



# Fidelity and IPS: does quality of implementation predict vocational outcomes over time for organizations treating persons with severe mental illness in the Netherlands?

Lars de Winter<sup>1</sup> · Christianne Couwenbergh<sup>1</sup> · Jaap van Weeghel<sup>1</sup> · Cris Bergmans<sup>1</sup> · Gary R. Bond<sup>2</sup>

Received: 22 December 2019 / Accepted: 18 May 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

**Purpose** Individual placement and support (IPS) is an evidence-based supported employment intervention. Quality of IPS implementation is assessed using a validated fidelity scale. Previous studies found a positive association between fidelity and employment outcomes at a single time-point. This study examines the longitudinal associations between IPS fidelity scores and employment outcomes.

**Methods** We examined fidelity and employment outcome data for 27 IPS programs in the Netherlands providing IPS. These programs received at least one fidelity assessment and reported quarterly employment outcomes for at least one year to a central registry between 2014 and 2019. We first examined changes over time for fidelity and employment outcome. Then we analyzed the longitudinal associations between the quarterly employment outcomes and the IPS fidelity assessments on multiple time-points using a mixed-model analysis for the 17 programs with at least two fidelity assessments.

**Results** Both IPS fidelity and employment outcomes improved over time, with the largest improvement in employment outcomes achieved after 18 months of implementation. We found a significant positive longitudinal association between IPS-fidelity and employment.

**Conclusions** Improvement of fidelity is associated with improvement of employment outcomes over time. Future research should be focused on the improvement of specific elements of IPS implementation and their influence on employment outcomes.

**Keywords** IPS · Fidelity · Employment · Work · Longitudinal

## Introduction

Employment is a key factor in rehabilitation for people with severe mental illnesses (SMI). Competitive employment offers clients financial independence, contributes to their recovery process and has beneficial effects on mental health, self-esteem and quality of life [1, 2]. However, employment rates for people with SMI are far lower than in the general population. National surveys in the Netherlands have found that only 10%–17% of clients with SMI in specialized secondary care were competitively employed,

with no indications of improvement over time [3–5]. Mental health providers often view people with SMI as unable to work competitively. Partly as a reflection of these beliefs, 89% of people with SMI receive disability benefits [6] and, until recently, they were often placed in sheltered employment settings [7]. The disincentives for people with SMI to seek competitive employment are especially strong in high welfare countries such as the Netherlands. In addition, social exclusion and stigmatization impair participation of people with SMI in the labor market [8, 9]. Nevertheless, many unemployed people with SMI want to be employed. For example, one study found that 70% of the SMI clients younger than 45 in the Netherlands expressed the desire to work in a regular job [5]. These findings document the needs for interventions targeting support in regaining and maintaining competitive employment for this population.

Of all vocational rehabilitation programs, the most clearly described model is Individual Placement and Support (IPS

---

✉ Lars de Winter  
Ldewinter@kcpheenos.nl

<sup>1</sup> Phrenos Center of Expertise for Severe Mental Illness, Utrecht, The Netherlands

<sup>2</sup> Westat, Lebanon, NH, USA

[10]). IPS follows eight basic principles: the goal of competitive employment, zero exclusion and eligibility based on client choice, attention to client preferences, rapid job search, integration of employment services and mental health treatment, personalized benefits counseling, targeted job development, and individualized, long-term support. IPS has been disseminated worldwide, to a variety of countries in North America, Europe, Asia, and Oceania [11]. The effectiveness of IPS has been investigated extensively. Numerous meta-analyses and systematic reviews have documented that IPS has shown beneficial effects on employment outcomes compared with other vocational services [e.g. 7, 12, 13, 14]. In the Netherlands, a 30-month randomized controlled trial also found beneficial effects of IPS on employment outcomes for people with SMI [2].

To achieve these beneficial effects, IPS should be implemented according to the model (i.e., with good fidelity), because the quality of program implementation (its fidelity to a well-defined model) influences program effectiveness [15, 16]. Therefore, a 25-item IPS fidelity scale was developed measuring the quality of implementation to model standards [17]. This IPS fidelity scale assesses the quality of staffing, organization and services for each IPS program. It has been implemented worldwide, including in the UK [18] and throughout the international IPS learning community [19]. A number of studies already showed significant positive associations between fidelity on the IPS 25-item scale and competitive employment [17, 20, 21].

In the Netherlands, IPS programs received fidelity assessments by independent reviewers every 2 years and they collected and reported program-level employment outcome data every 3 months [22]. In this study, the aims are to examine changes over time for fidelity, its item scores and employment outcomes, and to analyze the association between the changes in fidelity and employment within these IPS programs over the past 5 years. Previous studies have been limited by examining the cross-sectional association between IPS fidelity and employment outcomes. Therefore, to our knowledge this is the first study analyzing the influence of fidelity scores on employment outcomes longitudinally over a time range of multiple years. The main questions for this study are: 1. Do programs improve in fidelity and employment outcomes over time? 2. Does fidelity predict employment outcomes of IPS over time?

## Methods

### Overview and study flow

In this study, we investigated the longitudinal association between changes in fidelity and employment outcomes over a time-period of 5 years, between the first quarter of 2014

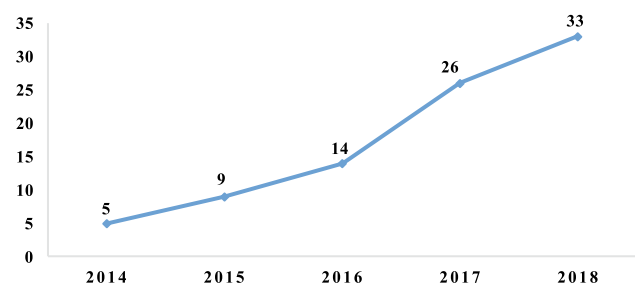
and the first quarter of 2019. The number of IPS programs in the Netherlands grew from five in 2014 to 33 at the end of 2018 (see Fig. 1). In 2019, the 33 IPS programs employed a total number of 219 IPS specialists serving 3138 clients.

Among these 33 IPS programs we identified 27 IPS programs that both collected employment outcome data over a period of at least 1 year and received at least one fidelity assessment. For these 27 programs, we were able to give an overview of the changes in fidelity and employment outcomes over the past 5 years.

For each IPS program, we collected the first measurement of employment outcomes and fidelity scores at the start of implementation of IPS. IPS programs received fidelity assessments every 2 years and collected employment outcome data every 3 months. Each IPS program initiated IPS at different time points between 2014 and 2019 and is in a different stage of implementation. This means that the sample size in this study is shrinking over time, as a decreasing number of IPS programs collected fidelity and employment outcome data over a longer time period. As a consequence of this shrinking sample size, 17 IPS programs received at least two fidelity assessments and only seven programs received at least three fidelity assessments. This means that for our analysis of the association between the changes in fidelity and employment in the IPS programs over time, we only included the 17 IPS programs with at least two fidelity assessments.

## Sample

The 27 IPS programs were part of mental health care organizations throughout the Netherlands. From these 27 IPS programs, 21 were provided in Flexible Assertive Community Treatment (F-ACT [23]) teams, early intervention in psychosis teams (EIP [24]) or both. Furthermore, four IPS programs were provided in housing services for people with severe mental illnesses and two IPS programs were provided in specialized addiction health care centers. All IPS programs provided IPS to clients with SMI. As of 2019,



**Fig. 1** Growth in the number of IPS programs in the Netherlands number between 2014 and 2018

the 27 IPS programs employed on average 8.8 IPS specialists ( $SD = 5.4$ ;  $min = 2$ ;  $max = 24$ ; mean full time equivalent of 6.0 staff positions) serving on average 116.2 clients ( $SD = 85.6$ ;  $range = 14-298$ ) per program.

## Measurements and data collection

### Fidelity

We assessed IPS fidelity using the 25-item IPS fidelity scale [25]. Each item is rated on a 5-point behaviorally anchored scale, ranging from 1 (no implementation) to 5 (full implementation), with intermediate numbers representing progressively greater degrees of implementation [25]. The IPS fidelity scale is scored by summing the item scores. The total score ranges from 25 to 125 points. IPS programs scoring between 74 and 99 are considered to have ‘fair’ fidelity, programs scoring between 100 and 114 have ‘good’ fidelity and programs scoring between 115 and 125 have ‘exemplary’ fidelity. IPS programs scoring below 74 are considered to provide ‘no IPS’ [25]. The IPS-25 scale has a good internal consistency and a moderate predictive validity [17].

Fidelity review procedures were as follows: Two assessors conduct fidelity reviews during a full-day visit at the IPS program. The assessors rely on different sources, including interviews with IPS specialists, staff members, clients, family members and directors. They also observe mental health treatment team meetings and vocational unit meetings and investigate program documents and client records. After completing independent ratings on the 25-item scale, assessors discuss and resolve any rating discrepancies to achieve consensus ratings.

In the current study, the fidelity reviews were conducted by nine trained external fidelity reviewers, who were not affiliated with the IPS program or organization under review.

### Employment rate outcome

Every 3 months each IPS program compiled program-level statistics, using a Dutch translation of the IPS Quarterly Employment Reporting Form [26]. This form includes information about the number of clients receiving IPS, the competitive employment and education rate of the IPS program, the drop-out rate and the full-time equivalent (FTE) of all employment specialists working in the IPS program. All definitions of the variables have been operationalized in a manual available for all IPS programs. The primary outcome for this study was the percentage of people who were competitively employed, defined as paid employment in a regular setting with at least minimum wage, at any time during the 3 month period.

The data were collected by IPS specialists and assembled by the IPS coordinator, who compiled all data for the IPS program. The compiled data were sent to a researcher of Phrenos Center of Expertise (LdW), who completed a quality control of accuracy and consistency with the outcome reporting manual before processing the outcome data.

### Analysis

#### Overview of the IPS program characteristics, employment outcomes and fidelity over time

This study focused on a 4-year follow-up period for which complete program characteristics, fidelity and employment outcome data were available.

We first provided a descriptive overview of the quarterly outcome data of program characteristics (i.e. caseload characteristics and number of IPS specialists) and employment outcomes for the first 4 years after the start of implementation of IPS, retrieved from the IPS Quarterly Employment Reporting Form. We reported the descriptive statistics and trends of improvements over time for the caseload size, number of IPS specialists and employment outcomes were analyzed using a mixed model analysis.

Descriptive statistics for the first, second and third fidelity assessment were calculated for the total fidelity score and specific item scores. Because not all organizations completed two or three fidelity assessments we also showed descriptive statistics for both fidelity and employment outcomes for IPS programs that completed two and three fidelity assessments. We analyzed differences between the first, second and third fidelity assessment for the IPS programs that completed at least two fidelity measurements and the IPS programs that completed at least three fidelity measurements separately using a one-way repeated measures analysis of variance (ANOVA). By analyzing the completer samples separately we are able to indicate potential differences between the total sample and the completer samples at specific time-points. To demonstrate potential above- and below-average performances of the IPS programs for specific fidelity items during the fidelity assessments, we compared these specific fidelity item scores with the mean item score during each fidelity assessment using independent sample t-tests.

#### Longitudinal association between fidelity and employment rate

The overall association between the total fidelity score and employment rate on all time-points was analyzed using Spearman correlations. Furthermore, for our primary analysis, investigating the influence of improvement of fidelity on the improvement of employment rate in the IPS programs

over the four year study period, we used a linear mixed model analysis on all 17 IPS programs that received at least two fidelity assessments. Because employment data were collected every 3 months while fidelity was measured every 2 years, we calculated the mean employment rate over the four quarterly time periods spanning the 6 months before and after each fidelity assessment to estimate employment outcomes corresponding to the same time frame as the fidelity assessments. The mixed model analysis includes all employment outcomes and fidelity assessments and limits attrition bias by adequately handling missing data by avoiding the listwise deletion of IPS programs with missing data. Furthermore, the mixed model analysis adequately corrects for statistical interdependence between the repeated measures within each IPS program [33]. To adequately analyze the influence of fidelity on program-level employment outcomes over time, we assessed the IPS program as subject level (level 1) and the sequence of fidelity as the level of time (level 2) in the analysis model. The employment rate of the IPS program was the primary outcome and the total fidelity score was a predictor in the mixed model analysis. To control for potential beneficial effects of the completer samples we also carried-out a repeated-measures ANOVA for the seven IPS programs that completed all three fidelity measures. We controlled for the assumptions of homoscedasticity and normality of residuals.

The influence of changes in fidelity on changes in employment outcomes over time is additionally explained using an exploratory descriptive post-hoc analysis based on the investigation of descriptive fidelity outcomes. To carry-out this analysis we clustered all IPS programs with at least two fidelity assessments into six potential categories (i.e. high or low fidelity at baseline which is improving, stable or decreasing over time). For these clusters, separate trend lines for employment rate over time are presented to show the differences in employment rate over time. The level of fidelity at baseline (high vs. low) was based on a fidelity cut-off score of at least 100 ('good' fidelity) based on the IPS-25 fidelity validation paper [17]. Improvement or deterioration over time was assessed by a difference between the first and last fidelity score of at least 13. The difference score was based on the standard deviation of the IPS-25 score from all IPS programs that participated in the IPS-25 fidelity validation paper [17].

## Results

### Overview of the IPS program characteristics, employment outcomes and fidelity over time

#### Number of clients and employment specialists

As shown in Table 1 the mean program-level caseload significantly increased over time from 76.4 clients in the first year to 153.5 clients in the fourth year after starting IPS ( $B = 22.41$ ;  $\chi^2 = 8.04$ ;  $p < 0.01$ ). By contrast, the mean number of employment specialists per program did not increase significantly over the years ( $B = 0.39$ ;  $\chi^2 = 27.74$ ;  $p = 0.55$ ). An average number of 22.4 clients ( $SD = 24.4$ ) per program are newly enrolled in IPS during each quarterly assessment period, whereas on average 14.0 clients ( $SD = 14.7$ ) per program ended IPS services during this period.

#### Employment rate outcome

The differences in employment rate over the years are shown in Fig. 2 for the total sample and for the completer samples. The mean employment rate increased from 26.0% in the first year to 45.1% in the fourth year after the start of implementation of IPS. The mixed model analysis has indicated a significant improvement over time ( $B = 0.81$ ;  $Wald \chi^2 = 20.67$ ;  $p < 0.01$ ). As the overall improvement was significant we dummy coded each quarterly employment rate outcome to show in which time period the outcomes specifically differed from the overall mean employment rate. This post-hoc analysis showed that the employment rate is significantly below average in the first ( $B = -10.59$ ;  $\chi^2 = 4.87$ ;  $p < 0.05$ ), third ( $B = -9.83$ ;  $\chi^2 = 4.31$ ;  $p < 0.05$ ) and fourth quarter ( $B = -10.07$ ;  $\chi^2 = 4.53$ ;  $p < 0.05$ ) of the first year and the first ( $B = -9.86$ ;  $\chi^2 = 4.10$ ;  $p < 0.05$ ) and second quarter ( $B = -11.98$ ;  $\chi^2 = 5.95$ ;  $p < 0.05$ ) of the second year after starting IPS.

#### Fidelity

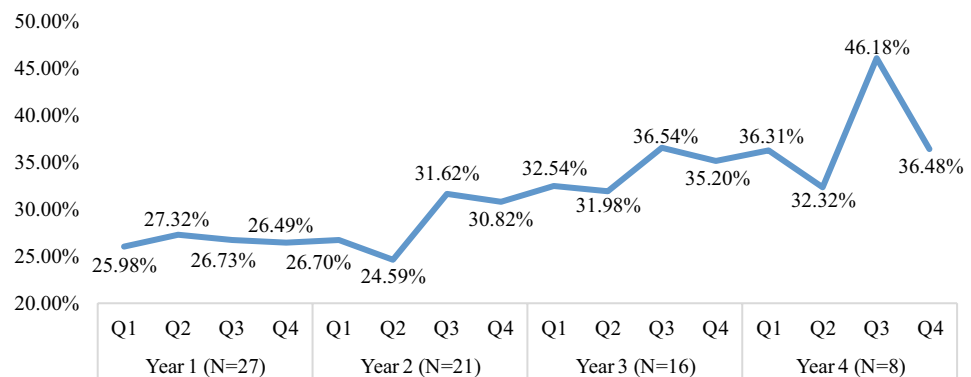
The analysis of the differences between the first, second and third fidelity assessments, presented in Table 2, indicated a significant improvement of fidelity over time for IPS programs that completed both two ( $F = 6.51$ ;  $df = 2$ ;  $p < 0.05$ ) and three fidelity assessments ( $F = 6.62$ ;  $df = 2$ ;  $p < 0.01$ ). Furthermore, we found a significant improvement of item 6 in the organization section (zero exclusion criteria;  $F = 3.20$ ;  $df = 2$ ;  $p < 0.05$ ) and item 7 in the organization section (agency focus on competitive employment;

**Table 1** Caseload of clients receiving IPS and number of employment specialists per IPS program over the years

Assesment period after initiating IPS programs	Number of IPS programs	Caseload (number of clients per IPS program)		Number of clients enrolled during this quarter*		Number of clients ending IPS service this quarter		Number of employment specialists	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Year 1 mean**	25	76.35	65.72	17.26	19.49	9.12	12.82	7.18	4.68
First-quarter year 1	27	73.11	64.36			7.06	7.67	7.05	4.43
Second quarter year 1	27	79.79	58.81	15.11	15.31	9.38	13.53	7.52	4.32
Third quarter year 1	25	87.04	56.57	12.78	14.74	9.15	14.73	7.43	4.25
Fourth quarter year 1	25	89.72	60.11	24.05	25.64	11.05	14.75	7.38	4.51
Year 2 mean**	14	95.43	62.16	24.00	27.93	15.50	15.22	7.85	5.35
First quarter year 2	21	95.52	66.55	23.00	35.33	12.62	17.52	7.55	4.74
Second quarter year 2	20	95.75	69.60	24.47	29.64	11.63	12.06	8.05	4.93
Third quarter year 2	18	102.89	77.60	22.29	18.19	12.88	11.71	9.00	5.19
Fourth quarter year 2	17	126.94	78.23	26.36	26.37	14.87	14.79	10.43	5.96
Year 3 mean**	8	148.25	77.74	22.78	25.67	17.15	14.80	10.00	6.95
First quarter year 3	16	142.88	85.40	25.93	33.85	15.87	17.94	9.87	6.23
Second quarter year 3	9	129.56	80.64	16.88	18.55	18.67	13.54	8.78	5.83
Third quarter year 3	8	141.88	82.53	20.88	24.99	17.00	15.33	9.38	5.88
Fourth quarter year 3	8	148.75	88.22	25.00	14.64	18.13	10.44	9.88	6.01
Year 4 mean**	3	153.50	60.10	21.73	22.30	22.17	16.00	7.00	2.83
First quarter year 4	8	148.25	81.41	21.88	20.76	22.38	20.35	10.57	6.32
Second quarter year 4	6	154.00	83.62	16.71	14.82	17.43	17.32	12.60	6.99
Third quarter year 4	4	152.75	17.67	34.00	34.58	26.25	10.28	8.67	3.51
Fourth quarter year 4	3	135.67	44.09	15.00	25.98	25.67	9.07	9.50	3.54

\*Data concerning the number of newly enrolled clients was available from the second quarter after the start of IPS

\*\*The annual means are the mean quarterly outcomes of the specific measurement that year

**Fig. 2** Quarterly competitive employment rate over the four-year study period

$F = 4.13$ ;  $df = 2$ ;  $p < 0.05$ ). For all other items we did not find any significant improvement over time in either completer samples.

Independent sample  $t$  test comparing specific fidelity items with the mean fidelity item score, within each fidelity assessment separately, has shown a significantly higher score than average on at least two fidelity assessments for item 3 of the staffing section (vocational generalists) and for item 2 (Disclosure) and item 11 (Individualized

follow-along support) of the services section. However, item 5 of the organization section (Role of employment supervisor) and item 6 of the services section (Frequent Employer Contact) had a significantly lower score than average on at least two fidelity assessments. There were no significant differences in any fidelity score (i.e. total score and specific items) between the total sample and the completer samples.

**Table 2** Fidelity outcome scores between different time measures

Fidelity score	Assessment 1			Assessment 2			Assessment 3			Analysis	
	<i>M</i> (SD)	<i>N</i>	Min–Max	<i>M</i> (SD)	<i>N</i>	Min–Max	<i>M</i> (SD)	<i>N</i>	Min–Max	<i>F</i>	<i>p</i>
<b>Total IPS-25 fidelity score*</b>											
Total sample	91.00 (10.79)	27	64–107	97.06 (9.79)	17	73–113	107.86 (4.53)	7	103–115	6.51	0.01
Completers 2 fidelity assessments	89.41 (12.15)	17	64–104	97.29 (8.40)	7	86–108				6.62	0.00
Completers 3 fidelity assessments	89.86 (10.84)	7	76–104								
<b>Staffing section</b>											
Item 1. Caseload size	<b>3.96 (1.37)<sup>+</sup></b>	27	1–5	4.24 (0.97)	17	2–5	4.57 (1.13)	7	2–5	1.75	0.17
Total sample	<b>3.59 (1.33)<sup>+</sup></b>	17	1–5	4.29 (0.76)	7	3–5				1.16	0.35
Completers 2 fidelity assessments	3.57 (1.40)	7	1–5								
Completers 3 fidelity assessments											
Item 2. Employment services staff	<b>3.96 (1.53)<sup>+</sup></b>	27	1–5	4.29 (1.10)	17	2–5	4.86 (0.38)	7	4–5	2.07	0.12
Total sample	<b>3.59 (1.58)<sup>+</sup></b>	17	1–5	4.86 (0.38)	7	4–5				1.19	0.34
Completers 2 fidelity assessments	3.86 (1.95)	7	1–5								
Completers 3 fidelity assessments											
Item 3. Vocational generalists	<b>4.37 (0.88)<sup>+</sup></b>	27	3–5	<b>4.35 (0.79)<sup>+</sup></b>	17	3–5	4.86 (0.38)	7	4–5	1.05	0.38
Total sample	<b>4.29 (0.92)<sup>+</sup></b>	17	3–5	4.29 (0.76)	7	3–5				1.77	0.19
Completers 2 fidelity assessments	4.00 (1.00)	7	3–5								
Completers 3 fidelity assessments											
<b>Organization section</b>											
Item 1. Integration through team assignment	3.85 (1.38)	27	2–5	4.35 (1.06)	17	2–5	5.00 (0.00)	7	5–5	1.08	0.37
Total sample	4.18 (1.29)	17	2–5	4.71 (0.49)	7	3–5				1.30	0.31
Completers 2 fidelity assessments	4.29 (1.11)	7	2–5								
Completers 3 fidelity assessments											
Item 2. Integration through frequent team contact	3.54 (1.25)	27	1–5	3.88 (0.93)	17	1–5	4.29 (0.76)	7	3–5	0.83	0.49
Total sample	3.62 (1.05)	17	1–5	3.86 (0.38)	7	3–4				0.71	0.56
Completers 2 fidelity assessments	4.07 (0.45)	7	3–5								
Completers 3 fidelity assessments											
Item 3. Collaboration between employment specialists and Vocational Rehabilitation counselors	3.89 (0.75)	27	2–5	<b>4.12 (0.78)<sup>+</sup></b>	17	2–5	4.57 (0.53)	7	4–5	1.72	0.18
Total sample	3.88 (0.86)	17	2–5	4.29 (0.49)	7	4–5				2.99	0.06
Completers 2 fidelity assessments	3.57 (0.98)	7	2–5								
Completers 3 fidelity assessments											
Item 4. Vocational unit	<b>3.04 (1.29)<sup>–</sup></b>	27	1–5	3.59 (1.18)	17	1–5	4.14 (0.69)	7	3–5	2.09	0.12
Total sample	2.88 (1.41)	17	1–5	3.71 (1.60)	7	1–5				0.75	0.54
Completers 2 fidelity assessments	3.14 (1.35)	7	1–5								
Completers 3 fidelity assessments											
Item 5. Role of employment supervisor	<b>1.88 (1.31)<sup>–</sup></b>	27	1–5	<b>2.18 (1.07)<sup>–</sup></b>	17	1–4	2.71 (0.76)	7	2–4	1.13	0.35
Total sample	2.06 (1.52)	17	1–5	2.29 (0.95)	7	1–3				1.08	0.39
Completers 2 fidelity assessments	2.83 (1.17)	7	1–4								
Completers 3 fidelity assessments											

Table 2 (continued)

Fidelity score	Assessment 1			Assessment 2			Assessment 3			Analysis	
	M (SD)	N	Min–Max	M (SD)	N	Min–Max	M (SD)	N	Min–Max	F	p
Item 6. Zero exclusion criteria	3.93 (0.87)	27	2–5	4.12 (0.86)	17	3–5	4.71 (0.49)	7	4–5	1.42	0.25
Total sample	4.00 (1.00)	17	2–5	4.29 (0.76)	7	3–5				3.20	0.048
Completers 2 fidelity assessments	3.43 (1.13)	7	2–5								
Completers 3 fidelity assessments*											
Item 7. Agency focus on competitive employment	2.98 (1.00)	27	1–5	<b>3.12 (0.93)<sup>+</sup></b>	17	2–5	4.00 (0.58)	7	3–5	4.13	0.01
Total sample	2.68 (0.92)	17	1–4	3.29 (1.11)	7	2–5				2.06	0.14
Completers 2 fidelity assessments*	2.79 (1.07)	7	1–4								
Completers 3 fidelity assessments											
Item 8. Executive team support for SE	3.04 (1.43)	25	1–5	3.24 (1.35)	17	1–5	4.14 (0.69)	7	3–5	1.87	0.15
Total sample	2.75 (1.53)	17	1–5	3.00 (1.15)	7	1–4				1.33	0.30
Completers 2 fidelity assessments	3.50 (1.38)	7	2–5								
Completers 3 fidelity assessments											
<b>Services section</b>											
Item 1. Work incentives planning	4.11 (0.93)	27	3–5	<b>4.06 (0.97)<sup>+</sup></b>	17	2–5	4.43 (0.98)	7	3–5	0.58	0.63
Total sample	4.06 (0.90)	17	3–5	3.86 (0.69)	7	3–5				1.02	0.41
Completers 2 fidelity assessments	4.43 (0.79)	7	3–5								
Completers 3 fidelity assessments											
Item 2. Disclosure	<b>4.67 (0.73)<sup>+</sup></b>	27	3–5	<b>4.94 (0.24)<sup>+</sup></b>	17	4–5	5.00 (0.00)	7	5–5	1.00	0.40
Total sample	<b>4.71 (0.69)<sup>+</sup></b>	17	3–5	5.00 (0.00)	7	5–5				1.42	0.27
Completers 2 fidelity assessments	4.57 (0.79)	7	3–5								
Completers 3 fidelity assessments											
Item 3. Ongoing, work-based vocational assessment	4.07 (0.62)	27	3–5	<b>3.88 (0.70)<sup>–</sup></b>	17	2–5	4.29 (0.49)	7	4–5	0.75	0.53
Total sample	4.06 (0.56)	17	3–5	4.00 (0.58)	7	3–5				0.64	0.60
Completers 2 fidelity assessments	3.86 (0.69)	7	3–5								
Completers 3 fidelity assessments											
Item 4. Rapid search for competitive job	2.81 (1.30)	27	1–5	3.00 (1.27)	17	1–5	3.29 (1.25)	7	1–5	2.04	0.12
Total sample	2.35 (1.37)	17	1–4	2.71 (1.25)	7	1–5				1.52	0.24
Completers 2 fidelity assessments	2.43 (1.27)	7	1–4								
Completers 3 fidelity assessments											
Item 5. Individualized job search	4.59 (0.64)	27	3–5	<b>4.69 (0.79)<sup>+</sup></b>	17	2–5	5.00 (0.00)	7	5–5	0.85	0.48
Total sample	4.59 (0.51)	17	4–5	4.67 (0.52)	7	4–5				0.98	0.43
Completers 2 fidelity assessments	4.71 (0.49)	7	4–5								
Completers 3 fidelity assessments											
Item 6. Job development—Frequent employer contact	<b>2.41 (1.08)<sup>–</sup></b>	27	1–5	<b>2.53 (0.72)<sup>–</sup></b>	17	2–4	<b>2.57 (0.79)<sup>–</sup></b>	7	2–4	0.74	0.54
Total sample	2.53 (1.23)	17	1–5	2.29 (0.49)	7	2–3				1.19	0.34
Completers 2 fidelity assessments	2.29 (1.25)	7	1–5								
Completers 3 fidelity assessments											

Table 2 (continued)

Fidelity score	Assessment 1			Assessment 2			Assessment 3			Analysis	
	M (SD)	N	Min–Max	M (SD)	N	Min–Max	M (SD)	N	Min–Max	F	p
Item 7. Job development—Quality of employer contact	3.37 (1.24)	27	1–5	<b>3.65 (1.11)<sup>+</sup></b>	17	1–5	4.14 (0.69)	7	3–5	0.52	0.67
Total sample	3.47 (1.46)	17	1–5	3.57 (1.40)	7	1–5				1.07	0.39
Completers 2 fidelity assessments	3.00 (1.41)	7	1–5								
Completers 3 fidelity assessments											
Item 8. Diversity of job types	4.07 (1.59)	27	1–5	4.35 (1.17)	17	1–5	4.71 (0.76)	7	3–5	0.30	0.82
Total sample	<b>4.24 (1.56)<sup>+</sup></b>	17	1–5	4.00 (1.53)	7	1–5				0.45	0.72
Completers 2 fidelity assessments	4.00 (1.91)	7	1–5								
Completers 3 fidelity assessments											
Item 9. Diversity of employers	4.04 (1.68)	27	1–5	4.59 (0.80)	17	3–5	5.00 (0.00)	7	5–5	1.19	0.33
Total sample	<b>4.06 (1.71)<sup>+</sup></b>	17	1–5	4.57 (0.79)	7	3–5				0.67	0.58
Completers 2 fidelity assessments	4.14 (1.86)	7	1–5								
Completers 3 fidelity assessments											
Item 10. Competitive jobs	2.30 (1.44)	27	1–5	3.12 (1.32)	17	1–5	3.43 (1.72)	7	1–5	1.24	0.31
Total sample	2.35 (1.46)	17	1–5	3.43 (1.81)	7	1–5				0.11	0.95
Completers 2 fidelity assessments	3.00 (1.53)	7	1–5								
Completers 3 fidelity assessments											
Item 11. Individualized follow-along supports	<b>4.41 (1.22)<sup>+</sup></b>	27	1–5	<b>4.71 (0.47)<sup>+</sup></b>	17	4–5	5.00 (0.00)	7	5–5	1.21	0.32
Total sample	4.24 (1.48)	17	1–5	4.71 (0.49)	7	4–5				1.09	0.38
Completers 2 fidelity assessments	4.00 (1.83)	7	1–5								
Completers 3 fidelity assessments											
Item 12. Time-unlimited follow-along supports	<b>4.11 (1.31)<sup>+</sup></b>	27	1–5	4.71 (0.59)	17	3–5	4.71 (0.76)	7	3–5	2.29	0.09
Total sample	<b>3.82 (1.51)<sup>+</sup></b>	17	1–5	4.43 (0.79)	7	3–5				1.46	0.26
Completers 2 fidelity assessments	3.43 (1.90)	7	1–5								
Completers 3 fidelity assessments											
Item 13. Community-based services	3.04 (1.22)	27	1–5	3.35 (1.11)	17	1–5	3.86 (0.69)	7	3–5	0.76	0.53
Total sample	3.18 (1.19)	17	1–5	3.86 (0.90)	7	3–5				0.57	0.64
Completers 2 fidelity assessments	3.29 (1.25)	7	2–5								
Completers 3 fidelity assessments											
Item 14. Assertive engagement and outreach by integrated treatment team	4.37 (0.88)	27	2–5	4.29 (0.92)	17	2–5	4.57 (0.53)	7	4–5	0.55	0.65
Total sample	4.53 (0.64)	17	3–5	4.00 (1.00)	7	2–5				1.20	0.34
Completers 2 fidelity assessments	4.57 (0.53)	7	4–5								
Completers 3 fidelity assessments											

\*Significant improvement over time ( $p < 0.05$ )

+Itemscore significantly above average during that measurement

–Itemscore significantly below average during that measurement



## Longitudinal associations between fidelity and employment rate

### Main analysis

We detected a significant overall medium positive association between the total fidelity score and employment rate across all fidelity assessments ( $r=0.48$ ;  $p<0.01$ ).

The total fidelity score met the assumption of homoscedasticity and normality of residuals. The intraclass correlation between the repeated measures of employment outcomes within the IPS programs was moderate to high ( $ICC=0.68$ ), which justifies the use of the mixed model analysis for the IPS program outcomes. The mixed model analysis has shown that the IPS-25 fidelity was a significant positive predictor for employment rate ( $B=0.76$ ;  $Wald \chi^2=4.82$ ;  $p<0.05$ ). The repeated measures ANOVA confirmed this result with a significant fidelity\*time interaction on employment rate for the seven IPS programs that completed three fidelity measures ( $F=4.30$ ;  $df=2$ ;  $p<0.01$ ). These results have indicated that the improvement of fidelity scores are associated with improvement of employment rate in the IPS programs. However, we did not find a significant interaction in the mixed model analysis, indicating no differences in improvement of employment rate over time between IPS programs with low and high fidelity scores during the study period.

### Post-hoc descriptive analysis

The significant positive influence of fidelity on employment outcomes and the lack of interaction between fidelity and time has also been explained by the descriptive post-hoc analysis clustering the IPS programs based on the fidelity score at baseline and the improvement over time. Based on the criteria described in the methods section we were able to assess three clusters: IPS programs with a high and stable fidelity, low and improving fidelity and low and stable

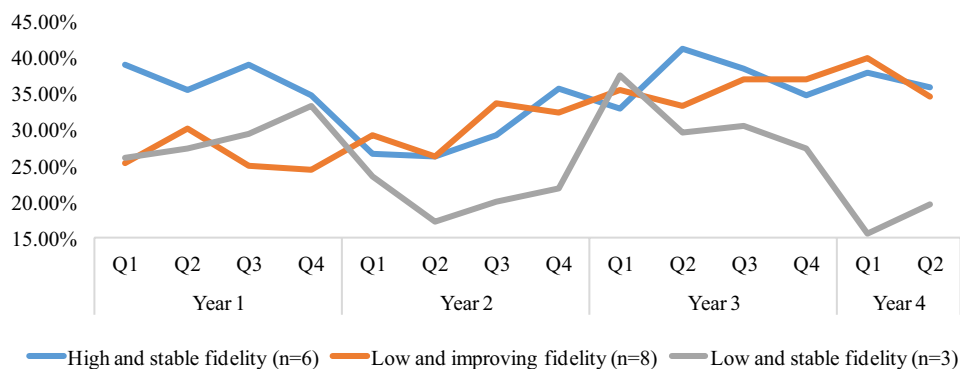
fidelity. For these three clusters, we were able to assess the employment rate from the IPS programs over a period of 3.5 years. The mean employment rate over time for these clusters have been presented in Fig. 3. The figure has shown relatively high employment rates that is stable over time for the high stable fidelity group, relatively low employment rate, that was stable over time, for the low stable fidelity group and a trend of improvement of employment rate in the low improving group. The trends of employment rate over time are in accordance with each cluster, confirming the positive predictive value of fidelity assessments for employment rate over time in the main analysis.

## Discussion

In this report, we investigated the longitudinal association between IPS fidelity scores and employment outcomes of IPS programs over a period of 4 years. Mixed model analyses have shown a positive association between improvement of fidelity scores and improvement in program-level employment rate over time. These results are supported by the repeated measures ANOVA, analyzing IPS programs that completed all three fidelity assessments. We did not find a significant interaction between fidelity assessments and employment rate over time. This indicates that improvement in program-level employment rate does not differ for IPS programs starting at low fidelity as for programs starting at high fidelity.

These results demonstrate that improvement in fidelity predicts improvement in employment outcomes over time, regardless of the number of years that the IPS programs already implemented IPS or the level of fidelity that they already achieved. The post-hoc descriptive analysis supports these findings indicating that IPS programs with a lower fidelity at the beginning, that improve during the study period also show a relative improvement of employment rate over time, whereas IPS programs with a high and stable fidelity during the study period show a relatively

**Fig. 3** Distribution of descriptive clusters of fidelity improvement for employment rate over time



stable trend of higher employment rate over time. However, as the findings of the post-hoc analysis are only based on descriptive trends we should interpret these outcomes with caution.

We also analyzed the overall fidelity score and the specific item scores over time. The results indicated that the overall fidelity score of the IPS programs significantly improves over time. Analysis of the specific item score showed that the items ‘zero exclusion criteria’ and ‘agency focus on competitive employment’ also improved over time. On the other hand, the items ‘role of employment supervisor’ and ‘frequent employer contact’ had a significantly lower score than average on at least two fidelity assessments.

These findings give some indications of specific components in the implementation of IPS that need some special attention in the Netherlands. The need for improvement for these components is partially confirmed by some qualitative studies about the client and IPS specialist experiences of IPS [27, 28]. First of all, clients confirmed that they found it helpful to have ongoing support and belief from the mental health care team [27], they also appreciated a positive relationship with and ongoing support of the employment specialist [29]. This is especially helpful to feel more confident in the workplace [28]. Furthermore, the focus on competitive employment and employers’ contact has also been an important focus in the improvement of IPS fidelity in the Netherlands. The availability of job coaches was a facilitating and important factor for employers in the decision-making process of employers for hiring employees with intellectual disabilities, if they were responsive to the employer’s need [30, 31]. Clients also confirmed that they benefit from the skills of tailor-made support from the employment specialists [27].

Despite the fact that we gave some indications about performances of specific components of fidelity we were not able to investigate the influence of these specific item scores on employment outcomes. Because of the large number of items alpha inflation is a major concern when analyzing all item scores separately. In addition, the majority of item scores show a limited range, which increases the risk of invalid results. Two studies have already investigated the associations between specific item scores of the IPS-25 fidelity scale and employment outcomes (i.e. [17, 20]). These two studies reported inconsistent results by indicating different fidelity items associated with better employment outcomes. This suggests that individual items on the IPS-25 fidelity scale might not be reliable and consistent predictors for employment outcome, whereas the total scale is modestly predictive, as replicated in multiple studies [17, 20, 21]. The inadequate predictive power for individual items is in line with the classical test theory, presuming that a multi-item scale has greater reliability than individual items, and reliability is an upper bound for validity [32]. Future research

should focus on more quantitative analyses of specific components of the IPS model to give insight into this topic.

This is the first study analyzing longitudinal trends and associations between fidelity and outcome data. Study limitations include the relatively small amount of IPS programs for the number of assessments. This negatively influences the statistical power of the study. Another limitation is in the method of data collection. The employment outcomes were aggregated on a program level by each IPS program. The research team had only access to these data aggregated by each IPS program and was not able to control the quality of data collection. Finally, the fidelity and employment outcome data were collected for quality improvement purposes, and did not receive the intensive data quality review that is commonly implemented in a research study. Furthermore, despite the fact that employment rate outcomes used in this study have been widely adopted for quality improvement and research purposes, the IPS Quarterly Employment Reporting Form [26] is not a validated measure of employment outcomes. Therefore, findings are based on correlations and do not establish causal relationships. The analyses should be considered as valuable, but exploratory.

Based on the current research we can conclude that improvement of implementation, and thereby fidelity, of the IPS model has a significant positive association with employment outcomes over time within the IPS programs regardless of the number of years that the IPS programs already implemented IPS or the level of fidelity that the IPS programs achieve during the study period. These findings contribute to the conclusion that IPS programs, achieving improvement of fidelity to the IPS model over time, can expect better employment outcomes.

## Compliance with ethical standards

**Conflict of interest** The authors declared that they did not have a conflict of interest for this study.

## References

1. Dunn EC, Wewiorski NJ, Rogers ES (2008) The meaning and importance of employment to people in recovery from serious mental illness: results of a qualitative study. *Psychiatry Rehabil J* 32(1):59–62
2. Michon H, van Busschbach JT, Stant AD, van Vugt MD, van Weeghel J, Kroon H (2014) Effectiveness of individual placement and support for people with severe mental illness in the Netherlands: a 30-month randomized controlled trial. *Psychiatry Rehabil J* 37(2):129–136
3. Kortrijk HE, Mulder NL, Kamperman AM, van Weeghel J (2018) Employment rates in flexible assertive community treatment teams in the Netherlands: an observational study. *Commun Ment Health J* 55(2):350–359. <https://doi.org/10.1007/s10597-018-0233-0>
4. Nugter MA, Engelsbel F, Bähler M, Keet R, van Veldhuizen R (2016) Outcomes of FLEXIBLE Assertive Community Treatment

- (FACT) implementation: a prospective real life study. *Commun Ment Health J* 52(8):898–907
5. Van Hoof F, Knispel A, Hulsbosch L, de Lange A, Michon H, Kroon H (2017) *Landelijke Monitor Ambulantisering en Hervorming Langdurige GGZ 2017 (National monitor Deinstitutionalization and reform of long term mental health care 2017)*. Utrecht: Trimbos-instituut. <https://assets.trimbos.nl/docs/3c0ea839-ee7c-4427-b393-8e7819afa323.pdf>
  6. Karpov B, Joffe G, Aaltonen K, Suvisaari J, Baryshnikov I, Näätänen P, Heikkinen M (2017) Level of functioning, perceived work ability, and work status among psychiatric patients with major mental disorders. *Eur Psychiatry* 44:83–89
  7. Suijkerbuijk YB, Schaafsma FG, van Mechelen JC, Ojajarvi A, Corbière M, Anema JR (2017) Interventions for obtaining and maintaining employment in adults with severe mental illness, a network meta-analysis. *Cochrane Database Syst Rev* 9:CD011867. <https://doi.org/10.1002/14651858.CD011867.pub2>
  8. Boardman J (2011) Social exclusion and mental health—how people with mental health problems are disadvantaged: an overview. *Mental Health Soc Inclusion* 15:112–121
  9. Elraz H (2018) Identity, mental health and work: how employees with mental health conditions recount stigma and the pejorative discourse of mental illness. *Hum Relat* 71(5):722–741
  10. Becker DR, Drake RE (2003) *A working life for people with severe mental illness*. Oxford University Press, New York
  11. Drake RE (in press) IPS International: special issue. *Psychiatric Rehabil J*
  12. Brinchmann B, Widding-Havneraas T, Modini M, Rinaldi M, Moe CF, McDaid D, Mykletun A (2019) A meta-regression of the impact of policy on the efficacy of Individual Placement and Support. *Acta Psychiatr Scand*
  13. Metcalfe JD, Drake RE, Bond GR (2017) Economic, labor, and regulatory moderators of the effect of individual placement and support among people with severe mental illness: a systematic review and meta-analysis. *Schizophr Bull* 44(1):22–31
  14. Modini M, Tan L, Brinchmann B, Wang MJ, Killackey E, Glozier N, Harvey SB (2016) Supported employment for people with severe mental illness: systematic review and meta-analysis of the international evidence. *Br J Psychiatry* 209(1):14–22
  15. Bond GR, Drake RE (2019) Assessing the fidelity of evidence-based practices: History and current status of a standardized measurement methodology. *Admin Policy Ment Health Ment Health Serv Res*. <https://doi.org/10.1007/s10488-019-00991-6>
  16. Durlak JA, DuPre EP (2008) Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol* 41(3–4):327–350
  17. Bond GR, Peterson AE, Becker DR, Drake RE (2012) Validation of the revised individual placement and support fidelity scale (IPS-25). *Psychiatric Serv* 63(8):758–763
  18. Hutchinson J, Gilbert D, Papworth R, Boardman J (2018) Implementing supported employment. Lessons from the making IPS work project. *Int J Environ Res Public Health* 15(7):1545–1561
  19. Bond GR, Drake RE, Becker DR, Noel VA (2016) The IPS learning community: a longitudinal study of sustainment, quality, and outcome. *Psychiatric Serv* 67(8):864–869
  20. Kim SJ, Bond GR, Becker DR, Swanson SJ, Langfitt-Reese S (2015) Predictive validity of the individual placement and support fidelity scale (IPS-25): a replication study. *J Vocat Rehabil* 43(3):209–216
  21. Lockett H, Waghorn G, Kydd R, Chant D (2016) Predictive validity of evidence-based practices in supported employment: a systematic review and meta-analysis. *Mental Health Rev J* 21(4):261–281
  22. Van Weeghel J, Bergmans C, Couwenbergh C, Michon H, de Winter L (2019) Individual placement and support in the Netherlands: past, present, and future directions. *Psychiatr Rehabil J*. <https://doi.org/10.1037/prj0000372>
  23. Van Veldhuizen JR (2007) FACT: a dutch version of ACT. *Commun Ment Health J* 43:421–433
  24. Singh SP, Wright C, Burns T, Joyce E, Barnes TRE, Singh S, Burns T (2003) Developing early intervention services in the NHS: a survey to guide workforce and training needs. *Psychiatr Bull* 27(7):254–258
  25. Becker DR, Swanson SJ, Reese SL, Bond GR, McLeman BM (2015) *Supported employment fidelity review manual. A companion guide to the evidence-based IPS Supported Employment Fidelity Scale. 3rd Edition, December 2015*. Retrieved at 18–6–2019 through [https://ipsworks.org/wp-content/uploads/2017/08/ips-fidelity-manual-3rd-edition\\_2-4-16.pdf](https://ipsworks.org/wp-content/uploads/2017/08/ips-fidelity-manual-3rd-edition_2-4-16.pdf)
  26. IPS Employment Center (2019) *IPS quarterly employment reporting form*. Retrieved at 10–14–2019 through <https://ipsworks.org/wp-content/uploads/2018/01/IPS-Learning-Community-Outcomes2018.pdf>
  27. Areberg C, Björkman T, Bejerholm U (2012) Experiences of the individual placement and support approach in persons with severe mental illness. *Scand J Caring Sci* 27(3):589–596
  28. Scott M, Falkmer M, Girdler S, Falkmer T (2015) Viewpoints on factors for successful employment for adults with autism spectrum disorder. *PLoS ONE* 10(10):e0139281
  29. Sveinsdottir V, Bull HC, Evensen S, Reme SE, Knutzen T, Lystad JU (2019) A short history of individual placement and support in Norway. *Psychiatr Rehabil J*. <https://doi.org/10.1037/prj0000366>
  30. Corbière M, Brouwers E, Lanctôt N, van Weeghel J (2014) Employment specialist competencies for supported employment programs. *J Occup Rehabil* 24(3):484–497
  31. Ellenkamp JJ, Brouwers EP, Embregts PJ, Joosen MC, van Weeghel J (2016) Work environment-related factors in obtaining and maintaining work in a competitive employment setting for employees with intellectual disabilities: a systematic review. *J Occup Rehabil* 26(1):56–69
  32. Streiner DL, Norman GR, Cairney J (2015) *Health measurement scales: a practical guide to their development and use*, 5th edn. Oxford University Press, New York
  33. West BT, Welch KB, Galecki AT (2014) *Linear mixed models: a practical guide using statistical software*. CRC Press, New York