Signs of Safety Evaluation - Addendum

About

This addendum adds to the existing evaluation report of Signs of Safety available on gov.uk.[1] More information about the project can be found on the report page.[2]

About Works for Early Intervention and Children's Social Care

What Works for Children's Social Care (WWCSC) and the Early Intervention Foundation (EIF) are merging. The new organisation is operating initially under the working name of What Works for Early Intervention and Children's Social Care. Our new single What Works centre will cover the full range of support for children and families from preventative approaches, early intervention and targeted support for those at risk of poor outcomes, through to support for children with a social worker, children in care and care leavers.

Introduction

What Works for Children's Social Care (WWCSC) was commissioned by the Department for Education to examine the following research questions on the impact of Signs of Safety (SofS):

- What, if any, is the impact of SofS on the duration of assessments?
- What, if any, is the impact of SofS on the likelihood of an initial child protection conference (ICPC for children who have already been assessed and whose case is designated as open?
- What, if any, is the impact of SofS on the likelihood of a case being re-referred if it has previously been assessed as 'no further action' (NFA)?
- What, if any, is the impact of SofS on the likelihood of a re-referral leading to a child protection plan (CPP) or to a child becoming looked after (CLA)?
- What, if any, is the impact of SofS on the likelihood of a child receiving kinship care instead of non-kinship care?

This addendum details the findings of the evaluation of Signs of Safety on its impact on initial child protection conferences (ICPC) and an exploratory analysis of Signs of Safety's impact on entry to care rates. The findings concerning the remaining research questions are discussed in Baginsky et al. (2020).[3]

1
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/956625/Sofs_revised_evaluation_report_270121.pdf
[2] https://whatworks-csc.org.uk/research-report/impact-of-signs-of-safety-on-outcomes-for-children/
[3] See https://whatworks-csc.org.uk/wp-content/uploads/Sofs_revised_evaluation_report_270121.pdf.

Analytical approach

We used routinely collected data on nine local authorities piloting SofS as part of the Department for Education's English Innovation Programme. We compared the outcomes of children in local authorities who use SofS with the outcomes of children in similar local authorities who do not, using a matched difference-in-differences (DiD) design. Comparator local authorities were identified based on parallel trends in outcomes before SofS was implemented in the pilot sites. Data was accessed via the ONS' secure research service (SRS). More information on the analytical approach can be found in the trial protocol.[4]

Difference-in-differences

When assessing whether a new approach is working, we could measure the outcomes for children and young people before and after and see whether they've improved. The trouble with this is that any changes you see before and after the approach is implemented could be due to other changes that are happening over time that are not related to the approach (e.g. the local authority's finances improving or the support at schools for vulnerable children improving). A difference-in-differences approach solves this problem by comparing how outcomes change over time for those affected by the intervention and how the outcomes change over time for those not affected by the new approach. Looking at the difference in outcomes over time for those not affected simulates what we would expect in the absence of the new intervention.

The exploratory analysis was conducted in light of the findings of Baginsky et al. (2020) that children were less likely to receive kinship care instead of non-kinship care after the introduction of SofS, to explore whether this finding can be explained by changes in entry to care rates.

---

[4] https://whatworks-csc.org.uk/wp-content/uploads/WWCSC-Signs-of-Safety-Trial-Protocol-updated-v2.pdf

Key findings

For the two research questions that this addendum covers, we find:
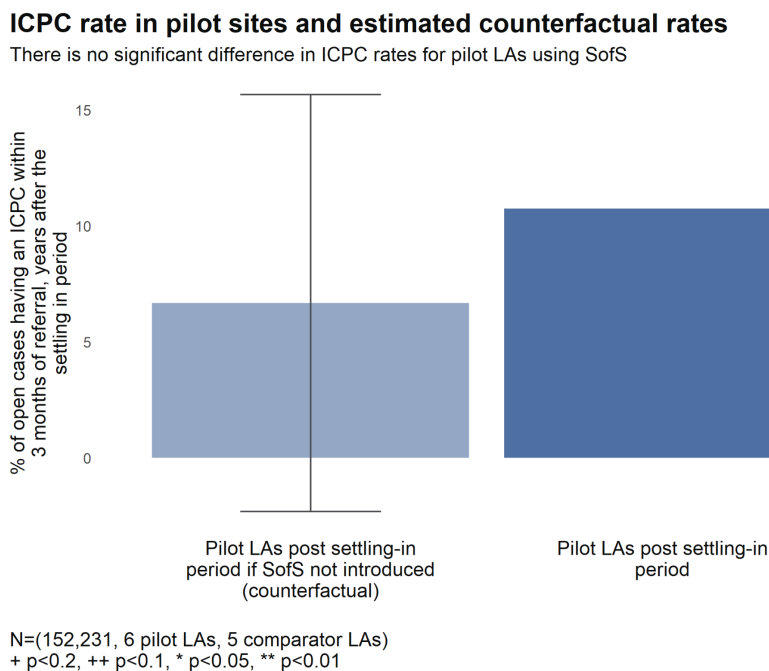
- No evidence that the introduction of SofS significantly affected the probability of initial child protection conferences (ICPC)
- No evidence that the introduction of SofS significantly affected the probability of children entering care in the nine pilot sites.

Our research suggests that Signs of Safety did not, on average, affect these two outcomes across the nine pilot sites. As the analysis focuses on outcomes across all nine pilot sites, it focuses on the average effect of SofS rather than potential impacts of SofS on individual local authorities. Detailed results for both outcomes are described below. The findings concerning the first research questions outlined above are discussed in Baginsky et al. (2020).

ICPC rates

We found no evidence that the implementation of SofS affected the probability of an ICPC for children and young people who had an open case. **The analysis suggests that SofS had no significant impact on the rate of children and young people having an ICPC within three months of their referral date.** The results remain robust to different model specifications and sensitivity analyses. The assumptions in our statistical model were met, which allowed a causal interpretation of results. More details on the analysis results can be found in technical appendix A.[5]

Figure 1: Impact of SofS on the ICPC rate

ICPC rate in pilot sites and estimated counterfactual rates

There is no significant difference in ICPC rates for pilot LAs using SofS

N=(152,231, 6 pilot LAs, 5 comparator LAs)
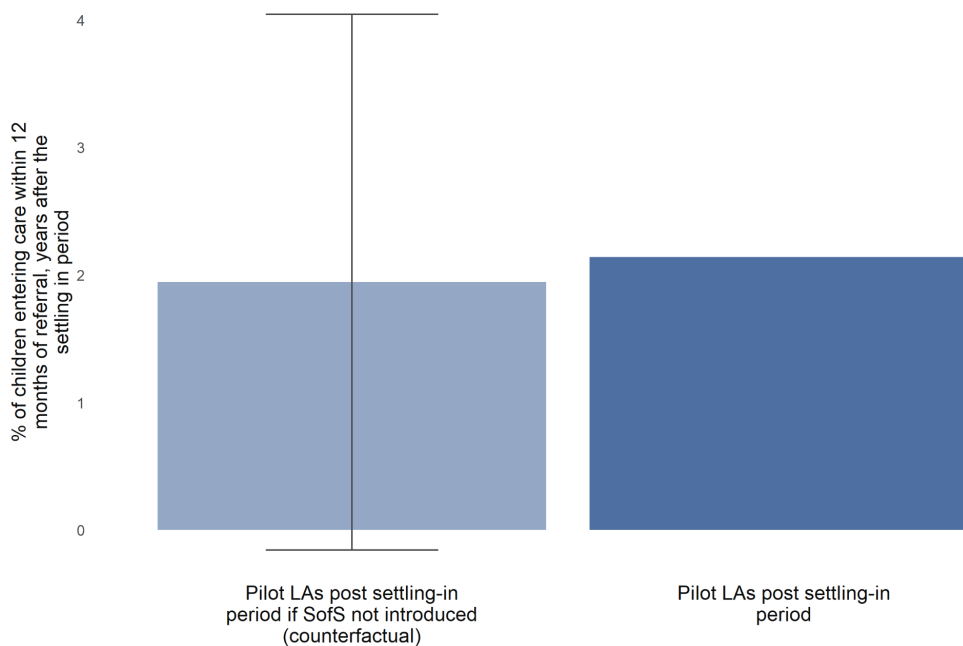+ p<0.2, ++ p<0.1, * p<0.05, ** p<0.01

Entry to care

The main analysis (Baginsky et al., 2020) found that kinship care decreased in local authorities that implemented SofS, contrary to the theory of change. To explore potential explanations of this finding, we conducted an exploratory analysis to investigate whether changes in the use of kinship care can be explained by changes in the rate of children entering care. **We did not find any evidence that the introduction of SofS affected the probability of children entering care** and thus cannot shed further light on the dynamics behind the kinship care finding. The analysis suggests that SofS had no significant impact on the rate of children and young people going into care within twelve months of their referral date. The results remain robust to different model specifications and sensitivity analyses. The assumptions in our statistical model were met, which allowed a causal interpretation of results. More details on the analysis results can be found in technical appendix A.

Figure 2: Impact of SofS on the rate of entry to care

CLA rate in pilot sites and estimated counterfactual rates

There is no significant difference in CLA rates for pilot LAs using SofS

N=(233,637, 7 pilot LAs, 6 comparator LAs)
+ p<0.2, ++ p<0.1, * p<0.05, ** p<0.01

Technical Appendix

Full analysis results

This section details the full analysis results for the two outcomes. The analytical approach and all associated decision rules are described in more detail in the trial protocol[6] as well as the technical report.[7]

Matching results

We matched local authorities that implemented SofS with local authorities that did not implement SofS and that exhibited similar trends in the rate of ICPCs and the rate of children looked after (CLA), respectively, before the implementation of SofS in the pilot sites. Upon conclusion of the local authority level matching, we used the following pilots for our main analysis. This is the result of the local authority level matching on pre-SofS trends of the outcome variable (or a proxy thereof).

Table 1: Key to pilot sites' involvement in each evaluation question after matching

Pilots	Outcome: ICPC rate	Outcome: CLA rate
Pilot 1	✓	✓
Pilot 2	✓	✓
Pilot 3	x	x
Pilot 4	✓	x
Pilot 5	✓	✓
Pilot 6	✓	✓
Pilot 7	x	✓
Pilot 8	x	✓
Pilot 9	✓	✓

Within the groups of matched LAs, we match individual children using coarsened exact matching (CEM) to provide a more efficient estimate of the impact of SofS. We assessed the quality of the resulting balance using the multivariate imbalance scores (where 0 represents a perfectly balanced sample) and the local common support (where 100% represents that all children in pilots have a counterpart with the same demographics in a comparator LA). The resulting multivariate imbalance scores for the different populations are reported in Table 2 below. The scores are calculated using

---

[6] https://whatworks-csc.org.uk/wp-content/uploads/WWCSC-Signs-of-Safety-Trial-Protocol-updated-v2.pdf

[7] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/956624/Signs_of_Safety_revised_evaluation_appendices_January_2021__2_.pdf

the uncoarsened dataset. After matching, the imbalance score of the coarsened dataset would be equal to zero. The matching improves the balance of the data but does not fully account for the differences between the pilot and comparator LAs so we also control for individual-level covariates in the regressions.

Table 2: Multivariate imbalance scores and local common support pre and post CEM

Evaluation question	L1 imbalance score		Local common support	
	Prior to CEM	**Post CEM**	**Prior to CEM**	**Post CEM**
ICPC rates	0.61	0.58	40%	58%
CLA rates	0.52	0.5	49%	65%

Data source: Office for National Statistics – National Pupil Database. The L1 imbalance score and the local common support are calculated for the original population prior to matching, and for the matched population after conclusion of the CEM for the uncoarsened variables. Note that the imbalance score is 0 when it is calculated using the coarsened variable

Summary statistics

Since we use different samples and different pilot LA comparator LA matches for the two research questions we sought to answer, the sample size and covariate balance varies between the relevant samples. Table 3 gives an overview of the sample for the evaluation question on ICPC rates. The sample consists of any child that was referred within the relevant time period whose case was designated as open after assessment. Means are only reported for the time frame before SofS was implemented, since there might be changes in the composition of the population in pilot sites due to the implementation of SofS.

While the pilot sites and comparator LAs seem very similar on most covariates, the samples differ with respect to the baseline rate of the outcome and the first category of need prior to the settling-in period.

Table 3: Weighted pre-treatment summary statistics for the two evaluation questions broken down by pilots and comparators

Variables	ICPC rates - Pilot sites	ICPC rates - Comparator sites	CLA rates - Pilot sites	CLA rates - Comparator sites
Outcome	8%	11%	2%	4%
Gender – male	51%	51%	51%	51%
Gender – female	47%	47%	48%	48%
Gender – not recorded/unborn	-	-	-	-
Gender – missing	-	-	-	-
Age	7.46	7.45	7.83	7.83
Disabled	15%	15%	15%	15%
Academic year – secondary school	28%	28%	30%	30%
Academic year – primary school	36%	36%	38%	38%
Academic year – before school age	35%	35%	31%	31%
Academic year – missing	1%	1%	1%	1%
Low income (measured by FSM) – no	13%	13%	15%	15%
Low income (measured by FSM) – yes	30%	30%	38%	38%
Low income (measured by FSM) – below school age so no recorded	32%	32%	27%	27%
Low income – missing	25%	25%	19%	19%
Ethnicity – any other ethnic group	-	-	-	-

Ethnicity – Asian	3%	3%	3%	3%
Ethnicity – Black	8%	8%	6%	6%
Ethnicity - Chinese	-	-	-	-
Ethnicity – missing	6%	6%	7%	7%
Ethnicity – mixed	4%	4%	5%	5%
Ethnicity – unclassified	3%	3%	4%	4%
Ethnicity – White	75%	75%	75%	75%
Main need – not stated	7%	1%	23%	3%
Main need – abuse or neglect	50%	36%	40%	51%
Main need – child's disability/illness	3%	7%	5%	7%
Main need – parental disability/illness	2%	2%	2%	3%
Main need – family in acute stress	9%	10%	9%	10%
Main need – family dysfunction	26%	36%	14%	20%
Main need - low income	1%	4%	2%	3%
Main need – socially unacceptable behaviour	-	-	-	-
Main need – absent parenting	1%	-	2%	2%
Main need – cases other than children in need	-	2%	2%	1%
Main specification (which determines sample size)	Linear	Linear	Logistic	Logistic

Number of observations in this pre-treatment group	24,764	59,774	80,243	71,666
Number of observations of treatment and comparator LAs over all periods	152,231	152,231	233,637	233,637

Source: Office for National Statistics – National Pupil Database (April 2008-March 2019). Population consists of all children referred during the observation period who have been assessed and whose case was designated as open after assessment. Numbers with '-' are negligible and/or suppressed due to statistical disclosure reasons. Percentages are rounded to the nearest whole number and so categories may add to greater than 100%. NA is stated where the variable was not used. Where all values in a row are NA or '-', the row is omitted. All summary statistics are weighted statistics.

Analysis results

For the ICPC analysis, the assumptions of our statistical model were met (three out of four leads were not significant at the five per cent level), allowing a causal interpretation of the results. Our main specification is the linear model, as the incidence of ICPCs in our sample is 11% and thus above the cutoff of 10%. More details on the chosen decision rules and general analytical approach can be found in the trial protocol.[8] Since the Breusch-Godfrey test is not significant, we do not consider any multilevel models.

For the analysis of CLA rates, the assumptions of our statistical model were met (five out of five leads were not significant at the five per cent level), allowing a causal interpretation of the results. Given the low incidence of children entering care out of all children referred (2%), our main specification is the

---

[8] https://whatworks-csc.org.uk/wp-content/uploads/WWCSC-Signs-of-Safety-Trial-Protocol-updated-v2.pdf

logistic regression. Since the Breusch-Godfrey test is significant, we use cluster robust standard errors clustered at the individual level as an additional specification.[9]

Table 4: Test statistics for various tests conducted for each evaluation question

Test	Null hypothesis	Result for outcome: ICPC rate	Result for outcome: CLA rate
Breusch-Godfrey test	No serial correlation of order 1 in the errors	No serial correlation detected (p=0.77)	Serial correlation detected (p=0.01)
Number of insignificant leads	Parallel trends test – the majority of treatment leads should be insignificant for parallel trends assumption to hold	3 out of 4 leads are insignificant	5 out of 5 leads are insignificant

Source: Office for National Statistics – National Pupil Database (April 2008–March 2019)

Rate of Initial Child Protection Conferences (ICPC): Primary analysis

The analysis suggests that SofS did not significantly affect the likelihood of an ICPC taking place within three months of the referral date. This finding is consistent across different specifications. None of the lagged treatment effects is significant at the five per cent level, and excluding pilot 4 from the analysis does not affect the significance of the treatment effect.

---

[9] While the test for serial correlation is significant, we refrain from using fixed effects regressions as set out in the analysis plan. As the likelihood of recurring in the sample will be significantly correlated with the outcome measure (entering care), a fixed effects regression only covering a part of the original sample population is not suitable in this context.

Table 5: DiD regression table – estimating the impact of SofS on the rate of ICPCs (primary analysis)

R² =0.0398	Linear model	Lagged treatment	Excluding pilot 4
(Intercept)	0.1368**	0.1619***	0.2138***
Pilot LA	-0.0635*	-0.0739**	-0.0765**
DiD (pilot LA * post settling-in period)	0.0407	N/A	0.0389
Year dummy 2013	0.0075	0.004	0.0128
Year dummy 2014	0.057	0.0568	-0.0262*
Year dummy 2015	0.0726***	0.0679***	N/A
Year dummy 2016	0.0043	-0.0121	-0.0009
Year dummy 2017	0.0468	0.0427	0.0414
Year dummy 2018	0.0657***	0.0831***	0.0638***
DiD in year t	N/A	0.0704	N/A
DiD in year t+1	N/A	-0.0056	N/A
DiD in year t+2	N/A	-0.0494	N/A

Source: Office for National Statistics – National Pupil Database (April 2008 – March 2019)

Population: all children assessed and with open cases. N=152,231, 6 pilot sites, 5 comparator sites. Asterisks indicate p-values: *p<0.1, **p<0.05, ***p<0.01

ICPC rates: Sensitivity analysis

Pilot 4 started using SofS prior to the Innovation Programme and the qualitative work by Dr Mary Baginsky suggested that the experience of SofS in Pilot 4 may not be representative of the experience of the other pilots. We excluded it from the analysis to check whether the results were robust to its exclusion. When excluding pilot 4, the coefficient for the treatment effect remains not significant, further supporting the main findings

ICPC rates: Secondary analysis

We found no significant impact of the implementation of SofS on ICPC rates in the alternative models we used in the secondary analysis. **Even when controlling for the degree of embeddedness or the quality of delivery of Signs of Safety, the treatment coefficient remains not significant.**

There is also **no indication that there is an effect of SofS on ICPC rates that varies by Ofsted rating o**f the local authority implementing the model, or that the presence of alternative practice models (Restorative Practice or Reclaiming Social Work) significantly biased downwards any potential effect of SofS. While controlling for alternative practice models decreases the p-value of the main coefficient of interest, the treatment effect is not significant at the five per cent level.

Table 6: DiD regression table – estimating the impact of SofS on the rate of ICPCs (secondary analysis)

Variables	Quality of delivery	Embeddedness of SofS	Alternative practice models	Ofsted
(Intercept)	0.0974*	0.1005**	0.0488	0.1213**
Pilot site	N/A	N/A	N/A	-0.0947***
Year dummy 2013	0.0206*	0.0197	0.01	0.0209*
Year dummy 2014	0.0593*	0.0559*	0.0376	0.0566*
Year dummy 2015	0.0657***	0.0597***	0.03	0.0603***
Year dummy 2016	-0.0167	-0.0221	0.0078	-0.0328
Year dummy 2017	0.0259	0.0192	0.05	0.0069
Year dummy 2018	0.0488***	0.0423***	0.0802**	0.031
DiD (pilot LA * post settling-in period)	N/A	N/A	N/A	0.0559
Quality of delivery	-0.0882**	N/A	N/A	N/A
Quality of delivery * post-settling-in period	0.0511	N/A	N/A	N/A
Embeddedness score	N/A	-0.1238***	N/A	N/A
Embeddedness score * post settling-in period	N/A	0.0948	N/A	N/A
Comparator LA with a similar practice model	N/A	N/A	0.0951**	N/A
Comparator LA without a similar practice model	N/A	N/A	0.0626***	N/A
Comparator LA with a similar practice model * post settling-in period	N/A	N/A	-0.0907*	N/A

Comparator LA without a similar practice model * post settling-in period	N/A	N/A	-0.0074	N/A
Ofsted rating	N/A	N/A	N/A	-0.0909***
Pilot LA * post settling-in period * Ofsted rating	N/A	N/A	N/A	0.0311

Source: Office for National Statistics – National Pupil Database (April 2008 – March 2019)
Population: all children assessed with open cases, N=152,231, 6 pilot sites, 5 comparator sites.
Asterisks indicate p-values: *p<0.1, **p<0.05, ***p<0.01

CLA rates: Primary analysis

The analysis suggests that SofS did not significantly affect the likelihood of children entering care within 12 months of the referral date. This finding is consistent across different specifications. None of the lagged treatment effects is significant at the five per cent level, and additional specifications such as random effects models and introducing LA fixed effects leave the DiD estimate not significant.

Table 7: DiD regression table – estimating the impact of SofS on the rate of entry to care (primary analysis)

$R^2 = 0.0284$	Logistic regression - clustered	Linear regression - clustered	Logistic regression - not clustered	Lagged treatment - logistic regression
(Intercept)	0.99*	0.13***	0.99	1.18
Pilot LA	-0.30***	-0.01***	-0.30	-0.19
DiD (pilot LA * post settling-in period)	0.08	0.004	0.08	N/A
Year dummy 2009	-0.09	-0.002	-0.0 6	-0.10
Year dummy 2010	-0.52	0	-0.09	-0.19
Year dummy 2011	-0.53***	-0.01***	-0.52**	-0.64***
Year dummy 2012	-0.31***	-0.01***	-0.53*	-0.66**
Year dummy 2013	-0.31*	-0.004	-0.31	-0.44
Year dummy 2016	-0.50***	-0.02***	-0.50	-0.67
Year dummy 2017	-0.04	-0.01**	-0.04	-0.14
Year dummy 2018	-0.44**	-0.02***	-0.44	-0.33
DiD in year t	N/A	N/A	N/A	0.12
DiD in year t+1	N/A	N/A	N/A	-0.35
DiD in year t+2	N/A	N/A	N/A	-1.47

Source: Office for National Statistics - National Pupil Database (April 2008 – March 2019)
Population: all children referred. N=233,637, 7 pilot sites, 6 comparator sites.
Asterisks indicate p-values: *p<0.1, **p<0.05, ***p<0.01

CLA rates: Sensitivity analysis

The sensitivity analysis set out in the trial protocol does not apply to this research question, as pilot 4 has not been used as a pilot site due to a lack of suitable comparator sites and the sensitivity analysis of excluding pilot sites 1 and 2 is not applicable as it concerns the assessment stage and is thus not relevant for the outcome in question.

CLA rates: Secondary analysis

As the CLA analysis was of an exploratory nature, we limit the analysis to the primary analysis set out in the trial protocol, and did not conduct any secondary analysis.