



HM Prison &
Probation Service

An evaluation of a brief intervention to protect good decision-making in on-the-day sentencing proposals to courts

A randomised control trial

Georgia Barnett and Flora Fitzalan Howard

Evidence-Based Practice Team, Insights Group, HMPPS

Ministry of Justice Analytical Series
2023



His Majesty's Prison and Probation Service is committed to evidence-based practice informed by high-quality social research and statistical analysis. We aim to contribute to the informed debate on effective practice with the people in our care in prisons, probation and youth custody.

Disclaimer

The views expressed are those of the authors and are not necessarily shared by the Ministry of Justice (nor do they represent Government policy).

First published 2023



© Crown copyright 2023

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Any enquiries regarding this publication should be sent to us at National.Research@noms.gsi.gov.uk

This publication is available for download at <http://www.justice.gov.uk/publications/research-and-analysis/moj>

ISBN 978-1-911691-13-6

Acknowledgements

We would like to thank the probation staff who agreed to take part in this trial at a particularly challenging time, while living and working through a global pandemic. Probation court report writers do a demanding job in often pressured circumstances, carefully balancing the needs of the courts, the public and the defendants they're assessing. We are extremely grateful for your commitment to this research, and for the hard work you do, day in and day out, to protect the public and give people the best chance of desisting from crime.

We would also like to thank the court Senior Probation Officers, who helped to get staff on board, collate information and generally help in any way they could, and in particular, the staff at Luton and Northampton courts whose help with the pilot of the task we trialled was invaluable in shaping the final intervention. Special thanks also go to Fuschia Booth and the Pre-sentence Report pilot team, in particular James Tennant, Doreen Taylor, Andy Harris, Zoe Walker, Michelle Szyman, Justine Rhodes, Beth Shelley and Nicola Morant whose help and support with the trial was invaluable.

We also thank Kate Netten and Lydia Baxter for their significant support with data collection and coding.

The authors

Georgia Barnett and Flora Fitzalan Howard are forensic psychologists who work in the HMPPS Evidence-Based Practice Team in the Insights Group, a team that aims to enable and embed evidence-based practice across the Prison, Probation and Youth Custody Services.

Contents

1. Summary	1
2. Introduction	4
2.1 Context	4
2.2 Review of the literature	5
2.3 Research questions and hypotheses	7
3. Method	8
3.1 Sample	8
3.2 Measures	9
3.3 Materials	12
3.4 Procedure	13
3.5 Analysis	14
3.6 Limitations	14
4. Results	17
4.1 Efficacy of randomisation	17
4.2 Efficacy of Implementation	17
4.3 Attrition	18
4.4 Time taken to complete the task	18
4.5 Sentencing proposals	18
4.6 The influence of ethnicity on sentencing proposals	19
4.7 Impact of the intervention on proposal weight and orientation	21
4.8 Report writers' views of the intervention	26
5. Discussion and implications	28
5.1 Summary of findings	28
5.2 Conclusion	30
5.3 Future research	31
References	32
Appendix A	35
Ethical Considerations	35

Appendix B	37
Flow diagram of sample attrition	37
Appendix C	38
Sentencing Proposal Coding Framework	38
Appendix D	40
Questions in the experimental task	40
Appendix E	42
Randomisation and trial protocol	42
Appendix F	43
Semi-structured interview schedule	43
Appendix G	44
Characteristics of the report writers in the experimental and control groups	44
Appendix H	46
Characteristics of report subjects for reports written under experimental and control conditions	46
Appendix I	48
Results of multinomial regression analyses examining the influence of ethnicity on orientation and weight of sentencing proposals	48
Appendix J	53
Results of multinomial regression analyses examining the influence of the intervention, and interaction between the intervention and report subjects' ethnicity, on orientation and weight of sentencing proposals	53
Appendix K	57
Results of multinomial regression analyses examining the influence of completing the intervention as intended on orientation and weight of sentencing proposals	57

List of tables

Table 1. Characteristics of probation report writers in the study	8
Table 5. Orientation and weight of sentencing proposal by report subjects' ethnic group.	19
Table 6. Comparison of the orientation and weight of sentencing proposals put forward to the courts in pre-sentence reports for those report subjects of ethnic minority or white background.	20
Table 13. Comparison of the orientation and weight of the initial and alternative sentencing proposals put forward to the courts in pre-sentence reports.	24
Table 14. Number and proportion of report subjects in different ethnic groups by intervention condition with different orientation and weight of sentencing proposals in on-the-day pre-sentence reports.	25
Table 15. Comparison of number and proportion of reports produced under the intervention condition which complied or failed to comply with the experimental task, on people from different ethnic groups, with different orientation and weight of sentencing proposals in on-the-day pre-sentence reports.	26
Table 2: Sanctions grouped by rehabilitative orientation and weight	38
Table 3. Characteristics of report writers in the experimental and control conditions	44
Table 4. Characteristics of the report subjects in the experimental and control conditions	46
Table 7. Summary of multinomial logistic regression analysis examining impact of ethnicity on sentencing proposal orientation	48
Table 8. Summary of multinomial logistic regression analysis examining impact of ethnicity on sentencing proposal weight	50
Table 9. Summary of multinomial logistic regression analysis examining impact of intervention, and interaction between intervention and ethnicity, on sentencing proposal orientation	53
Table 10. Summary of multinomial logistic regression analysis examining impact of intervention, and interaction between intervention and ethnicity, on sentencing proposal weight	55
Table 11. Summary of multinomial logistic regression analysis examining impact of the intervention when completed as intended on sentencing proposal orientation	57
Table 12. Summary of multinomial logistic regression analysis examining impact of the intervention when completed as intended on sentencing proposal weight	59

List of figures

Figure 1. CONSORT flow diagram of sample attrition at each stage of the trial 37

1. Summary

Background

External scrutiny bodies, academic research, and practitioners have raised concerns about the potential impact of speedy reporting practices and poor-quality reports for people from ethnic minorities on sentencing recommendations. The Government's Smarter Sentencing White Paper, published in 2020, set out a commitment to improve pre-sentence reports. This report documents the results of a randomised control trial that tested a brief intervention aiming to protect against disproportionality in sentencing proposals for people from ethnic minorities in on-the-day pre-sentence reports.

Method

One hundred and eight report writers from 14 courts in England and Wales took part in the trial, producing 985 on-the-day pre-sentence reports over the course of 10 months. Report writers were randomly assigned to one of two groups. The first (control) group wrote reports following usual, standard practice. The second (experimental) group undertook an additional 'consider the alternative' task prior to making their sentencing recommendation. The task was designed to interfere with mental shortcuts and counter a tendency to revert to the same few proposals they might commonly use. The task asked report writers to identify and make the best argument for a feasible alternative proposal to the one they first thought of, before deciding which of the two to recommend to the court. It was expected that, if the intervention worked, the task would have more of an impact on proposals put forward for people from ethnic minority groups than for white people. A small number of report writers were also interviewed at the end of the trial, to understand their experience of completing the task, barriers and enablers, and recommendations for developing practice.

Findings

The task took on average seven minutes to complete, however, a quarter of those assigned to do the intervention prior to making a sentencing proposal to court failed to do so for any of their reports. Just over half completed the intervention as intended for some, but not all, of their reports. This suggests that while it is possible for

probation report writers to implement the task as intended in a busy court setting, there are significant issues with consistent and reliable application of the task.

Interviews with 11 of the report writers in the intervention condition indicated that while they judged the intervention to be quick and easy to complete, those with more experience of writing court reports felt that this was unnecessary and difficult to fit into the time constraints of on-the-day reports. The three people interviewed who were newer to writing pre-sentence reports were more positive about the task, and felt it added value, prompting them to reflect on the sentencing proposal they were putting forward to the courts. This suggests that any such addition to the on-the-day reporting process would need to be preceded by a concerted effort to communicate the nature and purpose of the task to facilitate buy in, with a focus on those more experienced staff.

In the sample of reports produced for this trial, there was some evidence of disproportionality in the nature of the sentencing proposals being put forward to courts for people from different ethnicities. More specifically, after controlling for additional variables that might influence the nature of sentencing proposals, people from white backgrounds were more likely to receive a proposal oriented more heavily to rehabilitation, rather than punishment, than were people from ethnic minorities.

The intervention had more of an effect on sentencing recommendations put forward to courts for people from ethnic minorities than for white people, but not in the way expected. While the task had no impact on how burdensome that proposal was likely to feel (the weight of the proposal), it did have a differential impact on the orientation of the proposals put forward for people from ethnic minorities. Reports written under the intervention condition on people from ethnic minorities were more likely to recommend a sentence containing a predominantly punitive sanction, than a sentence that was predominantly rehabilitative or equally rehabilitative and punitive. This effect was most pronounced in reports written by report writers who failed to comply with the intervention.

It is possible that the introduction of the task triggered psychological reactance among some of the report writers in the intervention condition who saw the task as

unnecessary.¹ This could help explain the high rates of non-compliance with the intervention, and which may have increased the salience and subsequently the influence of the ethnicity of the report subject on the sentencing recommendation. Indeed, the association between ethnicity and orientation of the sentencing proposal in the on-the-day pre-sentence report was stronger among those in the intervention condition who did not, compared with those who did, comply with the task. Alternatively, it is possible that those failing to comply with the intervention were more biased or less likely to notice and act on their biases, than those who completed the task.

Limitations

A key limitation is the low levels of compliance with the intervention task; only 58.5% of the reports written under the intervention condition were completed following full and proper completion of the task. In addition, the lack of a control task means it is possible that those in the intervention condition were under greater time pressure when completing on-the-day reports, than those in the control condition. The former had to complete an additional (albeit relatively brief) task during an already tight time frame for the production of on-the-day reports. Finally, only 10% of the report writers were interviewed to ascertain issues with implementation of the intervention.

Conclusion

The trial found that the intervention was not implemented as intended in a real world setting and failed to have the desired effect on the sentences proposed to the courts in on-the-day pre-sentence reports. Most of the probation report writers interviewed felt that the intervention was both unnecessary and difficult to fit into the limited time they had to complete on-the-day reports. The fact that report writers found it challenging to fit a short task into their practice emphasises the extreme time constraints under which these reports are written. The findings also underscore the importance of the way in which anti-discrimination initiatives are introduced, to avoid potential backfire effects. The study raises concerns about the impact of speedy reporting practices on the proportionality of sentencing proposals put forward to courts for people from different ethnic groups.

¹ Reactance is an unpleasant motivational arousal that emerges when people experience a threat to or loss of their freedom to act in the way they choose. It serves as a motivator to restore one's freedom.

2. Introduction

2.1 Context

Pre-sentence reports (PSRs) are assessments conducted by probation staff that provide an independent recommendation about the options available to the court for the sentencing of a person convicted of crime. As well as supporting effective decision-making in the courts, PSRs form the basis of risk management and sentence plans and are used to inform decisions by appeal courts and the Parole Board. To achieve more timely and swifter justice, short-format or ‘on-the-day’ probation advice to courts has become more common. In a review of PSRs, Her Majesty’s Inspectorate of Probation (HMIP) reported that 58% PSRs produced for courts from June 2018-19 were delivered orally, requested and delivered on the day of sentencing (HMIP, 2020).

The Lammy Review (2017) of race and the criminal justice system, HMIP, practitioners and academics have all raised concerns about the impact of speedier reporting practices on the quality of PSR reports (HMIP, 2020; Robinson, 2019). Specifically, HMIP concluded that oral reports are less analytical and less tailored to the needs of individuals being sentenced than reports produced over a longer time. In addition, a recent thematic report by HMIP on race equality in probation services raised concerns about the quality of pre-sentence reports on people from minority ethnic backgrounds and warned that failure to consider all factors relevant to individuals being assessed could lead to more punitive sentences (HMIP, 2021b). Taken together, these things suggests that speedier reporting practices and possibly poorer quality reports for people from ethnic minorities have the potential to lead to poorer and disproportionately punitive sentencing recommendations.

The Smarter Sentencing White Paper set out the Government’s plans to “*uphold a fairer justice system that works for everyone*”, including a pilot to improve PSRs (Ministry of Justice, 2020, p. 6). The current research report describes the outcome of one part of that wider PSR pilot: a discrete trial commissioned by His Majesty’s Prison and Probation Service (HMPPS) to test an intervention designed to protect effective decision-making in time-pressured conditions. The study aimed to

determine whether the introduction of a brief intervention into the process of delivering on-the-day PSRs could influence the orientation (rehabilitative/punitive) and weight (more or less burdensome) of sentencing recommendations for people from minority ethnic backgrounds.

The trial was conducted during the COVID-19 pandemic, which necessitated some changes to the way pre-sentence reports were conducted and delivered to courts. Most report writers were using a mix of face-to-face and remote contact with the people they were assessing, most commonly conducting interviews for reports over the phone, rather than in person. This change in practice did not have a demonstrable impact on the quality of the reports produced (HMIP, 2021a).

2.2 Review of the literature

Research into those factors that affect the way we attend to, perceive and process information suggests that external scrutiny bodies, practitioners and academics are right to raise concerns about the potential impact of speedier reporting practices on the advice provided to courts. A wide range of disciplines including, but not limited to, behavioural science, psychology, sociology, anthropology and neuroscience have provided important insights into environmental, social and individual-level influences on human decision-making. In a review paper, Dror (2020) set out eight key sources of bias in expert judgement, including (i) those that affect everyone, as a result of the way our brains have developed to adapt to a complex and changing world, (ii) those specific to our individual experiences and environment, and (iii) those that are specific to the case being analysed. For example, we are influenced by our prior experience, to the extent that knowledge of previous cases can shape our expectations and subsequently our judgement of current cases (de Lange, Heilbrun & Kok, 2018). Negativity bias means that we tend to remember more easily and more strongly negative events and behaviour than positive ones (Baumeister, Finkenauer & Vohs, 2001). Confirmation bias means we favour evidence that supports our existing beliefs or initial judgement and tend to overlook or minimise evidence that doesn't (e.g., Hart et al., 2009). Narrow thinking is a common human tendency to focus our attention on one or two factors when making a decision, rather than

considering a wide range of factors, which can result in quicker, but less well-considered, decisions (e.g., Larrick, 2009).

There is clear evidence that the context within which we make a decision can affect our vulnerability to these, and other, biases. Research suggests that fatigue, distraction and time pressure can all make us less careful in our thinking, and more prone to taking mental short cuts (Soll, Milkman & Payne, 2015). This suggests that the time pressure associated with speedier, on-the-day, pre-sentence reports, is a condition that has the potential to affect the quality of decision-making. As well as experiencing pressure because of the tight timeframes associated with PSRs, feedback from report writers suggest that these constraints impact on the completeness and quality of information available on which to base a sentencing proposal, as well as the depth and nature of engagement with the subjects of reports (Robinson, 2019). A thematic review of the quality of pre-sentence reports suggested that those presenting on-the-day, oral, reports to courts had less time to reflect on the information available to them, than did those writing the fuller, longer-form reports, and that this contributed to more formulaic sentencing proposals that were less well-tailored to the individual subject of the report (HMIP, 2020a).

Research into what works to improve the quality of decision-making, reduce errors and protect against biases and other factors, suggests there are things we can do to mitigate these influences. This body of evidence indicates that prompts that encourage decision-makers to consider other information, or to identify different options, can help to counter a range of biases, including our tendency to engage in narrow thinking (Larrick, 2009). A tendency to recommend the same few sentencing proposals risks missing nuances or individual differences that could determine what is an appropriate sentence, particularly for those who differ in meaningful ways from the majority, or who have additional needs. This could, at least in part, explain the poorer quality of pre-sentence reports for some people from ethnic minorities (HMIP, 2021b).

Different prompts that have been effective in broadening thinking include asking decision makers to i) generate alternatives to their original thinking, ii) think about why their initial judgement might be wrong, and iii) make the decision more than once

and average out their answers (Milkman, Chugh & Bazerman, 2009). While there is some high-quality research in this area, much of it looks only at the short-term impact of strategies to manage bias and relatively little has taken place in ‘real world’ settings (Aczel et al., 2015). This study provides a robust test of a task intended to protect good decision-making in on-the-day pre-sentence reports and represents an attempt at trialling an evidence-informed strategy to mitigate the influence of bias and mental short cuts in a real world, criminal justice setting.

2.3 Research questions and hypotheses

The trial aimed to answer three questions, using a mixed methods approach:

- 1) Can a brief intervention to mitigate the influence of biases in decisions about sentencing proposals for people from ethnic minorities be implemented by report writers, as designed, in the PSR process for on-the-day reports?
- 2) Does ethnicity influence the nature of sentencing proposals put forward to courts in on-the-day pre-sentence reports, when controlling for other influential factors?

Hypothesis a) Report subjects’ ethnicity will significantly predict the orientation and/or weight of sentencing proposals put forward to the courts in on-the-day pre-sentence reports, with those from ethnic minorities more likely to receive punitive and heavier sentencing proposals.

- 3) Can a brief intervention to protect against bias in decision-making affect the sentencing recommendations put forward for people from different ethnic groups, when compared with usual PSR practice?

Hypothesis b) Compared to standard practice, the intervention will have a greater impact on the recommendations put forward to courts for minority ethnic report subjects, than for white report subjects.

3. Method

This study received approval from the Ministry of Justice National Research Committee. Ethical considerations are outlined in Appendix A.

3.1 Sample

All probation court report writers taking part in the Ministry of Justice Pre-sentence Report Pilot across 14 courts were asked to take part in the trial ($N = 115$). Five people did not consent to participate, one further person withdrew their consent during the trial, and one left their role near the start of the trial, leaving a final sample of 108 report writers. The report writers produced a total of 985 on-the-day reports during the 10-month study period, from March to December 2021. Each report writer produced between 1 and 47 reports during this time. The average number of reports per report writer was 9.12 ($SD = 9.96$).

The characteristics of the sample are presented in Table 1, although some information was missing for some report writers, who did not want to provide, or otherwise did not respond to requests for, this information.

Table 1. Characteristics of probation report writers in the study

Characteristic	M (SD)
Age ($n = 69$; min = 25, max = 69)	46.39 (11.45)
Years in service ($n = 69$; min = 1, max = 34.75)	14.10 (9.37)
Years as a pre-sentence report writer ($n = 68$; min = 1, max = 29)	6.01 (5.55)
	Percentage (n)
Gender	
Female	72.2% (78)
Male	23.1% (25)
Missing	4.6% (5)
Ethnicity	
Asian	-
Black	6.5% (7)
Mixed ethnicity	-
White	58.3% (63)
Missing	29.6% (32)

Note: Where numbers in any cell were under five the data are redacted to protect participants' anonymity.

Almost three quarters of the report writers were women. Just under 30% of report writers did not provide ethnicity data; of those who did ($n = 76$), 82.9% were white and 17.1% were from an ethnic minority group.

Figure 1 (Appendix B) depicts the allocation of report writers to the experimental or control condition. Fifty-four report writers were randomly assigned to undertake the intervention task; two did not engage in the trial; one withdrew from the study shortly after allocation, and another left the job at the start of the study period. All of those who were allocated to the control condition completed at least one report during the trial period and were therefore included in the final sample.

A subsample of 11 report writers was interviewed to help understand any implementation issues with the task. This was a purposive sample; three report writers were selected who had completed the task as intended throughout the trial, three who had not complied with the task, and four who had sometimes complied and sometimes failed to comply with the task (see section 3.2.3. for information on definition and measure of compliance). Women made up the majority of this subsample (72.7%; $n = 8$). Interviewees had been working for the Probation Service for between two and a half to 30 years ($M = 13.7$ years, $SD = 9.4$). The age of the people interviewed, for whom we had these data ($n = 9$), ranged from 29 to 63 ($M = 49.0$ years, $SD = 11.0$). A third of the people interviewed had been writing pre-sentence reports for three years or less, while another third had been writing reports for court for over 12 years ($M = 8.6$ years, $SD = 9.1$).

3.2 Measures

3.2.1 Rehabilitative orientation and weight of sentencing recommendations

Punishment and rehabilitation are two of the five purposes of sentencing that the court is required to consider when imposing a sentence.² In this study the outcomes of interest were a measure of whether the sanctions proposed in the PSR were more rehabilitative or more punitive, and an appraisal of their weight (based on the duration and demand the sanction(s) places on an individual). In advance of data

² 1) Punishment, 2) reduction of crime including through deterrence, 3) reform and rehabilitation, 4) public protection and 5) reparation to the victim of the crime (Section 57., Sentencing Act 2020).

collection, a coding system for sentencing proposals based on these two dimensions, was devised by the report authors. Table 2 in Appendix C indicates the coding applied to sanctions according to rehabilitative orientation and weight. Three probation officers with several years' experience of writing presentence reports reviewed the framework and agreed that the suggested coding was an adequate reflection of the orientation and weight of the different sanctions described.

For cases in which custody was considered, this was coded as a punitive outcome while a community sentence³ was considered a rehabilitative outcome. Treatment requirements, offending behaviour programmes and rehabilitative activity requirements (RAR) were deemed to be rehabilitative, while fines, unpaid work and curfews were deemed punitive. The weight of the sanctions was judged on the amount and duration of time and effort required of an individual subject to them. The outcomes were initially coded by two research assistants. The report authors double coded all outcomes blind to the experimental condition of the report writer or ethnicity of the report subjects. In 92.1% of the cases both sets of coders agreed on the outcome. Discrepancies were discussed and coding was agreed by both report authors.

3.2.2 Time taken to complete task

The time taken to complete the additional task was noted by the report writers who entered this into the Effective Practice Framework digital tool as soon as they had completed the intervention questions.⁴

3.2.3 Compliance with the task

Adherence to the experimental protocol was determined by assessing the quality of the responses to the questions that formed the intervention task. Responses to these questions were rated independently by two assistants, both for completeness and

³ While community sentences are also punitive, these are considered to be more rehabilitative, and less punitive, than custodial sentences.

⁴ The Effective Practice Framework tool is a digital tool which helps staff to decide on what to propose in a pre-sentence report or licence plan, by providing a shortlist of sentence and intervention options that fit the characteristics and needs of the individual being assessed. It is a rule-based augmented decision-making tool, which uses algorithms to help practitioners make more accurate and consistent decisions.

relevance of evidence provided. Based on these ratings, each task was rated as compliant or non-compliant. Responses were rated as compliant only if two different sentencing proposals were provided, each with a distinct argument for why this was a suitable proposal for the subject of the report. If two different proposals were identified, but no rationale was provided, or if the rationale was not arguing for the proposal (e.g., explained why the proposal was not suitable or relevant in this case), the response was rated as non-compliant. To ensure reliability of the ratings, the compliance ratings were double coded by the report authors. The rate of agreement between the two sets of coders was 84.5%. Where there were discrepancies the report authors' coding was preferred.

3.2.4 Covariates

In addition to these measures, information was collected on report subjects' age and risk of reoffending as assessed by the Offender Group Reconviction Score - 3 (OGRS-3; Howard, Francis, Soothill & Humphreys, 2009), which forms part of the Offender Assessment System (OASys; Home Office, 2006), the assessment of risk and needs related to offending for men and women under the supervision of HMPPS in England and Wales. OGRS-3 is an actuarial assessment of risk of any proven reoffending, that combines information on the age, sex and criminal history of people convicted of crime to estimate the percentage likelihoods of proven reoffending (any conviction or caution for a new offence) committed within one and two years of the start of a community sentence or discharge from custody. In addition, the report subjects' index offence – the main offence for which they were being sentenced on this sentencing occasion – as well as the disposal they received for their last offence, if they had a previous conviction, were obtained from official records. Index offences were initially coded into seven categories: acquisitive, drugs, motoring, robbery, sexual, violent or other crimes, in line with the convention for coding HMPPS offender segmentation data (NOMS, 2013). Analysis indicated that there were too few cases in some of the categories, so robbery, sexual and violent crimes were subsequently collapsed into one category, producing five offence type groups. Previous disposals were classified as (i) no previous disposal, (ii) absolute or conditional discharge, (iii) custody, (iv) community order, (v) suspended sentence order, or (vi) fine, compensation or other. As there were too few cases in some of the disposal type categories for some of the analysis, a collapsed version of this variable was created,

consisting of three categories: no previous disposal, custody or non-custodial disposal (community sentence).

3.3 Materials

3.3.1 Consider the Alternative Task

The experimental task was developed for the purposes of the trial and intended to protect against biases by both (i) slowing down the decision-making process by asking report authors to answer some additional questions before coming to a decision and (ii) broadening their thinking by asking them to identify an alternative (but feasible) sentencing proposal to the one they first identified. Research suggests that asking people to consider the opposite of, or consider an alternative to, the decision they are about to make can effectively counter a number of biases (Milkman, Chugh & Bazerman, 2009). This intervention could be construed as an example of a 'nudge plus', prompting decision makers to reflect on alternatives consciously, rather than intervening solely with automatic, unconscious processes (Banerjee & John, 2021). This approach has the advantage of offering greater autonomy to the decision maker than traditional nudges.

In order to encourage participants to seriously consider the alternative option, report writers were prompted to articulate why each proposal was suitable for the report subject, before committing to one to put forward to the courts (see Appendix D for the questions comprising the experimental task). A small pilot of the task took place with probation report writers at Northampton and Luton Magistrates' courts. Over the course of a week, two report writers at these courts undertook the task on paper when completing on-the-day pre-sentence reports. The research authors reviewed the completed tasks and made changes to the wording designed to improve comprehension of, and fidelity to, the task.

Those report writers in the control condition did not alter their practice in any way. They followed usual practice, filling in report subjects' details in the digital Effective Practice Framework tool before picking proposals from a list of suitable sanctions generated by the tool.

3.4 Procedure

3.4.1 Trial protocol

Participants were assigned to either condition 1 (experimental) or condition 2 (practice as usual). Report writers in condition 1 were required to fill in an additional set of questions (the experimental task; Appendix D) in the Effective Practice Framework.⁵ All reports allocated to condition 2 were completed as normal, with no change to the report writers' usual practice.

The report writers were briefed at the start of the trial by the pilot programme management staff and senior probation officers at each site, using 'how to' videos and written briefing materials. In order to maximise fidelity to the trial protocol, the researchers were available to the report writers throughout the trial to answer any questions and correct any deviations from the protocol (although no participants got in touch). Once the protocol started, the experimental and control conditions remained unchanged.

Information on random allocation, fidelity checks and minimising bias, see Appendix E.

3.4.5. Process evaluation

At the end of the trial period, researchers interviewed, either by telephone or video conference, a subsample of the report writers in the experimental condition. The semi-structured interview schedule focussed on understanding what participants thought of the task, how easy they found it to complete, what interfered with completing it, how it could be improved and what their thoughts were about integrating this or a similar task into the PSR process as future standard practice. Research notes were taken during the interviews for later analysis. A copy of the semi-structured interview is available in Appendix F. On average, the interviews lasted 30 minutes.

⁵ The Effective Practice Framework tool is a digital tool which helps staff to decide on what to propose in a pre-sentence report or licence plan, by providing a shortlist of sentence and intervention options that fit the characteristics and needs of the individual being assessed. It is a rule-based augmented decision-making tool, which uses algorithms to help practitioners make more accurate and consistent decisions.

In addition, report writers were asked to estimate how long the task took them to complete, to provide a rough estimate of the duration of the task. Proportions of people complying with the task, and number of reports produced following proper completion of the intervention was also gathered to provide insight into how the task was implemented, and whether it was implemented as intended.

3.5 Analysis

To reduce the risk of bias resulting from selectively removing people from the experimental group, all sentencing recommendations were analysed regardless of whether report writers in the intervention condition complied fully with the intervention task each time.

Multinomial logistic regression was used to determine whether ethnicity or experimental condition predicted proposal type. These analyses controlled for other factors that could influence the type of sentencing proposals put forward to the courts (see measures for details). To test hypothesis b, the analyses examined whether there was an interaction between the experimental condition and ethnicity of the report subject and the proposal orientation or weight, by including an interaction term in these models.

The first author used thematic analysis to analyse the interviews with report writers, following the steps outlined by Braun and Clarke (2006); (1) becoming familiar with the data; (2) generating coding categories; (3) generating themes; (4) reviewing themes; (5) defining and naming themes; (6) locating exemplars.

3.6 Limitations

For several reasons, it was not possible to assess the mechanism through which any change in outcome might have been achieved. The experimental task aimed to reduce the influence of implicit biases, including confirmation bias and narrow thinking on the decision about whether to recommend a more rehabilitative/punitive or heavy/light sanction at the pre-sentence report stage. It was not possible to measure these factors pre- and post-task, as this would have required resources above and beyond what is available and feasible to implement in a busy court

setting. Doing so would have compromised the delivery of on-the-day reports directed by the court. In addition, existing measures of implicit bias have been criticised, as the scores on implicit association tests have not demonstrated a relationship with discriminatory practices and therefore cannot be interpreted as meaningful indicators of a propensity to discriminate (e.g., Blanton, Jaccard, Strouts, Mitchell & Tetlock, 2015). For example, a meta-analysis of 492 studies examining procedures to change implicit bias found that changes in measures of this type of bias did not necessarily translate into changes in behaviour (Forscher et al., 2019).

Two steps were taken to mitigate the impact of this limitation on determining whether the intervention was responsible for any change in outcome. First, a gold standard method of evaluation was used – a well-implemented randomised control trial. Second, information on some of the other major factors which were expected to have an impact on whether someone is recommended a more punitive or heavier, over a more rehabilitative or lighter, sanction, was obtained. These variables included risk of reoffending, age, current offence, and last disposal type. To determine whether randomisation was successful the report writers in the experimental and control conditions were compared on characteristics (time in service, gender, ethnicity, and length of time as a PSR writer) that might influence the nature of the proposals they put forward to the courts for people from white or minority ethnic backgrounds.

Another limitation is that the report writers were not blind to the condition that they were in; half of the report writers were assigned to do an extra task and knew that the impact of this task was being tested. All report writers participating in the trial were fully informed of the nature of the research, and that it aimed to determine whether a simple addition to the process could affect the nature of PSR recommendations for report subjects from ethnic minority groups. However, on the assumption that PSR writers would not have any vested interest in demonstrating that the additional task is of benefit (indeed some reaction was expected against the inclusion of more stages to the process, which is difficult to achieve in the time available as it is), it was not expected that the explicit demand characteristics of the extra task would have an impact on the report recommendation. Arguably the knowledge that the participants are taking part in a trial looking at proportionality in decision-making may have had the effect of increasing all participants' critical appraisal of their report

recommendations and would therefore have affected reports in both conditions equally, making it more difficult to detect any effect.

In addition, the outcome measure consisted of a judgement on both the orientation and likely perceived weight of the sentencing proposal on those subject to the recommended sanctions. While a reliable coding framework was developed, which led to consistent coding of proposals on these two dimensions by different coders, the only test of validity was a check of the face validity of the coding, by experienced probation officers. There could be significant variation in the way in which different sentences are perceived and experienced by those subject to them. Further tests of the validity of this way of conceiving sentencing options would be helpful in determining the degree to which this is a meaningful outcome.

The lack of a control task means it is possible that those in the intervention condition were under greater time pressure when completing on-the-day reports, than those in the control condition. The former had to complete an additional (albeit relatively brief) task during an already tight time frame for the production of on-the-day reports.

Finally, only 10% of the report writers were interviewed to ascertain issues with implementation of the intervention. While this was a purposive sample, designed to provide a representative set of views of those who took part, additional factors affecting implementation may have been missed.

4. Results

4.1 Efficacy of randomisation

Report writers were randomised into the intervention ($n = 52$) or control groups ($n = 56$) in order that the groups would be comparable on (measurable and non-measurable) variables which could impact the sentencing proposals they decided to put forward to the courts. Participant characteristics are described in Table 3 (Appendix G). Comparison of those assigned to the experimental and control conditions on several demographic variables confirmed that randomisation was successful (see Appendix G for details), i.e., the groups did not differ statistically significantly on variables thought to have a potential influence on PSR proposals made. However, the groups did differ significantly on the number of reports they completed during the trial, a difference that appeared to be down to a small number of highly prolific report writers in the control condition (see Appendix G).

The subjects of the pre-sentence reports were not randomised. However, those characteristics that could influence the proposals put forward by the report writers were compared, to determine whether there were any differences between the subjects of the reports written by report writers in the two conditions (see Table 4, Appendix H). Characteristics of the report subjects were included in subsequent analyses.

4.2 Efficacy of Implementation

In total 58.5% ($n = 220$) of the reports completed by the experimental group were produced following full and proper completion of the task. A fifth (19.2%, $n = 10$) of report writers in the intervention group fully complied with the consider-the-alternative task for every on-the-day report they completed during the trial period. A quarter ($n = 13$) of the report writers in the intervention condition failed to comply with the task at all, while the remaining 56% ($n = 23$) sometimes complied and sometimes failed to comply with the task.

4.3 Attrition

The rates of attrition from the sample were low. Only two people dropped out of the trial altogether, and both were in the experimental group. One withdrew from the trial, while the other left their role. However, 12% of report writers in both the control ($n = 7$) and experimental conditions ($n = 6$) completed only one on-the-day pre-sentence report during the 10-month trial period. It is possible that some of these report writers left the role, but that the researchers were not informed of this.

4.4 Time taken to complete the task

Report writers' records of the time it took them to complete the additional task indicated that this ranged from one to 40 minutes. The average time it took those in the intervention condition to complete the additional questions was 7.44 minutes ($SD = 4.86$).

4.5 Sentencing proposals

Over the course of the trial, the vast majority (96.1%) of people were recommended a community order in their on-the-day pre-sentence report. Only 0.4% ($n = 4$) of the reports recommended custody, while 2.3% recommended a financial sanction only. An absolute or conditional discharge was recommended in 0.6% of reports, and the remaining 0.6% reports recommended other sanctions or were missing this information.

Just over half (52.6%) of the reports proposed predominately rehabilitative sentencing options, while just under a third (32.3%) were classed as being predominately punitive in their orientation.

Two thirds (66.1%) of the sentences proposed were classed as placing a moderate burden on people in terms of the time and effort it would take to comply with the requirements. Just under a quarter (24.3%) of recommendations were classed as being heavily burdensome. Only 6.7% of recommendations were categorised as representing a light burden on the report subject.

Information on the final disposal imposed by the court was missing for 14.6% of the report subjects. However, of those for whom this information was available, just under a quarter (23.2%) were given a different sentence to the one recommended in their pre-sentence report. Most commonly courts decided to impose suspended sentence orders in place of community orders. In 26 cases the court ordered a custodial sentence when the pre-sentence report recommended a community sanction.

4.6 The influence of ethnicity on sentencing proposals

Table 5 shows the orientation and weight of sentencing proposals put forward for people in different ethnic groups.

Table 5. Orientation and weight of sentencing proposal by report subjects' ethnic group.

Ethnicity of report subjects	Asian % (n)	Black % (n)	Mixed ethnic group % (n)	Other % (n)	White % (n)
Orientation of proposals					
Rehabilitative	39.2% (20)	41.5% (44)	61.5% (24)	38.1% (16)	56.0% (413)
Equal	17.6% (9)	23.6% (25)	-	16.7% (7)	13.1% (97)
Punitive	43.1% (22)	34.9% (37)	30.8% (12)	45.2% (19)	30.9% (228)
Weight of proposals					
Light	9.6% (5)	6.6% (7)	-	14.6% (6)	6.5% (47)
Medium	73.1% (38)	74.5% (79)	63.2% (24)	63.4% (26)	67.3% (483)
Heavy	17.3% (9)	18.9% (20)	34.2% (13)	22.0% (9)	26.2% (188)

Note: Where cell numbers were less than five the figures have been redacted.

There was no difference in the weight of the sentencing proposals report writers recommended for those in different ethnic groups ($\chi^2 = 21.39, p = .24, df = 8$). However, there was a statistically significant difference in the orientation of the proposals put forward for report subjects of different ethnicities ($\chi^2 = 11.09, p = .01, df = 8$). Being in the Asian or other ethnic group categories was associated more strongly with receiving a punitive sentencing proposal, and less strongly associated with receiving a predominantly rehabilitative sentencing proposal than being in the

white or mixed ethnic group categories. This analysis does not, however, take into account any difference in other factors that could affect the type of proposal put forward to the courts, including index offence, risk of reoffending, age, last disposal type or gender of those in the different ethnic groups in this sample.

Collapsing the ethnic groups into two categories (see Table 6) produced similar results; there was no difference in the weight of the proposal recommended in the on-the-day pre-sentence reports for people in minority or white ethnic groups ($\chi^2 = 2.37, p = .31, df = 2$). However, being from an ethnic minority was more strongly associated (statistically significantly) with receiving a more punitive or equally punitive and rehabilitative sentencing proposal, while being white was more strongly associated with receipt of a more rehabilitative sentencing proposal in the on-the-day pre-sentence reports produced during this trial ($\chi^2 = 11.31, p = .01, df = 2$). Again, this analysis did not take into account any other factors that could affect the nature of the proposals recommended to the court.

Table 6. Comparison of the orientation and weight of sentencing proposals put forward to the courts in pre-sentence reports for those report subjects of ethnic minority or white background.

	Minority ethnic report subjects (<i>n</i> = 238) % (<i>n</i>)	White report subjects (<i>n</i> = 738) % (<i>n</i>)
Orientation of proposal		
Rehabilitative	43.7% (104)	56.0% (413)
Equal	13.1% (44)	18.5% (97)
Punitive	37.8% (90)	30.9% (228)
Weight of proposal		
Light	8.0% (19)	6.5% (47)
Medium	70.5% (167)	67.3% (483)
Heavy	21.5% (51)	26.2% (188)

To determine whether ethnicity influenced the sentencing recommendations to court, when controlling for other factors that would be expected to influence this outcome (report subject age, risk of reoffending, gender, index offence type and last disposal

type), two multinomial regression analyses were conducted (see Appendix I for Tables 7 and 8 detailing the results).

The model predicting sentencing orientation was statistically significant (-2LL = 1405.29, χ^2 (32) = 195.81, $p = .001$). While gender ($\chi^2 = 6.17$, $p = .05$, $df = 2$), index offence type ($\chi^2 = 34.84$, $p = .001$, $df = 8$) and risk of reoffending ($\chi^2 = 30.75$, $p = .001$, $df = 2$), significantly predicted the orientation of the sentence recommended in on-the-day reports, overall, ethnicity did not ($\chi^2 = 12.57$, $p = .13$, $df = 8$). However, the analysis indicated that being black compared to being white was statistically significantly associated with receiving a sentencing recommendation that included a punitive (as well as rehabilitative) element over a mainly rehabilitative sentencing proposal. Report writers were just over two times (1/0.44) as likely to recommend a predominantly rehabilitative sentence compared to a sentence containing a balance of both punitive and rehabilitative elements, for white report subjects than they were for black report subjects (see Table 7, Appendix I).

The model predicting sentencing weight (which contained the collapsed, three group version of the person's most recent disposal type) was significant (-2LL = 1182.84 χ^2 (26) = 71.24, $p = .001$). Ethnicity was not a significant predictor of proposed sentencing weight ($\chi^2 = 4.42$, $p = .82$, $df = 8$), while age ($\chi^2 = 7.74$, $p = .02$, $df = 2$), and risk of reoffending ($\chi^2 = 11.45$, $p = .003$, $df = 2$) were.

4.7 Impact of the intervention on proposal weight and orientation

To determine whether the intervention had an impact on the orientation and nature of sentencing proposals put forward to the courts, two multinomial logistic regression analyses were performed. Both analyses controlled for a range of possible influences on the sentencing proposal outcomes, comprising the report subjects' age, gender, risk of reoffending score (as measured using OGRS-3), index offence type, last disposal type and ethnicity (binary). The interaction between report subjects' ethnicity and the report writers' intervention condition was also included to test the hypothesis that the intervention would have a differential impact on the outcome of reports for those from ethnic minorities.

The first regression indicated that report subjects' risk of reoffending ($\chi^2 = 32.31, p = .001, df = 2$), offence type ($\chi^2 = 34.58, p = .001, df = 8$) and gender ($\chi^2 = 6.44, p = .04, df = 2$) had an impact on the orientation of their sentencing proposal. There was also an interaction between the ethnicity of the report subject and whether the report writer was in the intervention condition, and the subsequent orientation of the sentencing recommendation put forward to the court ($\chi^2 = 7.24, p = .03, df = 2$).

Table 9 (Appendix J) shows that for every one-point increase in risk of reoffending score, the odds of being recommended a mainly rehabilitative, rather than punitive, sentencing proposal increased by 1.02. This likely reflects the fact that accredited offending behaviour programmes are targeted to those with a higher level of risk of reoffending and greater criminogenic needs. In addition, the odds of a male report subject receiving a predominantly rehabilitative sentencing recommendation compared to a predominantly punitive sentencing proposal, were around half ($1/1.91 = 0.52$) that of a female report subject in this sample. That is, in this trial, on-the-day reports on women were twice as likely to recommend a primarily rehabilitative sentence compared to recommending a primarily punitive sentence than reports on men. Those receiving a sentence for robbery, sexual or violent crime had greater odds of receiving a predominantly rehabilitative rather than punitive sentencing recommendation in their on-the-day pre-sentence report than those who were being sentenced for other types of crime (those being sentenced for 'other' crimes were 0.45 times less likely to receive a mainly rehabilitative than mainly punitive sentencing recommendation).

The odds of receiving an equally rehabilitative and punitive compared to a predominantly punitive sanction were only affected by index offence type and the last disposal received for a previous offence. Those receiving a sentence for other types of crime had smaller odds (0.36 times less) of receiving an equally rehabilitative and punitive than a predominantly punitive sentencing recommendation than those who were being sentenced for robbery, sexual or violent crime. Having received an 'other' disposal type at the last sentencing occasion was associated with smaller odds (0.32 decrease in odds) when compared with receiving a community disposal, of receiving an equally rehabilitative and punitive sentencing recommendation rather than a predominantly punitive recommendation. This means that those who were sentenced

to a community disposal for their last offence were more likely to receive a sentencing recommendation that included some rehabilitative elements for their next offence than were those who received an 'other' type of disposal at their last sentencing occasion.

The second regression model found that only report subject age ($\chi^2 = 7.86, p = .02, df = 2$), and risk of reoffending ($\chi^2 = 12.66, p = .002, df = 2$), significantly predicted the weight of the sentencing outcome put forward to the courts in the on-the-day pre-sentence reports produced during the trial. Because neither ethnicity ($\chi^2 = 0.94, p = .63, df = 2$), nor intervention condition ($\chi^2 = 0.80, p = .67, df = 2$) significantly predicted the weight of the sentencing outcome in this model, an interaction between the two was not analysed. Table 10 (Appendix J) shows that being older slightly increased the odds of getting a lighter, compared with a heavier, and medium compared with a heavier, set of sanctions in the sentencing recommendation. Higher risk of reoffending, as measured using OGRS-3, was associated with slightly lower odds of getting a lighter or moderately weighted set of sanctions, compared with a heavier sentencing proposal. Compared with those being sentenced for other types of crime, those receiving a sentence for a drugs offence had statistically significantly higher odds of receiving a lighter or moderately burdensome sentencing recommendation in their on-the-day pre-sentence report than a heavier set of sanctions.

Repetition of these analyses to examine only those reports which involved faithful implementation of the task, found no significant effect of the intervention condition ($c^2 = 0.13, p = .94, df = 2$; $c^2 = 1.72, p = .43, df = 2$), or report subject ethnicity ($c^2 = 2.03, p = .36, df = 2$; $c^2 = 0.47, p = .79, df = 2$), on proposal orientation or weight, respectively (Tables 11 and 12, Appendix K).

To better understand the significant interaction between report subject ethnicity, intervention condition and the orientation of the sentencing proposal, the proportion and nature of the proposals put forward under the intervention condition and under the control condition were examined. For reports written under the intervention conditions the initial proposals were compared to those alternative proposals that

were put forward to the court (Table 13). The alternative proposal (the second option report writers in the intervention condition were asked to identify) was recommended to the courts in 61 of the on-the-day pre-sentence reports (16.2% of the reports produced under the intervention condition). Whether the proposal put forward to the court was the report writers' initial or alternative proposal made no difference to the rehabilitative orientation ($\chi^2 = 2.02, p = .36, df = 2$), nor weight ($\chi^2 = 3.00, p = .22, df = 2$), of the sanctions proposed (see Table 8).

Table 13. Comparison of the orientation and weight of the initial and alternative sentencing proposals put forward to the courts in pre-sentence reports.

	Initial proposal put forward to court (n = 315) % (n)	Alternative proposal put forward to court (n = 61) % (n)
Orientation of proposal		
Rehabilitative	54.3% (171)	47.5% (29)
Equal	11.7% (37)	18.0% (11)
Punitive	34.0% (107)	34.4% (21)
Weight of proposal		
Light	6.7% (21)	13.1% (8)
Medium	66.0% (208)	62.3% (38)
Heavy	27.3% (86)	24.6% (15)

To examine differences between the weight and orientation of the proposals put forward for those in different ethnic groups in the intervention and control conditions, the ethnic categories were collapsed because the number of people in some of the subgroups were too small for statistical analysis. As seen in Table 14, in the intervention condition, a statistically significantly smaller proportion of report subjects from ethnic minorities were proposed rehabilitative sentences than were white subjects ($\chi^2 = 14.40, p = .00, df = 2$). There was no difference in the weight of the sentencing proposals put forward by report writers in the intervention condition for white or minority ethnic report subjects ($\chi^2 = 1.48, p = .48, df = 2$). There were no significant differences between the proportions of ethnic minority and white report subjects in the control condition, for the orientation ($\chi^2 = 2.58, p = .28, df = 2$) or weight ($\chi^2 = 0.94, p = .62, df = 2$) of the proposals. These analyses do not take into account the impact of other factors on the orientation and weight of sentencing

proposals, such as age, gender, risk of reoffending, index offence type or last disposal type.

Table 14. Number and proportion of report subjects in different ethnic groups by intervention condition with different orientation and weight of sentencing proposals in on-the-day pre-sentence reports.

	Experimental condition		Control condition	
	White report subjects % (n)	Minority ethnic report subjects % (n)	White report subjects % (n)	Minority ethnic report subjects % (n)
Orientation of proposal				
Rehabilitative	57.9% (168)	37.2% (32)	54.7% (245)	47.4% (72)
Equal	10.0% (29)	22.2% (19)	15.2% (68)	16.4% (25)
Punitive	32.1% (93)	40.7% (35)	30.1% (135)	36.2% (55)
Weight of proposal				
Light	7.1% (21)	9.3% (8)	6.1% (26)	7.3% (11)
Medium	64.5% (187)	68.6% (59)	69.2% (296)	71.5% (108)
Heavy	28.3% (82)	22.1% (19)	24.8% (106)	21.2% (32)

One possible explanation is that the intervention prompted report writers to add a punitive element to their sentencing proposal – turning primarily rehabilitative or equally punitive and rehabilitative recommendations into recommendations that contained additional punitive sanctions - but only for those report subjects from ethnic minorities. This would mean that, compared to white report subjects, report subjects from ethnic minorities were less likely to get a predominantly rehabilitative sentencing proposal.

To better understand the way in which the intervention may have impacted on report writer’s recommendations, a series of chi-square analyses was conducted, comparing the orientation of the sentencing proposals put forward for white and minority ethnic report subjects for those reports that complied with the intervention task, failed to comply with the intervention or were in the control group. There were differences in the proportion of report subjects in different ethnic groups who were recommended rehabilitative, equal or punitive sentencing proposals both for those

whose reports were written following proper completion of the intervention ($\chi^2 = 6.56$, $p = .04$, $df = 2$) and those whose report writers did not complete the intervention as it was intended ($\chi^2 = 8.98$ $p = .01$, $df = 2$). Whether or not the report writer complied with the intervention task, white report subjects were more likely than report subjects from ethnic minorities to receive a predominantly rehabilitative sentencing recommendation and less likely to receive a predominantly punitive proposal. This difference was larger when the report writer did not comply with the intervention task (Table 15). When report writers complied with the intervention, there was a 2.6 percentage point difference in the proportion of people from white compared with ethnic minority backgrounds who were proposed a predominately punitive sentence. This difference rose to 17.6 percentage points when report writers did not complete the intervention as intended. However, numbers were small in this sub analysis.

Table 15. Comparison of number and proportion of reports produced under the intervention condition which complied or failed to comply with the experimental task, on people from different ethnic groups, with different orientation and weight of sentencing proposals in on-the-day pre-sentence reports.

	Report writer complied with the intervention		Report writer did not comply with the intervention	
	White report subjects	Minority ethnic report subjects	White report subjects	Minority ethnic report subjects
Orientation of proposal				
Rehabilitative	54.2% (91)	38.5% (20)	63.1% (77)	35.3% (12)
Equal	11.9% (20)	25.0% (13)	7.4% (9)	17.6% (6)
Punitive	33.9% (57)	36.5% (19)	29.5% (36)	47.1% (16)

4.8 Report writers' views of the intervention

While the majority of the report writers interviewed (ten of the eleven) felt that the task was quick and easy to complete, most described it in mainly negative terms ($n = 8$). There were two key reasons for this; report writers felt additional pressure as the task further constrained the time available to complete on-the-day reports, and more experienced report writers questioned the value of the task.

Time constraints

A consistently and frequently identified theme that came through strongly from interviews with report writers, was the lack of time they felt they had to fulfil all the requirements of an on-the-day report. Eight of the eleven people interviewed indicated that there was too little time to gather all the relevant information and compile a report, before the addition of an extra task, or felt that the time constraints on on-the-day reports created unhelpful pressure, which the task exacerbated.

Value

Eight of the eleven report writers interviewed felt that the task was unnecessary because they believed their experience was sufficient to protect against bias and to ensure they make fair and effective proposals. Six indicated that while they believed they would not benefit from, or did not need to undertake, the task, they felt others would. Participants indicated that the task could help report writers coming in to write reports from another probation area (usually due to short staffing), or staff new to pre-sentence report writing. One interpretation of this finding is that participants were demonstrating the '*bias blindspot*' (Pronin, Lin & Ross, 2002); a term which refers to our tendency to find it easier to accept that biases affect other people, and harder to see these as relevant to ourselves. The three report writers who felt the task was personally valuable indicated that it prompted them to reflect on their proposals, slowed down their thinking and was easy to use. All three of these report writers were relatively new to court report writing.

5. Discussion and implications

5.1 Summary of findings

In this study, probation court report writers were asked to consider an alternative sentencing proposal to the one they initially decided upon. It was hoped that this would prompt deliberate and conscious reflection on their choices, which would counter implicit biases and narrow thinking that could lead to more punitive and harsher sanctions for people in ethnic minority groups. While there is some debate about the extent of discrimination in sentencing for ethnic minorities, the best available evidence suggests that this does exist in England and Wales (Sorsby, 2002; Lammy, 2017; Von Hirsch & Roberts, 1997). Despite its limitations, this study provides some support for this notion. In this sample, while controlling for a number of potentially influencing variables, being white was associated with a greater chance of receiving a predominantly rehabilitative sentencing recommendation in an on-the-day pre-sentence report, than coming from an ethnic minority group.

The trial suggests that the intervention was relatively brief, taking an average of just over seven minutes to complete. However, compliance with the task was low; fewer than a fifth of the report writers in the intervention condition completed the 'consider an alternative proposal' task as intended throughout the trial period. This suggests that the task cannot be implemented reliably as intended in on-the-day presentence report writing practice without further intervention, such as increasing the time available to write the reports or marketing of the potential benefits of such an intervention, to report writers.

The results suggest that the 'consider an alternative' task influenced the nature of the sentencing proposals that probation report writers put forward to the courts in on-the-day sentencing reports for people from ethnic minorities, but not in the way expected. Being in the intervention condition appears to have had the effect of prompting report writers to identify additional punitive elements to a sentencing proposal, moving from predominantly rehabilitative or equally rehabilitative and punitive sentencing options to proposals that contain mostly punitive elements. However, this effect is confined to reports on people from ethnic minorities; the same effect is not observed for those

report subjects of white background. This effect was not observed when examining only the outcomes of those reports that complied with the intervention task. There are a number of possible explanations for this finding. It is possible that those failing to comply with the intervention were more biased or less likely to notice and act on their biases, than those who completed the task. Alternatively, for those report writers in the intervention condition who did not comply with the task, the trial may have had the effect of raising awareness of ethnicity as a factor that might influence their decision making, making this more prominent and more likely to subsequently influence the proposal they put forward to the courts, without the mitigating effect of the intervention. Finally, the trial may have triggered *psychological reactance* among some of those in the intervention condition, which could explain in part the high rates of non-compliance with the task. Psychological reactance occurs when people feel their autonomy is under threat and can result in hostility or defiance (Brehm & Brehm, 1981). Some studies have found that psychological reactance to initiatives aimed at reducing discriminatory behaviour can increase levels of implicit and explicit prejudice (Legault, Gutsell & Inzlicht, 2011; Lindner, Graser & Nosek, 2014). It is possible that the introduction of the intervention had a backfire effect, increasing the salience and influence of ethnicity in reports writers' decisions about proposals to recommend to the courts, and reducing the likelihood of proper completion of the task.

To be most effective, prompts for conscious reflection and deliberation on choices should encourage ownership of the reflective process, and commitment to and investment in the outcome that the prompts aim to bring about (in this case, fairness and proportionality, through mitigation of the influence of bias on sentencing proposals put forward to the courts) (Banerjee & John, 2021). Interviews with a subsample of the report writers who were asked to complete the intervention indicated that most were not convinced of the merits of the approach, partly because it added time to a task that was already difficult to achieve in the allotted timeframe, and partly because they did not feel it was necessary, as they did not believe their proposal recommendations were or could be influenced by biases. This supports the notion that the task may have induced a backfire effect resulting from report writers' perceptions that the task was unnecessary, and a potential threat to their

professional autonomy and ability to do the job well in the (limited) time available to complete on-the-day reports.

The fact that some of the report writers tasked with completing the intervention felt that it was difficult to incorporate a seven-minute task into their on-the-day report writing practice underscores the severe time constraints under which these reports are produced. Interviewed participants indicated that time pressures were felt keenly, and some participants raised concerns about the impact of these time constraints on the quality of their reports. In line with these concerns, this study provides some evidence of disproportionality based on ethnicity, in sentencing recommendations put forward to the courts in on-the-day pre-sentence reports.

5.2 Conclusion

A brief intervention that aimed to protect against biases in decisions about the type and nature of sentence to recommend to the courts in time pressured conditions, was in practice relatively brief and easy to follow. However, almost half of the reports produced under the intervention condition did not follow proper and full completion of the task. Most of the probation report writers interviewed felt that the intervention was both unnecessary and difficult to fit into the constricted time they had to complete on-the-day reports. The fact that report writers found it challenging to fit a short task into their practice emphasises the time constraints under which these reports are written. The intervention itself had a mixed effect on the sentencing proposals put forward to the courts. For those report subjects from ethnic minorities, the intervention condition had the effect of prompting report writers to opt for a sentence that encompassed more punitive elements, over predominantly rehabilitative sanctions. This effect was stronger on the outcome of sentencing proposals for those report writers who were asked to complete the intervention but did not comply with this request.

The trial suggests that the intervention was not implemented as intended in a real world setting and failed to have the desired effect on the sentences proposed to the courts in on-the-day pre-sentence reports. The findings also underscore the importance of the way in which anti-discrimination initiatives are introduced, to avoid unintended consequences.

5.3 Future research

Future research could help improve our understanding of how and under what conditions interventions to protect good decision-making can work in a real-world criminal justice setting. Specifically, it would be helpful to better understand how the way in which interventions are introduced affects their success. Are there specific conditions that make successful implementation more likely, and how important is the buy in of those staff expected to complete or facilitate the interventions?

Implementation science points to several necessary preconditions for the successful roll-out of new practices or interventions (e.g., Damschroder et al., 2009). It would be helpful to better understand how these conditions can be met, and the way in which they work, in a criminal justice context. Given report writers' concerns about the time constraints of on-the-day pre-sentence reports, further research could also examine the impact of time available to write reports on the quality of those reports, and on the nature and orientation of the sentencing proposals put forward to courts by probation practitioners.

References

- Aczel, B., Bago, B., Szollosi, A., Foldes, A., & Lukacs, B. (2015). Is it time for studying real-life debiasing? Evaluation of the effectiveness of an analogue intervention technique. *Frontiers in Psychology, 6*, 1120.
- Banerjee, S., & John, P. (2021). Nudge plus: Incorporating reflection into behavioral public policy. *Behavioural Public Policy, 1–16*.
- Baumeister, R. F., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323-370.
- Blanton, H., Jaccard, J., Strouts, E., Mitchell, G. & Tetlock, P. (2015). Towards a meaningful metric of implicit prejudice. *Journal of Applied Psychology, 100*, 1468 – 1481.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77-101.
- Brehm, S. S., & Brehm, J. W. (1981). *Psychological Reactance: A Theory of Freedom and Control*. New York, NY: Academic Press.
- Damschroder, L.J., Aron, D.C., Keith, R.E., Kirsh, S.R., Alexander, J.A., Lowery, J.C (2009) Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Science, 4:50*.
- de Lange, F. P., Heilbron, M., & Kok, P. (2018). How do expectations shape perception? *Trends in Cognitive Science, 22*, 764–779.
- Dror. I. (2020). Cognitive and human factors in expert decision making: Six fallacies and the eight sources of bias. *Analytical Chemistry, 92*, 7998–8004.

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., & Devine, P. G. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology: Attitudes and Social Cognition*, 117, 522-559.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, 135, 555-588.

Her Majesty's Inspectorate of Probation (2020). *The Quality of Pre-sentence Information and Advice Provided to Courts: Research and Analysis Bulletin 2020/04*. HMIP, Manchester; UK.

Her Majesty's Inspectorate of Probation (2021a). *A Thematic Review of the Quality and Effectiveness of Probation Services Recovering from the Impact of Exceptional Delivery Models Introduced Owing to the Covid-19 Pandemic*. HMIP, Manchester; UK.

Her Majesty's Inspectorate of Probation (2021b). *Race Equality in Probation: The Experiences of Black, Asian and Minority Ethnic Probation Service Users and Staff*. HMIP, Manchester; UK.

Lammy, D. (2017). *The Lammy Review: An Independent Review into the Treatment of, and Outcomes for, Black, Asian and Minority Ethnic Individuals in the Criminal Justice System*. Ministry of Justice, London; UK.

Larrick, R. P. (2009). Broaden the decision frame to make effective decisions. In E. A. Locke (Ed.), *Handbook of Principles of Organizational Behavior* (2nd Ed.). Chichester, UK: John Wiley & Sons.

Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, 22, 1472-1477.

Lindner, N. M., Graser, A., & Nosek, B. A., (2014). Age-based hiring discrimination as a function of equity norms and self-perceived objectivity. *PLoS ONE* 9: e84752. doi:10.1371/journal.pone.0084752

Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How can decision-making be improved? *Perspectives on Psychological Science*, 4, 379-383.

Ministry of Justice (2020). *A Smarter Approach to Sentencing*. Ministry of Justice, London; UK.

National Offender Management Service (2013). *NOMS Evidence and Segmentation: Companion Document to NOMS Commissioning Intentions from 2014*. NOMS Evidence and Segmentation (publishing.service.gov.uk)

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28, 369-381.

Robinson, G. (2019). Delivering McJustice? The probation factory at the Magistrates' court. *Criminology & Criminal Justice*, 19, 605–621.

Soll, J., Milkman, K. L., & Payne, J. W. (2015). The users' guide to debiasing, In G. Keren & G. Wu, (Eds.) *The Wiley-Blackwell Handbook of Judgment and Decision Making*. Chichester, UK; John Wiley & Sons.

Sorsby, A. (2022). *Data Insight: The Number of Convictions before the Imposition of a Short Sentence of Immediate Custody: Ethnicity and Gender Differences*. Available at:

Data_Insight_The_number_of_convictions_before_the_imposition_of_a_short_sentence_of_immediate_custody_ethnicity_and_gender_differences_August_2022.pdf (adruk.org)

Von Hirsch, A., & Roberts, J. V. (1997). Racial disparity in sentencing: Reflections on the Hood study. *The Howard Journal*, 36, 227-236.

Appendix A

Ethical Considerations

The research was approved by the Ministry of Justice National Research Committee. The following ethical issues were considered:

Risk of harm

There was no obvious reason to suppose that the inclusion of the debiasing task or participation in the trial generally could have any harmful effect on the report writers who agreed to take part. The task being trialled had demonstrated good outcomes in other settings, and there was no good reason to believe that this could negatively affect report recommendations, and therefore the people who are subject to those reports. However, the purpose of the trial was to determine what impact (positive or negative), if any, the task has on outcomes, so until the trial is complete, we could not know for certain.

Consent

All pre-sentence report writers who are active at the court during the course of the trial were given opportunities to opt out of the trial. The nature and purpose of the trial was explained verbally, during video briefing sessions, and potential participants were issued with written information about the trial. Participants were informed of the opt out mechanism; the voluntariness of participation was emphasised in communications, as was the fact that there would be no negative consequence should anyone wish not to take part.

As the subjects of the reports were not under study, we did not seek their consent. We did require some information about the subjects of the reports, in order to rule out that factors other than the experimental task were responsible for any change in proportion of punitive/heavy recommendations, but this is routine management information that was collected and held centrally by the NPS Quality and Effective Practice Team.

Confidentiality

All the information obtained in the semi-structured interviews and reports was kept confidential.

Anonymity

All the information obtained was made anonymous during the data collection and in the write-up. Participants were referred to as a number for the interview notes. If individual-level information was used during the write-up of the process evaluation aspect of the research, no information was included that could be directly attributable to any one participant.

Participants' right to withdraw

At any point prior, during or immediately after completion of the interviews the participants had the right to withdraw from the study. However, once the data had been analysed and incorporated into the results the researchers were unable to withdraw individual data, however, all the data remained anonymous.

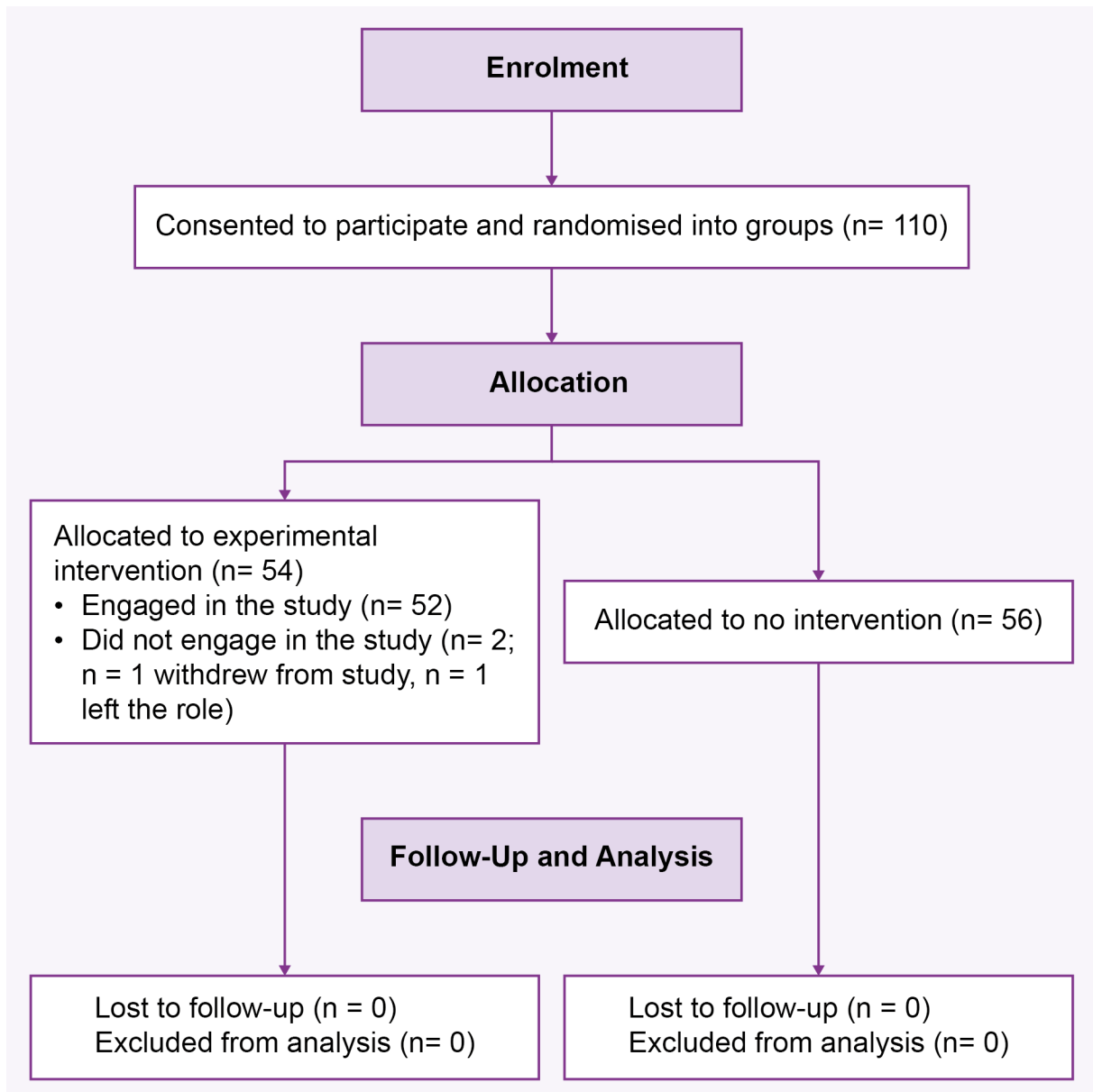
Data protection (including data security, retention and disposal).

An information sharing agreement adhering to the Ministry of Justice data protection policy and GDPR was set up and agreed prior to the sharing of information about on-the-day pre-sentence reports produced at trial sites by participants during the trial period, between the researchers and the MoJ National Reporting Team.

Appendix B

Flow diagram of sample attrition

Figure 1. CONSORT flow diagram of sample attrition at each stage of the trial



Appendix C

Sentencing Proposal Coding Framework

Table 2: Sanctions grouped by rehabilitative orientation and weight

Description	Orientation	Weight
Conditional discharge	Rehabilitative	Light
No conditions	Rehabilitative	Light
Suspended Sentence Order	Punitive	Medium
Exclusion or Prohibited Activity alone	Punitive	Light
Fine alone	Punitive	Light
Fine plus RAR (1-10 days)	Rehabilitative	Light
Fine plus RAR (11+ days)	Rehabilitative	Medium
Curfew alone or fine and curfew	Punitive	Medium
Curfew plus RAR (up to and including ten days)	Punitive	Medium
Curfew plus RAR (11-20)	Equal	Medium
Curfew plus RAR (21+)	Rehabilitative	Heavy
Curfew plus programme	Equal	Medium
Curfew plus RAR plus programme or MHTR, ATR, DRR	Rehabilitative	Heavy
Curfew and unpaid work plus RAR (any length)	Punitive	Heavy
Curfew and unpaid work	Punitive	Medium
Unpaid work* alone	Punitive	Medium
Unpaid work and fine or exclusion	Punitive	Medium
Unpaid work plus RAR (up to and including ten days)	Punitive	Medium
Unpaid work plus RAR (up to and including 10 days) plus fine	Punitive	Medium
Unpaid work and programme	Equal	Medium
Unpaid work plus RAR (11-20 days)	Equal	Medium
Unpaid work plus RAR (21+ days)	Rehabilitative	Heavy
Unpaid work plus RAR plus programme or MHTR, ATR, DRR or CRS	Rehabilitative	Heavy
Accredited programme	Rehabilitative	Medium
CTR (one or two)	Rehabilitative	Medium
RAR (1-10 days)	Rehabilitative	Light
RAR (11+)	Rehabilitative	Medium

Description	Orientation	Weight
RAR plus prohibited activity	Rehabilitative	Medium
RAR plus one additional rehab requirement (CRS, programme, or CTR (MHTR, ATR, DRR))	Rehabilitative	Medium
RAR plus one additional rehab requirement plus one punitive element	Rehabilitative	Heavy
RAR plus two additional rehab requirements (programme plus CRS or CTR (MHTR or ATR or DRR), or two programmes or two CTRs/CRSs)	Rehabilitative	Heavy
Attendance Centre Requirement	Rehabilitative	Light
Alcohol Abstinence and Monitoring Requirement	Equal	Medium
Alcohol Abstinence and Monitoring Requirement plus RAR	Rehabilitative	Medium

Note: RAR = Rehabilitative Activity Requirement; MHTR = Mental Health treatment Requirement; ATR = Alcohol Treatment Requirement; DTR = Drug Rehabilitation Requirement; CTR = Community Treatment Requirement; CRS = Commissioned Rehabilitative Service.

* Information on the number of hours of unpaid work proposed was not available

The number of hours of unpaid work someone was proposed was not taken into account in the coding framework because this information was not available. Unpaid work plus a Rehabilitative Activity Requirement (RAR) of 21 days or more was judged to be primarily rehabilitative, as in most cases the number of RAR hours would likely exceed the number of unpaid work hours. In addition, we considered that unpaid work was not wholly punitive, but had a rehabilitative element as it was bringing people together for purposeful activity and enabling them to “give back” to their community. Similarly, where RARs were proposed alongside a fine, the RAR activity was deemed to exceed activity/effect relating to the fine and this was therefore classed as a predominantly rehabilitative sentencing option.

Appendix D

Questions in the experimental task

The following questions appeared in the Effective Practice Framework (EPF) for those assigned to the experimental and control conditions of the trial.

Standard Pilot Questions

Pop up Q: “Is this an on the day report?” yes/no

Pop up Q: “Race/ethnicity” for both pilot and control group which only pops up if they say yes (above) and has drop down options of White, Black, Asian, Mixed and Other

Select from the following shortlist of sentences and professional override sentences

Select from the following shortlist of interventions and professional override interventions

Pop Up Q: “Number of RAR days” if they choose RAR as an intervention (fixed number)

Pop up Q: “Hours of Unpaid Work” if they choose unpaid work as an intervention (fixed number)

Pop up Q: “Type of curfew” if they choose curfew as an intervention (free text)

Experimental Group Pilot Questions

The following changes to the EPF are part of the PSR trial, and test a way of making it less likely that we’ll use mental shortcuts in our decisions in on-the-day reports

Pop up Q: “Is this an on the day report?” yes/no

Pop up Q: “Race/ethnicity” for both pilot and control group which only pops up if they say yes (above) and has drop down options of White, Black, Asian, Mixed and Other

Option A: Select from the following shortlist of sentences and professional override sentences

Option A: Select from the following shortlist of interventions and professional override interventions

Pop Up Q: "Number of RAR days" if they choose RAR as an intervention (fixed number)

Pop up Q: "Hours of Unpaid Work" if they choose unpaid work as an intervention (fixed number)

Pop up Q: "Type of curfew" if they choose curfew as an intervention (free text)

Question: "Why is this proposal the right proposal for this person?" (free text)

These next questions help you to identify a good alternative to your initial proposal

Option B: Select from the following shortlist of sentences and Professional override sentences

Option B: Select from the following shortlist of Interventions and Professional override interventions

Pop up Q: Number of RAR days (if they choose RAR as an intervention; fixed number)

Pop up Q: Hours of Unpaid Work (if they choose unpaid work as an intervention; fixed number)

Pop up Q: Length of curfew (if they choose curfew as an intervention; free text)

Why is this proposal the right proposal for this person? (free text)

Which option will you select: A or B? (free text)

Appendix E

Randomisation and trial protocol

Random allocation

A block randomisation technique was used, which is appropriate for use with small to moderate samples (for example, less than 50 per group), in order to balance the groups so that there were a roughly equal number of participants in each. The researchers predetermined a set number of people (a 'block', e.g., 6, 8, 10) to be assigned to a condition (e.g., experimental) using a random number generator in MS Excel. Those in the next block were then all assigned to the other condition (e.g., control). The block sizes were varied randomly using the random number generator.

Fidelity checks

Compliance with the experimental task was checked regularly for the duration of the trial (see measures section). No action was taken to correct deviations in completion of the intervention, in most part due to the constraints of the pandemic on access to report writers.

Minimising bias

The ideal RCT is 'double blind' which means that neither the participants nor the people implementing the conditions know who is in the experimental or the control conditions. This was not possible; those in the experimental condition knew they were being asked to complete an intervention, and those in the control group were aware they were continuing as normal. The researchers coding the main outcome measure (sentencing recommendation) were, however, blind to the condition of the participants. In addition, the researchers were blind to the group participants were in when conducting the analysis, to reduce the influence of bias at this stage on the results.

Appendix F

Semi-structured interview schedule

Each broad question to be followed by further exploration and prompts

What was your experience of completing the task? Prompts:

- Positive
- Negative
- Differences noticed in own decision-making practice

What challenges or barriers did you encounter when doing the task? Prompts:

- Personal/individual
- Situational/contextual
- Other

How faithfully did you implement the task? Prompts:

- Done every time, ever missed, etc

How long did the task take you? Prompts:

- Variations in time taken, and for what reasons

How feasible would this be to make standard practice in the PSR process? Prompts:

- What would need to be in place to make this feasible

What improvements could be made to the task if this was used again in future?

Prompts:

- Explanation given
- Layout/formatting
- Other

Appendix G

Characteristics of the report writers in the experimental and control groups

Table 3. Characteristics of report writers in the experimental and control conditions

	Experimental <i>M (SD)</i>	Control <i>M (SD)</i>
Age (<i>n</i> = 69)	47.86 (10.95)	44.79 (11.93)
Years in service (<i>n</i> = 69)	15.72 (8.88)	12.23 (9.71)
Years as a pre-sentence report writer (<i>n</i> = 68)	7.13 (6.16)	4.82 (4.63)
	Percentage (<i>n</i>)	Percentage (<i>n</i>)
Gender		
Female	76.9% (40)	67.9% (38)
Male	21.2% (11)	25.0% (14)
Missing	1.9% (1)	7.1% (4)
Ethnicity		
Asian	-	-
Black	-	-
Mixed Ethnic Group	-	-
White	65.4% (34)	51.8% (29)
Missing	25.0% (13)	33.9% (19)

Where numbers in any cell were under five the data is redacted to protect participants' personal information

Significance testing indicated that there were no significant differences between the report writers in the experimental and control groups in age ($t = -1.12, p = .27, df = 67$), time working for the Probation Service ($U = 731.5, p = .09$), or time as a PSR writer ($U = 725.5, p = .07$). There were also no statistically significant differences between the two groups in their gender ($\chi^2 (2,1) = 2.06, p = .37$), or ethnicity ($\chi^2 (4,1) = 5.52, p = .24$),

However, the report authors in the experimental group produced far fewer on-the-day reports throughout the study ($n = 376$), than those in the control group ($n = 613$). On

average, those in the experimental group produced 7.23 ($SD = 7.67$) reports over the trial period, compared with 10.88 ($SD = 11.48$) for those in the control group (t (*equal variances not assumed*) = 1.93, $p = .05$, $df = 96.5$, Cohen's $D = 0.37$). This discrepancy largely appears to be down to a small number of highly prolific on-the-day report writers in the control condition.⁶

⁶ Six of the report writers in the control condition wrote over 30 on-the-day reports during the trial period, compared with only one of those in the intervention condition.

Appendix H

Characteristics of report subjects for reports written under experimental and control conditions

Table 4. Characteristics of the report subjects in the experimental and control conditions

	Experimental (n=376) M (SD)	Control (n=609) M (SD)
Age (<i>n</i> = 911)	36.27 (10.76)	36.48 (11.25)
OGRS 3 Score (<i>n</i> = 985)	40.41 (24.41)	40.11 (25.84)
	<i>n</i> (%)	<i>n</i> (%)
Gender		
Female	13.0% (49)	15.8% (96)
Male	87.0% (327)	84.2% (513)
Ethnicity		
Asian	5.3% (20)	5.3% (32)
Black	10.1% (38)	11.3% (69)
Mixed Ethnic Group	4.3% (16)	3.8% (23)
Other	3.2% (12)	5.1% (31)
White	77.1% (290)	74.4% (453)
Not recorded	-	-
Index Offence		
Acquisitive	6.9% (26)	10.0% (61)
Drugs	5.3% (20)	3.6% (22)
Motoring	27.9% (105)	25.5% (155)
Sexual	1.3% (5)	0.8% (5)
Violent	34.3% (129)	36.1% (220)
Breach	3.2% (12)	2.6% (16)
Other	6.4% (24)	6.7% (41)
Not recorded	14.6% (55)	14.6% (89)

	Experimental (n=376) M (SD)	Control (n=609) M (SD)
Last disposal type		
Absolute/conditional discharge	-	4.3% (26)
Community order	21.5% (81)	19.0% (116)
Custody	14.6% (55)	13.1% (80)
Fine	10.6% (40)	6.9% (42)
First offence	43.9% (165)	49.9% (304)
Suspended Sentence Order	8.2% (31)	6.6% (40)
Other	-	-

Note: OGRS 3 = Offender Group Reconviction Scale – 3. Where cell numbers were less than five the figures have been redacted.

Significance tests indicated no difference between the subjects of the reports in the experimental and control groups in age ($t(909) = 0.28$, $p = .78$), OGRS-3 score ($U = 116759.5$, $p = .52$), gender ($\chi^2 = 1.38$, $p = .24$, $df = 1$), ethnicity ($\chi^2 = 3.23$, $p = .67$, $df = 5$), nor index offence type ($\chi^2 = 5.77$, $p = .57$, $df = 7$). However, the two groups did differ in the type of disposal they last received ($\chi^2 = 16.20$, $p = .01$, $df = 6$), with a smaller proportion of report subjects in the control than in the intervention condition having received a fine at their most recent sentencing for a prior offence.

Appendix I

Results of multinomial regression analyses examining the influence of ethnicity on orientation and weight of sentencing proposals

Table 7. Summary of multinomial logistic regression analysis examining impact of ethnicity on sentencing proposal orientation

	<i>b</i> (<i>SE</i>)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Rehabilitative vs. Punitive sanctions				
Intercept	1.20 (0.86)	-	-	-
Report subjects' age	0.02 (0.01)	1.00	1.02	1.04
Report subjects' risk of reoffending	0.02 (0.01)**	1.00	1.02	1.03
Report subjects' gender (female)	0.21 (0.31)	0.67	1.23	2.25
Report subjects' index offence (acquisitive)	-0.21 (0.57)	0.27	0.82	2.47
Report subjects' index offence (drugs)	-1.06 (0.58)	0.11	0.35	1.08
Report subjects' index offence (motoring)	-0.74 (0.44)	0.20	0.48	1.13
Report subjects' index offence (robbery, sex or violence)	-0.15 (0.43)	0.37	0.86	1.98
Report subjects' last disposal (no previous offence)	-0.87 (0.55)	0.14	0.42	1.23
Report subjects' last disposal (absolute or conditional discharge)	-0.85 (0.76)	0.10	0.43	1.91
Report subjects' last disposal (custody)	-0.52 (0.57)	0.20	0.60	1.83

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Report subjects' last disposal (community order)	-0.88 (0.53)	0.15	0.41	1.17
Report subjects' last disposal (suspended sentence order)	-0.41 (0.63)	0.20	0.67	2.28
Report subjects' ethnicity (Asian)	-0.23 (0.50)	0.30	0.79	2.12
Report subjects' ethnicity (Black)	0.81 (0.32)**	0.24	0.44	0.83
Report subjects' ethnicity (Mixed Ethnic Group)	0.86 (0.77)	0.53	2.36	10.59
Report subjects' ethnicity (Other)	-0.35 (0.62)	0.21	0.71	2.40
Equal vs. Punitive sanctions				
Intercept	2.12 (0.90)*	-	-	-
Report subjects' age	0.01 (0.01)	0.99	1.01	1.03
Report subjects' risk of reoffending	-0.01 (0.01)	0.98	0.99	1.00
Report subjects' gender (female)	-0.44 (0.34)	0.33	0.65	1.26
Report subjects' index offence (acquisitive)	-0.30 (0.59)	0.23	0.74	2.34
Report subjects' index offence (drugs)	-1.03 (0.59)	0.11	0.36	1.12
Report subjects' index offence (motoring)	-0.43 (0.44)	0.27	0.63	1.50
Report subjects' index offence (robbery, sex or violence)	-0.95 (0.44)	0.16	0.39	0.92
Report subjects' last disposal (no previous disposal)	-0.35 (0.59)	0.23	0.71	2.23
Report subjects' last disposal (absolute or conditional discharge)	-1.32 (0.94)	0.04	0.27	1.69
Report subjects' last disposal (custody)	-0.40 (0.63)	0.20	0.67	2.28
Report subjects' last disposal (community order)	-1.17 (0.58)	0.10	0.31	0.98
Report subjects' last disposal (suspended sentence order)	-0.68 (0.69)	0.13	0.51	1.96

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Report subjects' ethnicity (Asian)	0.23 (0.50)	0.48	1.26	3.36
Report subjects' ethnicity (Black)	-0.49 (0.33)	0.32	0.62	1.18
Report subjects' ethnicity (Mixed Ethnic Group)	0.87 (0.80)	0.50	2.39	11.38
Report subjects' ethnicity (Other)	0.43 (0.58)	0.50	1.54	4.78

Note: $R^2 = 0.21$ (Cox-Snell), 0.25(Nagelkerke) Model χ^2 (32) = 195.81, $p = .001$

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 8. Summary of multinomial logistic regression analysis examining impact of ethnicity on sentencing proposal weight

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Light vs. Heavy sanctions				
Intercept	-2.43 (1.11)*			
Report subjects' age	.04 (.02)**	1.01	1.04	1.07
Report subjects' risk of reoffending	-0.03 (0.01)**	0.96	0.98	0.99
Report subjects' gender (female)	0.66 (0.42)	0.86	1.93	4.36
Report subjects' index offence (acquisitive)	0.65 (0.68)	0.51	1.92	7.32
Report subjects' index offence (drugs)	1.79 (0.77)*	1.32	5.99	27.13
Report subjects' index offence (motoring)	-0.16 (0.54)	0.29	0.85	2.46

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Report subjects' index offence (robbery, sex or violence)	0.14 (0.51)	0.42	1.15	3.14
Report subjects' last disposal (no previous offence)	0.71 (0.65)	0.57	2.03	7.25
Report subjects' last disposal (non-custodial)	0.26 (0.56)	0.43	1.30	3.89
Report subjects' ethnicity (Asian)	0.83 (0.62)	0.67	2.29	7.78
Report subjects' ethnicity (Black)	0.23 (0.53)	0.44	1.26	3.57
Report subjects' ethnicity (Mixed Ethnic Group)	-0.71 (1.08)	0.06	0.49	4.10
Report subjects' ethnicity (Other)	0.67 (0.70)	0.50	1.95	7.63
Medium vs. Heavy sanctions				
Intercept	0.43 (0.58)			
Report subjects' age	0.02 (0.01)*	1.00	1.02	1.04
Report subjects' risk of reoffending	-0.01 (0.01)**	0.98	0.99	1.00
Report subjects' gender (female)	0.09 (0.26)	0.65	1.10	1.84
Report subjects' index offence (acquisitive)	0.64 (0.38)	0.89	1.90	4.03
Report subjects' index offence (drugs)	1.03 (0.55)	0.95	2.80	8.29
Report subjects' index offence (motoring)	0.26 (0.30)	0.72	1.30	2.35
Report subjects' index offence (robbery, sex or violence)	0.18 (0.29)	0.69	1.20	2.10
Report subjects' last disposal (no previous disposal)	0.53 (0.32)	0.91	1.70	3.16
Report subjects' last disposal (non-custodial)	0.12 (0.25)	0.69	1.13	1.84
Report subjects' ethnicity (Asian)	0.29 (0.42)	0.58	1.33	3.04
Report subjects' ethnicity (Black)	0.30 (0.31)	0.73	1.35	2.49

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Report subjects' ethnicity (Mixed Ethnic Group)	-0.22 (0.41)	0.36	0.80	1.79
Report subjects' ethnicity (Other)	0.01 (0.49)	0.39	1.01	2.62

Note: $R^2 = 0.08$ (Cox-Snell), 0.11 (Nagelkerke) Model $\chi^2 (26) = 71.24, p = .001$

* $p < .05$, ** $p < .01$, *** $p < .001$

Appendix J

Results of multinomial regression analyses examining the influence of the intervention, and interaction between the intervention and report subjects' ethnicity, on orientation and weight of sentencing proposals

Table 9. Summary of multinomial logistic regression analysis examining impact of intervention, and interaction between intervention and ethnicity, on sentencing proposal orientation

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Rehabilitative vs. Punitive sanctions				
Intercept	-0.92 (0.62)			
Report subjects' age	0.01 (0.01)	0.99	1.01	1.03
Report subjects' risk of reoffending	0.03 (0.01) ^{***}	1.01	1.02	1.04
Report subjects' gender (female)	0.65 (0.26) ^{**}	1.14	1.91	3.20
Report subjects' index offence (acquisitive)	0.09 (0.38)	0.52	1.10	2.30
Report subjects' index offence (drugs)	0.01 (0.46)	0.41	1.01	2.47
Report subjects' index offence (motoring)	-0.25 (0.29)	0.44	0.78	1.36
Report subjects' index offence (robbery, sex or violence)	0.80 (0.28) ^{**}	1.29	2.24	3.87
Report subjects' last disposal (no previous offence)	-0.55 (0.37)	0.28	0.58	1.18
Report subjects' last disposal (absolute or conditional discharge)	0.44 (0.73)	0.38	1.55	6.44

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Report subjects' last disposal (custody)	-0.16 (0.39)	0.40	0.86	1.83
Report subjects' last disposal (community order)	0.27 (0.37)	0.63	1.31	2.72
Report subjects' last disposal (suspended sentence order)	0.28 (0.45)	0.54	1.32	3.22
Report subjects' ethnicity (ethnic minority)	-0.24 (0.26)	0.48	0.79	1.30
Experimental condition (control)	-0.02 (0.20)	0.66	0.98	1.46
Report subjects' ethnicity*experimental condition	-0.43 (0.42)	0.29	0.65	1.48
Equal vs. Punitive sanctions				
Intercept	-1.97 (0.92)*			
Report subjects' age	-0.01 (0.01)	0.97	0.99	1.01
Report subjects' risk of reoffending	0.01 (0.01)	1.00	1.01	1.02
Report subjects' gender (female)	0.49 (0.34)	0.84	1.63	3.16
Report subjects' index offence (acquisitive)	0.31 (0.59)	0.43	1.37	4.33
Report subjects' index offence (drugs)	1.11 (0.58)	0.96	3.02	9.45
Report subjects' index offence (motoring)	0.49 (0.44)	0.69	1.63	3.83
Report subjects' index offence (robbery, sex or violence)	1.02 (0.44)*	1.17	2.76	6.48
Report subjects' last disposal (no previous disposal)	0.29 (0.59)	0.42	1.33	4.22
Report subjects' last disposal (absolute or conditional discharge)	1.31 (0.94)	0.59	3.70	23.33
Report subjects' last disposal (custody)	0.45 (0.63)	0.46	1.57	5.38
Report subjects' last disposal (community order)	1.14 (0.59)*	0.99	3.12	9.84
Report subjects' last disposal (suspended sentence order)	0.70 (0.69)	0.25	2.02	7.85

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Report subjects' ethnicity (ethnic minority)	-0.37 (0.35)	0.35	0.69	1.37
Experimental condition (control)	-0.44 (0.29)	0.37	0.64	1.13
Report subjects' ethnicity*experimental condition	0.96 (0.52)	0.94	2.61	7.25

Note: $R^2 = 0.21$ (Cox-Snell), 0.25 (Nagelkerke) Model χ^2 (24) = 195.68, $p = .001$

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 10. Summary of multinomial logistic regression analysis examining impact of intervention, and interaction between intervention and ethnicity, on sentencing proposal weight

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Light vs. Heavy sanctions				
Intercept	-2.37 (1.12)*			
Report subjects' age	0.38 (0.02)**	1.01	1.04	1.07
Report subjects' risk of reoffending	-0.03 (0.01)**	0.96	0.98	0.99
Report subjects' gender (female)	0.62 (0.41)	0.83	1.86	4.17
Report subjects' index offence (acquisitive)	0.66 (0.68)	0.51	1.94	7.33
Report subjects' index offence (drugs)	1.80 (0.77)*	1.34	6.02	27.10
Report subjects' index offence (motoring)	-0.20 (0.54)	0.29	0.82	2.37
Report subjects' index offence (robbery, sex or violence)	0.12 (0.51)	0.42	1.13	3.05

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Report subjects' last disposal (no previous disposal)	0.71 (0.64)	0.58	2.03	7.18
Report subjects' last disposal (custody)	0.24 (0.56)	0.43	1.27	3.81
Report subjects' ethnicity (ethnic minority)	0.34 (0.36)	0.70	1.40	2.83
Experimental condition	0.002 (0.32)	0.54	1.00	1.87
Medium vs. Heavy sanctions				
Intercept	0.52 (0.59)			
Report subjects' age	0.02 (0.01)*	1.00	1.02	1.04
Report subjects' risk of reoffending	-0.02 (0.01)**	0.98	0.99	0.99
Report subjects' gender (female)	0.10 (0.26)	0.66	1.10	1.84
Report subjects' index offence (acquisitive)	0.65 (0.38)	0.90	1.91	4.05
Report subjects' index offence (drugs)	1.09 (0.55)*	1.01	2.98	8.82
Report subjects' index offence (motoring)	0.29 (0.30)	0.74	1.33	2.40
Report subjects' index offence (robbery, sex or violence)	0.20 (0.29)	0.70	1.23	2.14
Report subjects' last disposal (no previous disposal)	0.48 (0.32)	0.87	1.62	3.01
Report subjects' last disposal (custody)	0.11 (0.25)	0.68	1.11	1.81
Report subjects' ethnicity (ethnic minority)	0.14 (0.21)	0.76	1.15	1.73
Experimental condition	-0.14 (0.18)	0.61	0.87	1.23

Note: $R^2 = 0.08$ (Cox-Snell), 0.10 (Nagelkerke), Model χ^2 (22) = 68.60, $p = .001$

* $p < .05$, ** $p < .01$, *** $p < .001$

Appendix K

Results of multinomial regression analyses examining the influence of completing the intervention as intended on orientation and weight of sentencing proposals

Table 11. Summary of multinomial logistic regression analysis examining impact of the intervention when completed as intended on sentencing proposal orientation

	<i>b</i> (<i>SE</i>)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Rehabilitative vs. Punitive sanctions				
Intercept	-1.54 (66)*			
Report subjects' age	0.01 (0.01)	0.99	1.01	1.03
Report subjects' risk of reoffending	0.03 (0.01)***	1.02	1.03	1.04
Report subjects' gender (female)	0.54 (0.29)	0.97	1.71	3.01
Report subjects' index offence (acquisitive)	0.23 (0.41)	0.56	1.26	2.84
Report subjects' index offence (drugs)	0.27 (0.50)	0.49	1.31	3.45
Report subjects' index offence (motoring)	-0.15 (0.31)	0.47	0.86	1.57
Report subjects' index offence (robbery, sex or violence)	0.82 (0.30)**	1.26	2.27	4.10
Report subjects' last disposal (no previous offence)	-0.07 (0.35)	0.47	0.93	1.87
Report subjects' last disposal (non-custodial)	0.70 (0.31)*	1.09	2.02	3.73
Report subjects' ethnicity (ethnic minority)	-0.29 (0.22)	0.49	0.75	1.15

	<i>b</i> (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Experimental condition (intervention as intended)	-0.07 (0.21)	0.61	0.93	1.41
Equal vs. Punitive sanctions				
Intercept	-1.90 (0.91)*			
Report subjects' age	-0.00 (0.01)	0.97	1.00	1.02
Report subjects' risk of reoffending	0.01 (0.01)	0.99	1.01	1.02
Report subjects' gender (female)	0.38 (0.37)	0.72	1.47	3.02
Report subjects' index offence (acquisitive)	0.48 (0.60)	0.50	1.61	5.24
Report subjects' index offence (drugs)	0.93 (0.63)	0.74	2.53	8.64
Report subjects' index offence (motoring)	0.38 (0.45)	0.61	1.46	3.49
Report subjects' index offence (robbery, sex or violence)	0.96 (0.44)*	1.09	2.61	6.23
Report subjects' last disposal (no previous disposal)	0.14 (0.51)	0.42	1.16	3.15
Report subjects' last disposal (non-custodial)	0.85 (0.46)	0.95	2.33	5.72
Report subjects' ethnicity (ethnic minority)	-0.01 (0.28)	0.57	0.99	1.71
Experimental condition (intervention as intended)	-0.05 (0.28)	0.55	0.95	1.63

Note: $R^2 = 0.19$ (Cox-Snell), 0.22 (Nagelkerke) Model χ^2 (22) = 146.56, $p = .001$

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 12. Summary of multinomial logistic regression analysis examining impact of the intervention when completed as intended on sentencing proposal weight

	<i>b (SE)</i>	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Light vs. Heavy sanctions				
Intercept	-1.84 (1.28)			
Report subjects' age	0.04 (0.02)*	1.00	1.04	1.07
Report subjects' risk of reoffending	-0.03 (0.01)**	0.95	0.97	0.99
Report subjects' gender (female)	0.54 (0.49)	.066	1.71	4.44
Report subjects' index offence (acquisitive)	0.99 (0.73)	0.66	2.70	11.17
Report subjects' index offence (drugs)	1.18 (0.92)	0.54	3.26	19.86
Report subjects' index offence (motoring)	-0.25 (0.60)	0.24	0.78	2.50
Report subjects' index offence (robbery, sex or violence)	0.02 (0.56)	0.34	1.02	3.05
Report subjects' last disposal (no previous disposal)	0.48 (0.73)	0.38	1.62	6.81
Report subjects' last disposal (non-custodial)	-0.03 (0.64)	0.28	0.97	3.43
Report subjects' ethnicity (ethnic minority)	0.28 (0.41)	0.60	1.32	2.93
Experimental condition (intervention as intended)	-0.46 (0.43)	0.28	0.63	1.46
Medium vs. Heavy sanctions				
Intercept	0.83 (0.66)			
Report subjects' age	0.02 (0.01)	1.00	1.02	1.04
Report subjects' risk of reoffending	-0.02 (0.01)***	0.97	0.98	0.99
Report subjects' gender (female)	0.22 (0.29)	0.70	1.24	2.21

	b (SE)	95% CI for Odd Ratio		
		Lower	Odds Ratio	Upper
Report subjects' index offence (acquisitive)	0.74 (0.42)	0.92	2.10	4.83
Report subjects' index offence (drugs)	1.10 (0.61)	0.83	2.75	9.16
Report subjects' index offence (motoring)	0.31 (0.33)	0.72	1.36	2.59
Report subjects' index offence (robbery, sex or violence)	0.18 (0.31)	0.65	1.19	2.19
Report subjects' last disposal (no previous disposal)	0.32 (0.35)	0.70	1.38	2.73
Report subjects' last disposal (non-custodial)	0.70 (0.28)	0.62	1.07	1.85
Report subjects' ethnicity (ethnic minority)	0.07 (0.23)	0.69	1.07	1.68
Experimental condition (intervention as intended)	-0.23 (0.21)	0.53	0.79	1.20

Note: $R^2 = 0.08$ (Cox-Snell), 0.10 (Nagelkerke), Model χ^2 (22) = 56.65, $p = .001$

* $p < .05$, ** $p < .01$, *** $p < .001$