upper bounds (Δ90% LB and Δ90% UB) by 0.04. Additionally, the ΔSRMR increased by 0.01.

Discussion

Analyzing a Personal Opportunities scale across five years in a sample of people with IDD revealed shifts in response patterns and engagement in various activities over time. The proportion of participants reporting having a key to their home and ability to lock their bedroom door systematically increased, while the frequency of shopping, entertainment experiences, eating out, and running errands fluctuated over the years. Additionally, changes were observed in the extent to which individuals had control over their daily schedules, free time, and purchasing choices.

Assessing the measurement properties, CFA was used to confirm the factor structure of the personal opportunities scale, with the three-factor model demonstrating the best fit. Longitudinal MI tests were then conducted, indicating partial threshold and loading invariance, but not intercept invariance. The performance of the key and lock items was an important discovery - both items had to be freely estimated during the height of the COVID-19 pandemic (2020-2021) for the models to converge, suggesting a significant shift in the functioning of those items during the 2020-2021 data collection cycle. However, when data from this year were excluded, the decrease in model fit between threshold and loading invariance and threshold, loading, and intercept invariance models was significantly less. These findings highlight the importance of considering the impact of extraordinary circumstances, such as the COVID-19 pandemic, on the measurement properties of psychological constructs such as personal opportunities.

Implications for HCBS Outcome Measurement

While scaled variables offer important opportunities for monitoring outcomes and guiding recommendations for interventions and policy recommendations, testing for MI is necessary to ensure that these scales perform consistently (Meredith, 1993; Vandenberg & Lance, 2000). This study emphasizes the importance of testing for MI over time to understand the impact of major historical events (Mara & Peugh, 2020). Without establishing MI, it is impossible to know if observed changes in outcomes are due to actual changes in participants' lives or to changes in how they interpret and respond to survey items (Mara & Peugh, 2020; Meredith, 1993; Vandenberg & Lance, 2000).

Using the example of the Personal Opportunity scales, testing for MI between years demonstrates that while participants did experience significant differences during the COVID-19 year of data collection (2020-2021), the scale items themselves also behaved differently. Given these results, we recommend that researchers using scaled measurement to evaluate longitudinal outcomes employ statistical checks, such as testing for MI, to ensure the scale maintains integrity, particularly in light of catastrophic events such as the COVID-19 pandemic.

To address any inconsistencies with scale performance, researchers can instead focus on changes in individual survey items rather than employing scales to examine NCI-IPS outcomes over a timeframe that includes the COVID-19 year (2020-2021). If scales are used, we suggest that the COVID-19 year be excluded or that researchers apply appropriate measurement constraints to account for differences in scale performance. Additional research is necessary to disentangle the effects of the COVID-19 pandemic on personal opportunities for people with IDD from its effects on the NCI-IPS. For example, future research may use true longitudinal measures that survey the same people prior to and after the COVID-19 pandemic, as opposed to the random cohorts used in this analysis.

Implications for Policy and Practice

The need for reliable and valid IDD system measures to monitor quality over time so that policymakers and system managers can make informed, data-based policy and regulatory decisions is well documented in the literature (Authors, 2020a, Authors, 2020b; Tichá, 2023). As the IDD field relies more heavily on data to make decisions, researchers must exercise due diligence to ensure that measurement is reliable and valid over time. Major national and state-based policy changes need to be effectively evaluated through the lens of how they impact the lives and outcomes of people who used Medicaid IDD HCBS over time. Furthermore, given the extraordinary challenges that the COVID-19 pandemic presented to people with IDD and service systems (Fisher et al., 2022; Friedman et al., 2021; Rosencrans et al., 2021), the validation of scales enables the field to better understand the true impact of catastrophic events and to prepare for future challenges.

Variables related to physical privacy (having a key to one's home and the ability to lock one's bedroom door) are particularly interesting in the context of a pandemic and associated public health response. Rates of having a key and being able to lock one's door began to increase in the 2018-2019 data collection cycle, possibly due to increased attention to privacy in response to the CMS Final Settings Rule, finalized in 2014 (CMS, 2014). However, these variables did not fit a previously-established model of personal opportunities during the 2020-2021 data collection cycle–a time of massive disruption to routines, including more time spent in the home (Pfieffer et al., 2022). Beyond changes in the rates of people able to exercise their right to privacy, the poor fit of these variables suggests a substantial change in the way people conceptualized privacy during the COVID-19 pandemic. As a whole, these findings illustrate the intersections of high-level policies, global events, and the daily lives of people with IDD, all of which must be considered in any future decisions.

Limitations

As with any study, this one has several important limitations that offer opportunities for future research. Firstly, while this analysis of five years of NCI-IPS data allows us to compare outcomes over time, it is not a true longitudinal analysis. Participants are randomly selected from the state's roster of Medicaid HCBS service users, minimizing potential threats to external validity. However, the possibility of bias cannot be fully dismissed, particularly in years when COVID-19 disrupted data collection (Butler et al., 2022).

There are also limitations associated with the data itself. The NCI-IPS recruits participants from each state's DD Waiver users. Literature suggests that only about 20% of adults with IDD who live in the community use state-funded disability services, limiting the generalizability of studies that use this sampling frame (Larson et al., 2020). Additionally, the NCI-IPS allows for proxy responses to some questions. NCI-IPS surveyors are trained to support people with IDD to answer independently whenever possible and to maximize the validity of proxy responses, but it should be noted that there are concerns with using proxy responses to surveys (Scott & Havercamp, 2018).

We also conducted this analysis with data from one state. This limitation is particularly noteworthy when considering the impacts of a historic event like the COVID-19 pandemic. While the NCI-IPS provides interviewers with standardized training, states manage their own sampling and data collection (National Core Indicators, 2022). It is therefore possible that some states were more successful than others in navigating the shift to administering surveys virtually. It will be important for future research to expand these analyses to include larger cohorts of states and different sets of data to better understand the scope and depth of measurement changes during the COVID-19 pandemic and the implications for outcome measures.

Conclusion

The COVID-19 pandemic and the associated public health response has enormously impacted both Personal Opportunities for people with IDD and how these opportunities are measured. This study demonstrates the importance of investigating the impacts of major historical events on outcomes of interest and measurement strategies. Testing for MI over time is a valuable strategy to ensure that outcomes are measured consistently over time and that resulting policy and practice recommendations are based on sound data.

References

Allison, P. D. (2001). *Missing Data*. SAGE Publications.

Butler, A. M., Burcu, M., Christian, J. B., Tian, F., Andersen, K. M., Blumentals, W. A., Joynt Maddox, K. E., & Alexander, G. C. (2023). Noninterventional studies in the COVID-19 era: methodological considerations for study design and analysis. *Journal of Clinical Epidemiology*, *153*, 91–101. https://doi.org/10.1016/j.jclinepi.2022.11.011

Centers for Medicare and Medicaid Services. (2014). Home and community based services. https://www.medicaid.gov/sites/default/files/2019-12/final-rule-fact-sheet.pdf

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheng, S. L., Prohn, S. M., Dinora, P., Broda, M. D., & Bogenschutz, M. (2020). Measuring and tracking personal opportunity outcome measures over 3 years to guide policy and services that promote inclusive community living. *Inclusion*, *8*(4), 335-350. http://doi.org/ 10.1352/2326-6988-8.4.335

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Dinora, P., Bogenschutz, M., & Broda, M. (2020). Identifying predictors for enhanced outcomes for people with intellectual and developmental disabilities. *Intellectual and Developmental Disabilities*, *58*(2), 139-157. 10.1352/1934-9556-58.2.139

Dinora, P., Prohn, S. M., Bogenschutz, M., Broda, M. D., Lineberry, S., & West, A. (2023). An examination of support needs, supports, and outcomes for people with intellectual and developmental disabilities. *Intellectual and Developmental Disabilities*, *61*(1), 65-78. https://doi.org/10.1352/1934-9556-61.1.65

Doornik, J. A., & Hansen, H. (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, *70*, 927–939. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0084.2008.00537.x

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(1), 128–141. https://doi.org/10.1207/S15328007SEM0801_7

Fisher, M. H., Sung, C., Kammes, R. R., Okyere, C., & Park, J. (2022). Social support as a mediator of stress and life satisfaction for people with intellectual or developmental disabilities during the COVID- 19 pandemic. *Journal of Applied Research in Intellectual Disabilities*, *35*(1), 243-251. http://doi.org/10.1111/jar.12943

Friedman, C. (2021). The COVID-19 pandemic and quality of life outcomes of people with intellectual and developmental disabilities. *Disability and Health Journal*, *14*(4), 101117. 10.1016/j.dhjo.2021.101117

Houseworth, J., Stancliffe, R. J., & Ticha´, R. (2019). Examining the National Core Indicators' potential to monitor rights of people with intellectual and developmental disabilities according to the CRPD. *Journal of Policy and Practice in Intellectual Disabilities*, *16*(4), 342–351. https://doi.org/10.1111/jppi.12315

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jones, J. L., Shogren, K. A., Grandfield, E. M., Vierling, K. L., Gallus, K. L., & Shaw, L. A. (2018). Examining predictors of self-determination in adults with intellectual and developmental disabilities. *Journal of Developmental and Physical Disabilities, 30*(5), 601–614. https://doi.org/10.1007/s10882-018-9607-z

Kaye, H. S., & Harrington, C. (2015). Long-term services and supports in the community: Toward a research agenda. *Disability and Health Journal*, 8(1), 3–8. http://dx.doi.org/10.1016/j.dhjo.2014.09.003

Lake, J. K., Jachyra, P., Volpe, T., Lunsky, Y., Magnacca, C., Marcinkiewicz, A., & Hamdani, Y. (2021). The wellbeing and mental health care experiences of adults with intellectual and developmental disabilities during COVID-19. *Journal of Mental Health Research in Intellectual Disabilities*, *14*(3), 285-300. https://doi.org/10.1080/19315864.2021.1892890

Lakin, K. C., Doljanac, R., Byun, S.-Y., Stancliffe, R., Taub, S., & Chiri, G. (2008). Choicemaking among Medicaid HCBS and ICF/MR recipients in six states. *American Journal on Mental Retardation*, *113*(5), 325–342. https://doi.org/10.1352/2008.113.325-342

Larson, S.A., Eschenbacher, H.J., Taylor, B., Pettingell, S., Sowers, M., & Bourne, M.L. (2020). In-home and residential long-term supports and services for persons with intellectual or developmental disabilities: Status and trends through 2017. *Minneapolis: University of Minnesota, Research and Training Center on Community Living, Institute on Community Integration.* https://conservancy.umn.edu/bitstream/handle/11299/203355/Report_2012_InHome_and_Residential.pdf?sequence=1

Little, R. J. A. (1988). A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Mara, C. A., & Peugh, J. L. (2020). Validity of data collected from randomized behavioral clinical trials during the COVID-19 pandemic. *Journal of Pediatric Psychology*, *45*(9), 971-976. http://doi.org/10.1093/jpepsy/jsaa078

Mardia, K. V. (1974). Applications of Some Measures of Multivariate Skewness and Kurtosis in Testing Normality and Robustness Studies. *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, *36*(2), 115–128. http://www.jstor.org/stable/25051892

Mehling, M. H., & Tasse´, M. J. (2014). Empirically derived model of social outcomes and predictors for adults with ASD. *Intellectual and Developmental Disabilities*, *52*(4), 282–295. https://doi.org/10.1352/1934-9556-52.4.282

National Core Indicators. (n.d.). https://idd.nationalcoreindicators.org/

National Core Indicators. (2020). NCI remote survey pilot study summary results, 2019-20. https://legacy.nationalcoreindicators.org/upload/core-indicators/NCI_REMOTE_BRIEF_REPORT_201222_accessible_FINAL_2.pdf

National Core Indicators. (2022). How it works. https://idd.nationalcoreindicators.org/how-it-works/

National Quality Forum. (2016). *Quality in home and community-based services to support community living: Addressing gaps in performance measurement.* https://www.qualityforum.org/ Publications/2016/09/Quality_in_Home_ and_Community-Based_Services_to_ Support_Community_Living__Addressing_ Gaps_in_Performance_Measurement.aspx

Pfeiffer, B., Brusilovskiy, E., Hallock, T., Salzer, M., Davidson, A. P., Slugg, L., & Feeley, C. (2022). Impact of COVID-19 on community participation and mobility in young adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *52*(4), 1553-1567. https://doi.org/10.1007/s10803-021-05054-0

Prohn S.M., Dinora P., Broda M.D., Bogenschutz M., Lineberry S. (2022) Measuring four personal opportunities for adults with intellectual and developmental disabilities. *Inclusion. 10*(1):19-34. doi: 10.1352/2326-6988-10.1.19. PMID: 35721258; PMCID: PMC9201575.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-543.

Rosencrans, M., Arango, P., Sabat, C., Buck, A., Brown, C., Tenorio, M., & Witwer, A. (2021). The impact of the COVID-19 pandemic on the health, wellbeing, and access to services of people with intellectual and developmental disabilities. *Research in Developmental Disabilities*, *114*, 103985. https://doi.org/10.1016/j.ridd.2021.103985

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–514. https://doi.org/10.1007/BF02296192

Schalock, R. L., Luckasson, R., & Tassé, M. J. (2019). The contemporary view of intellectual and developmental disabilities: Implications for psychologists. *Psicothema*. http://doi.org/10.7334/psicothema2019.119

Scott, H. M., & Havercamp, S. M. (2018). Comparisons of self and proxy report on health-related factors in people with intellectual disability. *Journal of Applied Research in Intellectual Disabilities, 31*(5), 927-936. https://doi-org.proxy.library.vcu.edu/10.1111/jar.12452

Shogren, K. A., Luckasson, R., & Schalock, R. L. (2015). Using context as an integrative framework to align policy goals, supports, and outcomes in intellectual disability. *Intellectual and Developmental Disabilities, 53*(5), 367–376. https://doi.org/10.1352/1934-9556-53.5.367

Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-Group invariance with categorical outcomes using updated guidelines: An illustration using Mplus and the lavaan/semTools packages. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), 111–130. https://doi.org/10.1080/10705511.2019.1602776

Tichá, R., Pettingell, S. L., Lemanowicz, J., Feinstein, C. S., Bershadsky, J., Houseworth, J., & Zhang, A. Q. (2023). Using the independent monitoring for quality program to examine longitudinal outcomes for people with intellectual and developmental disabilities. *Intellectual and Developmental Disabilities, 61*(3), 238-249. https://doi-org.proxy.library.vcu.edu/10.1352/1934-9556-61.3.238

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-70.

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10-18.

Table 1. Personal Opportunity Variables

Variable	Description
Rights Domain	
Key	Do you have a key to your home?
Lock	Can you lock your bedroom if you want to?
Community Participation Domain	
Shopping	How many times did you go shopping in the last month?
Entertain	How many times did you go out for entertainment in the past month?
Eat Out	How many times did you go to a restaurant or coffee shop in the past month?
Errands	How many times did you go out on errands or appointments in the past month?
Choice Domain	
Schedule	Who decides your daily schedule?
Free Time	Who decides how you spend your free time?
Choose Buy	Do you choose what you buy with your spending money?

Table 2

Frequencies for Personal Opportunities Indicators, by Year of Data Collection

	17-18 (N=842)	18-19 (N=807)	19-20 (N=512)	20-21 (N=726)	21-22 (N=708)	Overall (N=3595)	X^2 (df)
Key							76.78 (8)***
0 No	371 (44.1%)	330 (40.9%)	153 (29.9%)	217 (29.9%)	206 (29.1%)	1277 (35.5%)	
2 Yes/Maybe	427 (50.7%)	472 (58.5%)	352 (68.8%)	496 (68.3%)	492 (69.5%)	2239 (62.3%)	
Missing	44 (5.2%)	5 (0.6%)	7 (1.4%)	13 (1.8%)	10 (1.4%)	79 (2.2%)	
Lock							84.62 (8)***
0 No	271 (32.2%)	234 (29.0%)	113 (22.1%)	131 (18.0%)	125 (17.7%)	874 (24.3%)	
2 Yes/Maybe	492 (58.5%)	538 (66.7%)	350 (68.3%)	551 (75.9%)	525 (74.1%)	2456 (68.3%)	
Missing	79 (9.4%)	35 (4.3%)	49 (9.6%)	44 (6.1%)	58 (8.2%)	265 (7.4%)	
Shopping							344.74 (12)***
1 (0 times)	60 (7.1%)	54 (6.7%)	48 (9.4%)	264 (36.4%)	112 (15.8%)	538 (15.0%)	
2	163 (19.4%)	148 (18.3%)	93 (18.2%)	133 (18.3%)	131 (18.5%)	668 (18.6%)	
3	210 (24.9%)	221 (27.4%)	135 (26.4%)	137 (18.9%)	160 (22.6%)	863 (24.0%)	
4 (more than 5 times)	364 (43.2%)	369 (45.7%)	233 (45.5%)	180 (24.8%)	293 (41.4%)	1439 (40.0%)	
Missing	45 (5.3%)	15 (1.9%)	3 (0.6%)	12 (1.7%)	12 (1.7%)	87 (2.4%)	
Entertain							695.73 (12)***
1 (0 times)	136 (16.2%)	149 (18.5%)	134 (26.2%)	526 (72.5%)	306 (43.2%)	1251 (34.8%)	
2	253 (30.0%)	279 (34.6%)	175 (34.2%)	83 (11.4%)	178 (25.1%)	968 (26.9%)	
3	183 (21.7%)	180 (22.3%)	109 (21.3%)	48 (6.6%)	99 (14.0%)	619 (17.2%)	
4 (more than 5 times)	224 (26.6%)	191 (23.7%)	87 (17.0%)	54 (7.4%)	111 (15.7%)	667 (18.6%)	
Missing	46 (5.5%)	8 (1.0%)	7 (1.4%)	15 (2.1%)	14 (2.0%)	90 (2.5%)	
Eat Out							726.73 (12)***
1 (0 times)	75 (8.9%)	84 (10.4%)	73 (14.3%)	433 (59.6%)	188 (26.6%)	853 (23.7%)	
2	213 (25.3%)	174 (21.6%)	114 (22.3%)	127 (17.5%)	182 (25.7%)	810 (22.5%)	
3	199 (23.6%)	216 (26.8%)	126 (24.6%)	88 (12.1%)	135 (19.1%)	764 (21.3%)	
4 (more than 5 times)	318 (37.8%)	325 (40.3%)	195 (38.1%)	67 (9.2%)	187 (26.4%)	1092 (30.4%)	
Missing	37 (4.4%)	8 (1.0%)	4 (0.8%)	11 (1.5%)	16 (2.3%)	76 (2.1%)	

	17-18 (N=842)	18-19 (N=807)	19-20 (N=512)	20-21 (N=726)	21-22 (N=708)	Overall (N=3595)	X^2 (df)
Errands							79.84 (12)***
1 (0 times)	94 (11.2%)	81 (10.0%)	62 (12.1%)	166 (22.9%)	102 (14.4%)	505 (14.0%)	
2	322 (38.2%)	333 (41.3%)	225 (43.9%)	292 (40.2%)	308 (43.5%)	1480 (41.2%)	
3	179 (21.3%)	205 (25.4%)	125 (24.4%)	127 (17.5%)	168 (23.7%)	804 (22.4%)	
4 (more than 5 times)	193 (22.9%)	172 (21.3%)	91 (17.8%)	127 (17.5%)	112 (15.8%)	695 (19.3%)	
Missing	54 (6.4%)	16 (2.0%)	9 (1.8%)	14 (1.9%)	18 (2.5%)	111 (3.1%)	
Schedule							39.48 (8)***
0 someone else chose	121 (14.4%)	118 (14.6%)	81 (15.8%)	93 (12.8%)	95 (13.4%)	508 (14.1%)	
1 person had input	372 (44.2%)	359 (44.5%)	249 (48.6%)	262 (36.1%)	288 (40.7%)	1530 (42.6%)	
2 person made choice	308 (36.6%)	324 (40.1%)	176 (34.4%)	357 (49.2%)	316 (44.6%)	1481 (41.2%)	
Missing	41 (4.9%)	6 (0.7%)	6 (1.2%)	14 (1.9%)	9 (1.3%)	76 (2.1%)	
Free Time							32.00 (8)***
0 someone else chose	48 (5.7%)	44 (5.5%)	26 (5.1%)	36 (5.0%)	40 (5.6%)	194 (5.4%)	
1 person had input	252 (29.9%)	227 (28.1%)	139 (27.1%)	138 (19.0%)	187 (26.4%)	943 (26.2%)	
2 person made choice	500 (59.4%)	525 (65.1%)	341 (66.6%)	537 (74.0%)	473 (66.8%)	2376 (66.1%)	
Missing	42 (5.0%)	11 (1.4%)	6 (1.2%)	15 (2.1%)	8 (1.1%)	82 (2.3%)	
Choose Buy							21.14 (8) **
0 someone else chose	89 (10.6%)	73 (9.0%)	51 (10.0%)	76 (10.5%)	60 (8.5%)	349 (9.7%)	
1 person had input	324 (38.5%)	318 (39.4%)	223 (43.6%)	305 (42.0%)	343 (48.4%)	1513 (42.1%)	
2 person made choice	383 (45.5%)	397 (49.2%)	225 (43.9%)	319 (43.9%)	286 (40.4%)	1610 (44.8%)	
Missing	46 (5.5%)	19 (2.4%)	13 (2.5%)	26 (3.6%)	19 (2.7%)	123 (3.4%)	

* p < .05, ** p < .005, *** p < .001

Table 3

Model Fit and Parsimony Statistics Comparing Longitudinal Measurement Invariance Models

Model	df	ChiSq.	*p*	CFI	TLI	RMSEA	90% LB	90% UB	*p*	SRMR
Original 5-group	120	291.15	0.00	0.97	0.96	0.05	0.04	0.05	0.68	0.05
T Invariant	128	302.88	0.00	0.97	0.96	0.05	0.04	0.05	0.77	0.05
TL Invariant	160	440.13	0.00	0.96	0.95	0.05	0.05	0.06	0.20	0.05
TLI Invariant	198	2,814.63	0.00	0.59	0.63	0.15	0.14	0.15	0.00	0.07

Comparison of Model Fit Between Models

Model Comparison	Δdf	ΔCFI	ΔTLI	ΔRMSEA	Δ90% LB	Δ90% UB	ΔSRMR
T Invariant - Original	8.00	0.00	0.00	0.00	0.00	0.00	0.00
TL Invariant - T Invariant	32.00	-0.02	-0.01	0.01	0.01	0.01	0.00
TLI Invariant - TL Invariant	38.00	-0.36	-0.32	0.09	0.09	0.09	0.02

Note. N = 3,595. T = Thresholds, L = Loadings, I = Intercepts. CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, RMSEA = Root Mean Square Error of Approximation, along with its 90% lower bound (LB) and upper bound (UB), *p*-value for close fit, and Standardized Root Mean Square Residual (SRMR).

Table 4

Model Fit and Parsimony Statistics Comparing Longitudinal Measurement Invariance Models, No Covid-19 Year

Model	df	ChiSq.	*p*	CFI	TLI	RMSEA	90% LB	90% UB	*p*	SRMR
Original 5-group	96	155.60	0.00	0.99	0.98	0.03	0.02	0.04	1.00	0.04
T Invariant	102	162.59	0.00	0.99	0.98	0.03	0.02	0.04	1.00	0.04
TL Invariant	129	212.82	0.00	0.98	0.98	0.03	0.02	0.04	1.00	0.04
TLI Invariant	156	686.69	0.00	0.89	0.90	0.07	0.07	0.08	0.00	0.05

Comparison of Model Fit Between Models

Model Comparison	Δdf	ΔCFI	ΔTLI	ΔRMSEA	Δ90% LB	Δ90% UB	ΔSRMR
T Invariant - Original	6	0.00	0.00	0.00	0.00	0.00	0.00
TL Invariant - T Invariant	27	0.00	0.00	0.00	0.00	0.00	0.00
TLI Invariant - TL Invariant	27	-0.09	-0.08	0.04	0.04	0.04	0.01

Note. N = 3,595. T = Thresholds, L = Loadings, I = Intercepts. CFI = Comparative Fit Index, TLI = Tucker-Lewis Index, RMSEA = Root Mean Square Error of Approximation, along with its 90% lower bound (LB) and upper bound (UB), *p*-value for close fit, and Standardized Root Mean Square Residual (SRMR).