Outcomes Star™ Psychometric Factsheet: Justice Star™

Background

The Justice Star is the Outcomes Star for people in the criminal justice system who are on short sentences, approaching release from long sentences or in the community. It was developed by Triangle during 2015-2016 with a diverse group of collaborators:

- Barnardo's Family Support
- Lancashire Women's Centres
- Leicestershire Police
- Lifeline
- North London Forensic Service
- Shaw Trust
- Sodexo Justice Services
- The Fortune Society (USA)
- UnitingCare West (Australia).

More information about the development of the Justice Star can be found in the Development Report (Burns, MacKeith & Lamont, 2017) and the overall principles behind the development of all versions of the Outcomes Star are described in MacKeith (2011).

Method and analytic strategy

Justice Star data routinely collected and entered onto the Star Online was analysed by Triangle to test the Star's validity as an outcomes measurement tool. A full explanation of the analytic strategy is provided in the accompanying document – Outcomes Star Psychometric Factsheets: Overview.

With the exception of inter-rater reliability, these psychometric tests were conducted using anonymised data (N = 368) collected by an organisation working with formerly incarcerated people in New York. The average time between 1st and 2nd Star readings was 89 days. Inter-rater reliability was assessed using a sample of 50 workers who were using the Justice Star in services run by a Community rehabilitation company in the UK.

Results

Do service users and workers view the Star as appropriate and useful?

Acceptability and Content validity:

Feedback questionnaires were received from 60 service users and 46 staff.

Almost all of service users (93%) reported that the Justice Star 'helped them to describe how life is for them' and that it provided a good/very good summary of their current life and needs (98%). 95% of workers felt that the Justice Star described the client base fairly or very well.

Does it make sense for the different outcome areas of the Star to be included in the same tool?

Factor Structure: The Kaiser-Meyer-Olkin value exceeded the recommended minimum value of 0.60 (Kaiser, 1970, 1974) and a significant Bartlett's Test of Sphericity (Bartlett, 1954) supported the suitability of the data for factor analysis. The analysis yielded a unidimensional factor structure explaining 48% of the variance in the data.

Inter-item correlations between Parenting and caring, Drugs and alcohol and A crime free life and the other areas were lower than is typically found in other versions of the Outcomes Star. For Parenting and caring this may be explained by service users being given a reading of 10 if they do not have children. In the Drugs and alcohol area, very few service users were below 5 (Motivated and taking responsibility), which may reflect a reporting bias.   As with all versions of the Star, decisions about the inclusion of outcome areas are strongly influenced by the practical needs of service users and those supporting them.

Internal Consistency Internal consistency was very good (Cronbach's α =.78).

Is each outcome area measuring a unique aspect of the service user's situation?

Item redundancy: No inter-item correlation exceeded the 0.7 threshold, suggesting no redundancy between areas (see Table 1).

Does the Star detect change occurring within a service?

Responsiveness to change: The Wilcoxen Signed Rank Test revealed a statistically significant increase in all outcome areas (see Table 2), with medium effect sizes in all areas except Parenting and caring. In the Parenting and Caring area, service users are given a reading of 10 if they do not have children and in this dataset one third of service users began at 10. When service users beginning at 10 (who cannot move forward) were excluded from the analyses, the effect size increased from 0.18 to 0.33 (medium).

Do workers have a consistent understanding of how to apply the scales?

Inter-rater reliability: The average Krippendorff's α of .83 across the cohort indicates very good inter-rater reliability, with only 8 of the 50 workers (16%) falling below the 0.67 threshold.

Conclusions

The results of these initial analyses are encouraging and suggest that the Justice Star is a valid outcomes measurement tool, with high acceptability and content validity, a unidimensional factor structure, good responsiveness and inter-rater reliability.

Additional research

Research is planned to examine the relationship between Star readings and other measures (convergent and predictive validity).
External research about the Star as an outcomes and keywork measure can be found on our website:
http://www.outcomesstar.org.uk/about-the-star/evidence-and-research/research-library/#all

Table 1. Correlation matrix for outcome areas (*N* =368)

	1	2	3	4	5	6	7	8	9
1 Accommodation									
2 Living skills & self-care	.54								
3 Mental health & well-being	.44	.53							
4 Friends & community	.37	.51	.54						
5 Relationships & family	.54	.44	.56	.45					
6 Parenting & caring	.10	.08	.09	.17	.12				
7 Drugs & alcohol	.23	.34	.25	.28	.25	.10			
8 Positive use of time	.44	.55	.47	.52	.47	.26	.28		
9 Managing strong feelings	.39	.48	.57	.48	.46	.08	.34	.44	
10 A crime-free life	.15	.15	.07	.20	.13	.12	.30	.37	.16

Table 2. Responsiveness of the Justice Star (*N* =368)

	First Star median	Final Star median	Z	Effect size r^1
Accommodation	8.00	9.00	-9.81***	0.36
Living skills & self-care	8.00	9.00	-9.36***	0.35
Mental health & well-being	9.00	9.00	-7.77***	0.29
Friends & community	7.00	8.00	-10.18***	0.38
Relationships & family	8.00	9.00	-9.43***	0.35
Parenting & caring	7.00	9.00	-4.87***	0.18[2]
Drugs & alcohol	7.00	9.00	-10.92***	0.40
Positive use of time	8.00	9.00	-9.84***	0.36
Managing strong feelings	8.00	9.00	-9.43***	0.35
A crime-free life	7.00	8.00	-12.52***	0.46

*** $p < .001$

[1] Cohen provided rules of thumb for interpreting these effect sizes, suggesting that an *r* of .1 represents a 'small' effect size, .3 represents a 'medium' effect size and .5 represents a 'large' effect size
[2] This increases to 0.33 when the 33% of service users beginning at 10 are excluded

References

Bartlett, M. S. (1954). A note on the multiplying factors for various χ 2 approximations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 296-298.

MacKeith, J., Burns, S. & Lamont, E. (2017) <u>The Justice Star Development Report</u>, Brighton: Triangle Consulting

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*(4), 401-415.

Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and psychological measurement*, *34*(1), 111-117.

MacKeith, J., (2014). Assessing the reliability of the Outcomes Star in research and practice. *Housing, Care and Support*, *17*(4), 188-197.