Taking an evidence-based approach to evidence-based policing research.

Author details

Laura Boulton (corresponding author)

School of Justice Studies, Liverpool John Moores University, Liverpool, United Kingdom, L3 5UZ, email: l.j.boulton@ljmu.ac.uk

ORCID: https://orcid.org/0000-0001-5623-8884

Twitter: @lj_boulton

LinkedIn: laura-boulton-13708066

Rebecca Phythian

School of Forensic and Applied Sciences, University of Central Lancashire, Preston, Lancashire, United Kingdom, PR1 2HE, tel: 01772894251, email: rphythian@uclan.ac.uk

ORCID: https://orcid.org/0000-0001-6423-2116

Twitter: @beckyphythian

LinkedIn: rphythian

Stuart Kirby

Crime Insights Ltd., Liverpool, email: skirby@crimeinsights.ltd

ORCID: https://orcid.org/0000-0002-3049-1248

Ian Dawson

Lancashire Constabulary, Corporate Development Department, Hutton, Preston, PR4 5SB, tel: 01772 412867, email: Ian.dawson@lancashire.pnn.police.uk

Author Biography:

Both Laura Boulton and Stuart Kirby were working at the University of Central Lancashire when this research was conducted.

Acknowledgements: The authors are very grateful to Lancashire Constabulary for their facilitation of this research. The authors would also like to give special mention to Dr. Emily Cooper who assisted with data coding.

Introduction

Whilst historically, police practice was founded on experience rather than conscious design and evaluation (Huey and Ricciardelli, 2016), the 21st Century has seen evidence-based approaches being used to improve decision-making (Huey and Mitchell, 2016). This evolution emerged from a widespread recognition that strategies founded upon scientific knowledge and 'what works' are more effective than subjective 'hunches' (Lum, 2009; Lumsden and Goode, 2018). Whilst numerous law enforcement agencies across the world have invested significant resources in Evidence Based Policing (EBP), the impact of this effort remains ambiguous. This article seeks to systematically explore research-based projects generated by a police force situated in the North of England, as it continues to develop its approach to EBP, examining the impact this research has had on policing.

The article is organised in the following way:

- The literature review sets out the rationale for EBP and explains the types of research evidence and methodological designs that are available to base these approaches on in order to provide a contextual understanding for the methodology and coding structure used in this paper. The literature review also explores how the police actually implement research findings into their practice and why this might be.
- The method section describes how this study evaluated a sample of EBP research projects conducted within one police force. Authors systematically coded a sample of studies conducted within this force in terms of: (i) the context in which the study was conducted; (ii) the design and rigour of the methodology used in each study; (iii) how the results of the studies were disseminated; and (iv) how the results were implemented (if they were) into practice and to what extent.
- The results from this evaluation are presented in four sections, each examining the four elements listed above.

- Lastly, the paper discusses how these results can be used to inform the design and planning of EBP research projects in order to increase the chance of implementation of findings to impact on real world police practice.

Literature Review

EBP promotes the use of research, evaluation, analysis and scientific processes to direct police decision-making (Lum and Koper, 2015). Sherman (1998) posited two ways of implementing EBP practice;

(i) using the results of scientifically rigorous evaluations of law-enforcement tactics and strategies to guide decisions, and

(ii) generating and applying analytic knowledge derived from an agency's analysis of its own internal issues and crime problems.

This provides a rationale for either; (a) continuing to implement strategies shown to have a positive impact, or (b) discontinuing those interventions found to be ineffective or damaging (Chalmers, 2003; Lum and Koper, 2015).

The reasons for a police force engaging in EBP are straightforward. Research-based approaches are viewed as generating more 'intelligent policing', whereby crime reduction becomes synonymous with reduced demand on officers and staff (Lum, 2009); indeed, the process of recording, accessing, analysing and managing police data promotes objectivity, transparency, legitimacy and accountability in relation to police governance and spending (Sherman et al., 2002; Chalmers, 2003). Similarly, adopting an evidence-based mindset adds integrity and accountability in decision-making, making practitioners more likely to challenge existing practice (Sherman and Murray, 2015).

EBP approaches have become pervasive across Europe, Australasia and parts of America (Knutsson and Tompson, 2017). Furthermore, developments in practitioners

utilising an evidence-based approach are apparent, such as those termed ‘evidence cops’ and ‘pracademics’ (Sherman, 2015; Lum and Koper, 2017). However, whilst numerous law enforcement agencies have invested significant resources in this approach, the impact of this effort remains ambiguous. To understand this conundrum, studies have typically explored practitioner attitudes to EBP. Despite finding that the concept is popular, the ability to transform it into mainstream operational practice is still fraught with difficulties. Therefore, this article takes a different approach by examining the internal process one particular police force follows. To do this, a number of operational challenges associated with EBP should first be understood.

Not all research is created equally

One of the obvious complications that police agencies face in implementing EBP is establishing what constitutes evidence. This can be complex as research projects comprise many methodological designs, are associated with varying levels of rigour, and generate various levels of opportunity for impact on practice (Lum and Koper, 2015). Methodological approaches can broadly be grouped into qualitative and quantitative research (for a detailed discussion of methodological approaches to EBP, see Brown, 2019). Qualitative research typically involves the analysis of data which is non-numerical and is concerned with an understanding of a specific context or concept (Howitt, 2012). Specifically, qualitative methods may be adopted ‘where there is a paucity or lack of research into a particular research topic’ (Howitt, 2012, p.135). As such, it mainly uses fewer participants, purposively sampled based on a specific set of inclusion criteria. In contrast, quantitative research focusses on numerical data, often associated with large sample sizes (see Sheldon et al., 2012). This is not to suggest that one methodological approach competes with the other, as qualitative data compliments and provides ‘substance to quantitative findings’ (Howitt, 2012,

p. 154). Brown (2019) suggests that choosing from different methodological approaches was generally determined by epistemological differences, disciplinary fragmentation, competitiveness and entrepreneurialism.

Conventional measures to establish the quality of quantitative research typically include: validity (how well a piece of research reflects the reality it claims to represent), reliability (the repeatability of the findings) and empirical generalisability (extent to which findings can be applied to settings other than those of the study). The Maryland Scale of Scientific Methods (Sherman, 2013) aims to quality assure the methodological validity of individual studies using a quantitative methodology (Lumsden, 2016). It uses a five-level hierarchy ranging from a simple before and after assessment (least reliable) to a Randomised Controlled Trial (most reliable: Kirby, 2013). Whilst widely considered the gold standard of research (Sherman, 1998; Heaton and Tong, 2015), some are concerned that Randomised Controlled Trials lack external validity when generalising to other contexts (Brown et al., 2018), because they may not be theoretically grounded or able to identify a causal mechanism explaining why the intervention did or did not work (Hough, 2010). Similarly, the police must balance the level of resources invested into an evaluation with the level of resources available for the intervention.

The criteria used to understand quantitative research cannot be applied to qualitative research because of inherent differences in terms of research philosophy, aims, design, sampling and analysis. Instead, Kitto, Chester and Grbich (2008) have defined criteria that can strengthen the rigour of a qualitative research project, which can potentially be used to evaluate the quality of a qualitative research project. It includes:

- Clarity and justification of the research question: aims and methods used are in alignment with one another;

- Procedural rigour (description of the way the research was conducted should be transparent);
- Representativeness (sampling technique should recruit participants who are relevant to the research question);
- Interpretative rigour (the use of inter-rater reliability in the coding of data or triangulation of multiple data collection methods);
- Reflexivity (openly acknowledge and address the influences impacting on the results, i.e. researcher's own biases); and
- Transferability (relevance of findings for current knowledge, policy or practice).

How police agencies generate and implement research

The second element to be explored is the implementation and impact of EBP research. Whilst there is significant investment focussed towards EBP worldwide (Brown et al., 2020), particularly in Britain, Australia, New Zealand and the USA (Sherman, 2015), and closer working relationships evident between practitioners and local Universities, the implementation of research into operational policing remains challenging (Knutsson and Tompson, 2017; Brown et al., 2018). Lum (2009, p. 3) noted that 'there is little indication that most American police leaders and their agencies systematically or regularly use tactics that are evidence-based', whilst Telep and Winegar (2016) found just over half of their sample of American police leaders used research only 'sometimes'. Cherney and colleagues (2018) explored the perceptions of senior Queensland police staff towards EBP; despite some promising findings, such as the majority of participants claiming they understood EBP (62%), a much smaller proportion reported implementing outcomes from research into practice (22%). The UK shows similar results, with limited use of research to inform initiatives, especially amongst front line officers (Palmer, Kirby and Phythian, 2019). In fact, Fleming

and Wingrove (2017) identified a decline in eagerness following attendance at an EBP workshop, associated with a perceived lack of implementation support within the force. In essence, these studies show a gap between the high level of research and its use by police forces (Lum, 2009; Weisburd and Neyroud, 2011; Huey et al., 2018).

The attrition of research implementation is accounted for by three general explanations. First it can be difficult, especially for police officers and staff without academic training, to determine what counts as evidence of effective practice (Sherman, 2015; Lumsden and Goode, 2018). Definitions of what counts as evidence can vary between policing and academia. Whilst researchers focus on large data samples, police officers tend to focus on more memorable outliers (Canter, 2004; Telep and Lum, 2014). Secondly, research findings are often disseminated in ways that are inaccessible to practitioners (Bullock and Tilley, 2009; Snilstveit et al., 2016). The sheer number of published studies, with varied and nuanced results, make it difficult for police practitioners to establish what results to base their decisions on (Lum, 2009). Further, police research studies are typically written in a detailed academic format and subjected to a slow publication process. Access is often limited through Higher Education portals, which conflicts with the fast-paced and dynamic nature of policing (Heaton and Tong, 2015). Finally, it has been highlighted that the organisational culture surrounding policing may be resistant to promote EBP into practice. Essentially, there are fundamental distinctions between police and academics in terms of their 'thought processes, typical modes of action and the central objectives', particularly in relation to what 'knowledge' is perceived to be (Canter, 2004, p.111-112). Palmer et al. (2019) reported front line officers were more likely to report it as another 'fad', whilst Heaton and Tong (2015, p. 61) suggested that, as the police are operationally goal-focussed, they abandon scientific methodology in favour of the 'necessity to demonstrate success'. For research to add value to

policing and to be ‘taken seriously’, there must be evidence of its operational impact (Canter, 2004, p.114).

Although studies have found police officers are generally receptive to the idea of EBP (Lumsden, 2016; Telep and Winegar, 2016), it appears that there is a gap in awareness and knowledge of EBP in frontline and operational police practitioners (Telep and Lum, 2014). As such their day-to-day decision-making is guided by their experience, which they value more highly than research-based evidence (Canter, 2004; Lumsden, 2016; Telep and Winegar, 2016). Therefore, officers may be more receptive to qualitative research which focuses on experiences and perceptions in the words of real people, rather than the abstract nature of statistical analyses underlying quantitative research (Koehle et al., 2010).

Whilst research exists which evaluates a single police initiative or programme, no current study considers how the combination of context, research design rigour, and dissemination of findings create an operational impact. Neither is there any process where police forces could be objectively compared in relation to their progression in EBP. To address this gap, this study explores the efficacy of EBP as an operating philosophy in relation to practical implementation issues. By conducting a detailed examination of a sample of EBP projects from one police force, this paper aims to better understand the nature of EBP in terms of the research that underpins it and how this is implemented into practice. To do so, it explores the use of research studies conducted by a Constabulary in the North of England and the impact this research was having on EBP practice both within and external to this Constabulary.

Method

Data

This study was situated in a Constabulary based in the North of England, which has approximately 5000 members of staff. The Constabulary has been undertaking coordinated EBP research since 2009 and in 2017 it formed a dedicated EBP unit with the aim of using more research-based practice in policing. As a result of academic collaborations, the unit recorded 45 projects between 2009 and 2018. In addition, Constabulary policy facilitated dozens of PhD and MSc research projects by staff members on policing issues, supervised by academics.

The purpose of the methodology was to design a process that could understand the quality of the research project before then examining whether the findings were utilised. A number of Subject Matter Experts (i.e. a Professor in Policing, a Superintendent) were consulted before data collection began in terms of appropriate measures of evaluation. Possible evaluation measures were initially drawn from literature, followed by a process of refining such measures based on academic and practitioner input. After critical discussions, the researchers created and re-iterated multiple versions of a Policing Research Evaluation Framework (PREF: see Figure 1) which was then used to collect, organise and populate the dataset. This used four categories (detailed below): context, type of research, dissemination and implementation, which were further subdivided allowing any police force to profile the quality and impact of its internally commissioned research.

The criteria for a project to be included in the sample were: evidence of a systematic or methodological approach and having been completed with sufficient time to have the findings reviewed, implemented and/or disseminated (i.e. three months +). 29 studies were selected, thereby providing secondary data for analysis. The projects included a range of topics, such as game-based learning for police training in child interviewing, an exploratory analysis of the use of body worn video, ethical practice in policing and patterns of repeat domestic abuse. The projects within the sample also ranged in terms of how they were funded

and who led them. The sample included commissioned research, MSc and PhD projects. In terms of commissioned research, this encompassed:

(i) EBP related research granted funding via external funding bodies (i.e. the Home Office) and delivered by academic partners;

(ii) Projects funded and conducted by academic partners for which the force provided data; and

(iii) Force funded projects in which academic partners and/or police staff were commissioned to conduct a piece of research or evaluation.

MSc and/or PhD projects typically involved either secondary datasets provided to the student by the force or police relevant primary data collected by the student themselves, and were either funded by Higher Education Institution partners and conducted by non-police students (i.e. University of Liverpool Forensic Psychology MSc programme etc.), or funded by the force itself and conducted by police staff (i.e. University of Cambridge and University of Central Lancashire EBP related Masters programmes etc.). Basic contextual information and descriptions of each study were gathered, and the associated final research reports were supplied by the Constabulary. The researchers assessed the report for information relating to the methodological design and rigour of the studies. Primary authors (practitioners and/or academics) of the research projects were contacted to gain an understanding of if and how the findings of each study had been disseminated and if the findings had been implemented to generate any impact on police practice or policy. This information was given qualitatively; further research papers and reports were provided, with outstanding details confirmed via email by primary authors.

The qualitative information was then thematically coded where possible into categories which were numerically coded for statistical analysis (described further below). A

systematic approach was adopted to extract the necessary details from the research projects; the variables (see Figure 1) were coded dichotomously (present/ not present). Inter-rater reliability was conducted for the thematic analysis and an acceptable level of agreement was found (> 0.6).

[Figure 1 near here]

Measures

Information was coded into variables relating to four broad categories: (i) context, (ii) research design and rigour, (iii) dissemination, and (iv) implementation. The proposed categories were formed on the basis of literature, as well as the input and experience of academic and practitioners. The researchers perceived the 'context' to be useful to capture standard contextual information of research. Contextual information was organised and coded into variables relating to the purpose of the research, the type of research project (e.g. MSc dissertation), key contact/author, length and cost of research as well as cost bearer of the research.

'Research design and rigour' were deemed to be critical to gaining an understanding of the type and quality of the research conducted. Design and rigour of the research was assessed in relation to the type of data used (e.g. primary), the analysis approach used (e.g. mixed method) and the design of the methodology. To review the rigour of the methods used, researchers considered where the project sat in relation to the Maryland scale (Sherman, 2013) and whether the project demonstrated:

(i) Objectivity: methods for reducing bias during data collection and/or analysis were described (i.e. inter-rater reliability, data triangulation, member checks);

(ii) Reproducibility: the detail provided was transparent enough to allow the method to be replicated and the findings reproduced;

(iii) Plausibility: the sample or data used are relevant and appropriate to achieve the research aims and objectives, and;

(iv) Transferability: findings can be transferred, or generalised, to other contexts (i.e. applying the findings to other forces).

The studies were considered in relation to dissemination of the findings. Information pertaining to 'dissemination' was necessary to achieve the aims of the current research, identifying key methods to disseminate research in academia and practice. Dissemination was coded dichotomously (yes/no) in relation to whether the findings had been disseminated via an internal report, an internal/stakeholder input, a practitioner conference, an academic conference, in the press/social media/intranet or published as a peer-reviewed journal article.

Finally, information was collected and coded in relation to how much impact the research had had on police practice by assessing whether any findings had been implemented (i.e. as a change to practice and/or policy) and if so, to what extent (i.e. within one department only, within force more broadly, across multiple forces, nationwide or internationally). Again, gathering details relating to implementation was vital to achieve the aim of the research. To do so, the reasons why the findings from some studies had not been implemented were coded in the three categories as described by Kirby (2013): (i) theory failure (study was found to be based on an invalid or inappropriate theory, idea, or strategy), (ii) implementation failure (insufficient care taken in relation to the staff, resources and other factors necessary for the effective delivery of the project), and (iii) evaluation or measurement failure (evidence of poor design, inappropriate comparison groups, or used a measure which was insensitive to change).

Analysis

Firstly, descriptive analysis was conducted to generate a greater understanding of the trends and patterns visible within the dataset and to guide further inferential analysis. Kruskall Wallis tests were used to examine any differences in length of time taken to complete a project across different factors. Chi-square tests were conducted to inferentially determine whether associations between the categorical variables coded were statistically significant. Each 2x2 chi-square analysis examined the association between a characteristic of the project (i.e. mixed methods) and a rigour measurement category (i.e. has objectivity); each variable was coded dichotomously (present; not present). Caution must be considered when interpreting these findings due to the small sample size (n = 29). Only significant differences and associations are reported here, all other findings were found to be non-significant.

Results

Context

The purpose of the projects in the sample were: (i) exploring a specific problem (n = 16, 55%); (ii) evaluation of a tool/initiative (n = 8, 28%); (iii) understanding the views of a sample of people on a particular topic (n = 5, 17%). Most of the projects were MSc research projects (n = 16, 55%), followed by commissioned research (n = 10, 34%) and finally PhD theses (n = 3, 10%). Out of the 29 projects, the exact cost of only one project could be determined. Overall, the cost bearer was most often a university/student (n = 11, 38%) or the Constabulary (n = 10, 34%). Only two projects (7%) were externally funded (i.e. Police Innovation Fund) and in six (21%) cases, the cost bearer was unknown. The type of project was significantly associated with objectivity, X^2 (2, n = 29) = 10.712, p < .01, and

reproducibility, X^2 (2) = 17.744, $p < .001$. Commissioned projects were significantly related to a lack of objectivity (7.1%) and reproducibility (5.6%) whereas the opposite was true with MSc projects (objectivity, 85.7%; reproducibility.77.8%).

Design and rigour

Descriptive statistics are illustrated in Table 1. Most of the research used secondary data , whilst primary data was used in approximately one quarter of cases and a mixture of both primary and secondary was used in fewest cases. The data was analysed using quantitative analysis in the majority of projects, followed by mixed methods analysis and qualitative analysis used the least (see Table 1).

[Table 1 near here]

Out of the 23 projects using quantitative methods and analysis of some sort, the majority could only be described as a level 1 on the Maryland scale, followed by four level 2 studies and one level 3 study. No studies could be seen to reflect level 4 or level 5 on the Maryland scale. To more appropriately reflect the designs of the studies sampled, we examined the projects according to whether the design demonstrated objectivity, reproducibility, plausibility and transferability. Out of the 29 studies sampled approximately half demonstrated objectivity and transferability, over 60% were reproduceable and all were considered plausible (see Table 1).

Data type was significantly associated with objectivity, X^2 (2) = 10.216, $p < .01$, reproducibility, X^2 (2) = 9.836, $p < .01$, and transferability, X^2 (2) = 8.368, $p < .05$. Specifically, studies using secondary data were associated with having objectivity and having reproducibility, but not having transferability (see Table 2). Data type was also significantly associated with being disseminated, X^2 (2) = 6.751, $p < .05$, and implemented, X^2 (2) = 8.368,

$p < .05$. Studies using secondary data were less likely to be disseminated at an academic conference, to change practice or to implement changes within department or in force.

[Table 2 near here]

The method used in the analysis was significantly associated with objectivity, X^2 (2) = 13.813, $p < .01$, reproducibility, X^2 (2) = 9.222, $p < .01$, transferability, X^2 (2) = 14.419, $p < .01$, dissemination, X^2 (2) = 8.235, $p < .05$, and implementation, X^2 (2) = 14.419, $p < .01$. Specifically, studies using qualitative analysis were associated with having no objectivity, whilst quantitative analysis studies were associated with having objectivity and reproducibility but no transferability. Comparatively, studies using mixed method analysis were associated with having transferability. Qualitative research was less likely to be disseminated generally, whilst quantitative studies were less likely to be presented at a conference. Lastly, mixed studies were associated with being presented at a conference and generating a change to practice (see Table 3).

[Table 3 near here]

Dissemination

Over 90% of projects had been disseminated to some extent but the level of dissemination varied across the sample (see Table 1).

Transferable studies were associated with force/stakeholder input disseminations and presentations at practitioner conferences yet were less likely to disseminate the findings via the press/social media/intranet (see Table 4). Dissemination was found to be associated with implementation in multiple ways. Findings that had been disseminated via a force/stakeholder input and a practitioner conference were found to be significantly associated with having

been implemented and resulted in a change to practice that was felt within department and within force.

[Table 4 near here]

Implementation

It was established that the findings from half of the 29 projects were implemented in some way. This implementation resulted in changes to practice on most occasions, with a couple of instances of changes to policy (see Table 1). From the studies which generated findings that had been implemented, themes from those implementation modes can be found in Table 5 (presented with examples of each category from within this sample).

[Table 5 near here]

Out of those implemented studies, the changes that were implemented were felt within the specific force on most occasions, across forces in approximately one quarter of projects, and nationwide once. No studies had generated impact that had been implemented internationally (see Table 1).

Out of those 14 (48%) projects which findings were not implemented, the failure to implement findings was cited as an implementation failure in approximately three-quarter of cases and as a measurement failure in one-fifth of cases (see Table 1). For implementation failure cases, reasons included: (i) lack of funding or resources to implement recommendations; (ii) high turnover of senior management, and (iii) a lack of project ownership once the research report is delivered and the owner changes departments. For measurement failure, reasons included a lack of adherence to a commissioned research plan (i.e. late submission of report or pulled out of project before the end of the research).

Objectivity was significantly associated with implementation: studies without objectivity were associated with findings that have been implemented and had generated a change to practice that was felt within the department and within the force. On the other hand, studies that had transferability were more likely to be implemented (100% of transferable studies were implemented) and instigate a change in practice, within the department and the force (see Table 6).

[Table 6 near here]

Discussion

Public sector organisations have a responsibility to use their resources effectively and efficiently. Literature questions whether the popularity and investment in EBP has materialised into mainstream policing and delivered the operational impact initially envisaged. Implementation failure is a common concept across many disciplines, including policing (Kirby, 2013). Therefore, this study took a new approach to examine what EBP means in practice. Specifically, it examined the type of research the police are developing, the extent to which they use the evidence generated from such work, and what factors influence the implementation of this research knowledge. Using this information, it provides a common framework to enable police forces to more objectively assess their progress in EBP.

The results show no simple formula exists to guarantee research makes a significant impact to practice and/or policy. For instance, whilst PhD projects take significantly longer to complete than MSc or commissioned projects, there was no clear evidence that PhD projects are significantly associated with wider dissemination, implementation or impact. Furthermore, studies which demonstrate less evident objectivity (i.e. commissioned projects),

were associated with findings that have been implemented and had generated a change to practice, most often as an approach amendment, that was felt within the department and within the force. Therefore, it seems that commissioned projects, despite demonstrating less objectivity, generate a greater practical impact. This may reflect the nature of their aims: often they are commissioned to address a specific need or gap previously identified as an area of need for support by the force. In this sense they are examples of co-production, in which practitioners and academics work together to design research with specific aims around a particular intervention. The force may therefore be more receptive to these findings as they have specifically asked for answers in relation to that topic. This would support past research which argue that co-production increases the possibility of police theory and research advancement when police officers are actively involved in the process of research design and interpretation (Wood et al., 2008; Goode and Lumsden, 2018). In contrast, MSc projects tend to be very objective and reproducible (likely to reflect the nature of the assignment brief associated with them, i.e. presentation guidelines), but were not associated with implementation which may result from a more theoretical focus.

Some patterns emerged when considering the impact of data type and analysis approach on outcomes of the research findings. For instance, whilst secondary data was associated with objectivity and reproducibility, the findings were not associated with being transferrable to other domains and, in line with this, were not found to have been implemented. This may reflect the nature of studies which typically use secondary data (i.e. MSc projects) and the lack of control that the researcher has in projects that analyse such data. In terms of analysis, whilst quantitative studies were associated with having objectivity and reproducibility, they had no transferability and the findings were not implemented and did not generate changes to practice. Furthermore, qualitative studies were associated with a lack of objectivity. Whilst the nature of qualitative research makes the separation of the

researcher, and therefore objectivity, inherently difficult, it appears that methods of bias mitigation (i.e. inter-rater reliability) were often not employed within this sample's qualitative studies.

In contrast, it appears that studies which use mixed methods were associated with transferability as well as findings which were implemented as a change to practice. From this it appears that research which utilises mixed methods generates more positive and impactful results, potentially relating to the 'messiness' of police data and the reassurance that triangulation of methods brings to the interpretation of generated findings (Schulenberg, 2007; Tilley and Laycock, 2017). This supports Brown et al. (2018, p.45) who suggest that a mixed method approach to EBP research enables the 'fleshing out' of situational and contextual factors that are often ignored in purely quantitative approaches to cross-validate statistical findings.

Transferability appeared to be a key feature of a study's reach in terms of dissemination and implementation. Transferable studies were disseminated more than non-transferable studies and the findings from 100% of transferable studies were implemented as a change to practice. This understandably suggests that studies which have a broader scope in terms of topic, reach a wider audience who apply that knowledge to their practice.

Lastly, dissemination was a strong indicator of implementation and practical impact of research. This makes intuitive sense: those findings that were disseminated were more likely to be implemented, whilst those that were not were associated with implementation failure. This effect could reflect either that dissemination increases implementation, or that better studies/studies which show the most useful findings are disseminated more. Either way, this gives strong backing to the recommendation of dissemination of findings. Specifically, force/stakeholder inputs appear to be an important way to share learning from

research which leads to implementation. Furthermore, in this sample, studies suffering from implementation failure issues were reported to have a lack of dissemination due to a high turnover of senior management. For example, if a senior leader who was encouraging of a specific research study at the start of the project had moved on to another department, role or division by the end of the research study, it was felt that the impetus of dissemination and implementation was lost. This suggests that implementation failure, because of poor dissemination, reflects a lack of organisational memory.

Limitations

The methodological rigour section of the framework was based on the researcher's evaluation of the information available. In relation to the methodological information available, the researchers observed that this differed depending on the purpose of the project; if the research was conducted for an MSc dissertation, the individuals were required to provide specific information for the assessment, compared to an internal practitioner report focused on recommendations to practice and containing less methodological content. Therefore, the available details of the research projects may not accurately reflect the rigour of the method. Moreover, the researchers experienced some difficulties in collecting data relating to the dissemination and implementation sections of the framework. Finally, the findings should be interpreted with caution due to the small sample used (n = 29) and the use of data from one UK police force, and therefore may not be representative of EBP work in other forces. Nevertheless, this exploratory analysis has identified key learning and recommendations to support the dissemination and implementation of future EBP research. This is essential if evidence-based practice is to be fully embedded and if the full value of adopting this approach is to be appreciated.

Implications

Policing knowledge is a much wider concept than solely academic knowledge; it also includes tactical skills, understanding the practical and political context, as well as individual and collective decision-making processes. Knowledge-for-knowledge sake can be good for student development, but police research needs to be directed towards implementation. Unless academic partners conducting EBP research understand the practical and contextual issues relating to the implementation of research findings, then the utility of such work is limited. Therefore, co-production relationships between academics and police practitioners should incite professional challenge to generate meaningful insight (Klein and Jarosz, 2011). Such insight is needed regarding which policing problems need an evidence-based solution, as well as insightful interpretations of how the evidence generated can be implemented operationally.

Practically, these findings suggest that EBP research projects can be designed and disseminated in a way that increases the likelihood of implementing the findings to change practice. In terms of design, EBP researchers should consider using mixed methods data collection and analysis techniques and focus on designs that are transferable across different domains. To support transferable research, practitioners should spend time understanding why the research is being commissioned in the first place (what is the problem hoping to be solved) and consider if appropriate systems are in place to support the implementation of the findings that emerge before the work is commissioned.

As for dissemination, results should be disseminated more widely at as many levels as practically possible to engage end users. Focussing on getting the message across to end users (i.e. police officers) is crucial because if they do not know about the findings and/or

recommendations of the research undertaken (or they cannot understand them as they are not tailored to their needs), they cannot implement them. Academics, in particular, should avoid solely focussing on publication of findings in a peer-reviewed journal if they want to have a practical impact with their research. By the time results are published in this way, there will be a loss in relevance of the findings due to the delay associated with peer review publication. What is more, such formats do not promote accessibility to the practitioners who are able to implement the findings to practice. To counter this, research should be disseminated at both practitioner and academic targeted conferences, enabling a more accessible and context-relevant communication of results to relevant audiences.

These recommendations have been fed back to the Constabulary studied in this research and were well-received. To build on these recommendations and to foster an environment in which EBP research is implemented, they have taken some specific practical steps which other forces may similarly find useful. Firstly, they have co-ordinated a strategic EBP group, including both police staff and academics, that meet quarterly to review research project proposals to ensure prioritisation and support through implementation before research projects begin. Secondly, they have designed an EBP Hub page on their internal intranet system to freely share completed research with all police officers and staff. Lastly, they have organised an annual EBP award conference where EBP research findings can be promoted and celebrated.

Conclusion

Overall this study demonstrates the importance of research projects on evidence-based changes to practice, but also highlights the challenges of ensuring the potential impact of these changes are felt within and across departments and forces. For EBP to achieve its goal

of ‘informing decision makers about the best scientific evidence regarding strategies to realise desire outcomes’ (Bueermann, 2012, p. 14), implementation of research findings is key. Without implementation, research is being done for the sake of it, with little theoretical or practical implications or impact, but still at a cost to higher education institutions and police forces. This research has provided a framework which constitutes four categories: context, design and rigour, dissemination and implementation. These are further subdivided allowing any police force to profile the quality and impact of its internally commissioned research. The summary from this analysis suggests that key features of research which are associated with implementation of findings are: (i) mixed method data collection techniques, (ii) transferable designs, and (iii) increased dissemination across multiple levels of engagement platforms to engage both practitioner and academic audiences.

References

Braga, A. (2016). 'The Value of 'Pracademics' in Enhancing Crime Analysis in Police Departments.' Policing: A Journal of Policy and Practice, 10(3): 308–314.

Brown, J. (2019). 'Evidence-Based Policing: Competing or Complementary Models.' In Fielding, N., Bullock, K. and Holdaway, S. (eds), Critical Reflections on Evidence-Based Policing. UK: Routledge, pp. 95-114.

Brown, J. et al. (2018). 'Extending the Remit of Evidence-Based Policing.' International Journal of Police Science and Management, 20(1): 38-51.

Brown, J., McDowall, A., Gamblin, D. and Fenn, L. (2020). 'Assessing Transmission and Translation of Learning about Evidence-Based Policing by Graduate Trainee Police Officers'. Policing: A Journal of Policy and Practice, 14 (1): 119–134. DOI: 10.1093/police/pay072

Bueermann, J. (2012). 'Being Smart on Crime with Evidence-Based Policing.' NIJ Journal, 269: 12-15.

Bullock, K. and Tilley, N. (2009). 'Evidence-Based Policing and Crime Reduction.' Policing: A Journal of Policy and Practice, 3(4): 381-387.

Canter, D. V. (2004). 'A Tale of Two Cultures: A Comparison of the Cultures of the Police and Academia.' In Villiers, P. and Adlam, R. (eds.), Policing a Safe, Just and Tolerant Socitey: An International Model. Winchester: Waterside Press, pp. 109-121.

Cesario, S. (2002). 'Evaluating the Level of Evidence of Qualitative Research.' Principles and Practice, 31(6): 708-714.

Chalmers, I. (2003). 'Trying to Do More Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, and Up-To-Date Evaluations.' Annals of the American Academy of Political and Social Science, 589: 22–40.

Fleming, J. and Wingrove, J. (2017). ''We would if we could … but not sure if we can': Implementing Evidence-Based Practice.' Policing: A Journal of Policy and Practice, 11(2): 202-213.

Goode, J. and Lumsden, K. (2018). 'The McDonaldisation of Police-Academic Partnerships: Organisational and Cultural Barriers Encountered in Moving from Research on Police to Research with Police.' Policing and Society, 28(1): 75-89.

Cabinet Office (2019). 'What Works Network'. Available at https://www.gov.uk/guidance/what-works-network (accessed 20 June 2019).

Cherney, A., Antrobus, E., Bennett, S., Murphy, B. and Newman, M. (2018). 'Evidence Based Policing: Queensland Police.' Technical Report. Available at https://espace.library.uq.edu.au/view/UQ:716040 (accessed 1 May 2020).

Heaton, R. and Tong, S. (2015). 'Evidence-Based Policing: From Effectiveness to Cost-Effectiveness.' Policing, 10(1): 60-70.

Home Office (2017). 'Police Workforce, England and Wales: 31 March 2017'. Available at: https://www.gov.uk/government/statistics/police-workforce-england-and-wales-31-march-2017 (accessed 20 June 2019).

Honess, R. (2018). 'Evidence-Based Policing: Using Evidence and Research to Improve Policing Practice.' Ethical Record, 123(2): 20-22.

Hough, M. (2010). 'Gold Standard or Fool's Gold? The Pursuit of Certainty in Experimental Criminology.' Criminology and Criminal Justice, 1: 11-22.

Howitt, D. (2012). 'Using Qualitative Methods to Research Offenders and Forensic Patients'. In Sheldon, K., Davies, J. and Howells, K. (eds), Research in Practice for Forensic Professionals. London: Routledge, pp. 132-158. DOI: 10.4324/9780203805329

Huey, L and Mitchell, R.J (2016). 'Unearthing Hidden Keys: Why Pracademics are an Invaluable (if Under Utilized) Resource in Policing Research.' Policing, 10(3): 300-307.

Huey, L. and Ricciardelli, R.J. (2016). 'From Seeds to Orchards: Using Evidence-Based Policing to Address Canada's Policing Research Needs.' Canadian Journal of Criminology and Criminal Justice, 58(1): 119-132.

Huey, L. et al. (2018). ''If you're gonna make a decision, you should understand the rationale': Are Police Leadership Programs Preparing Canadian Police Leaders for Evidence-Based Policing?' Policing: A Journal of Policy and Practice, 1-11. DOI: 10.1093/police/pay086

Kirby, S. (2013). Effective Policing? Implementation in Theory and Practice. Basingstoke: Palgrave Macmillian.

Kitto, S.C., Chesters, J. and Grbich, C. (2008). 'Quality in Qualitative Research.' Medical Journal of Australia,188(4): 243-246.

Klein, G. and Jarosz, A. (2011). 'A Naturalistic Study of Insight.' Journal of Cognitive Engineering and Decision Making,5(4): 335-351.

Knutsson, J. and Tompson, L. (2017). Advances in Evidence-Based Policing. Taylor and Francis, London.

Koehle, G., Six, T. and Hanrahan, K. (2010). 'Citizen Concerns and Approval of Police Performance.' Professional Issues in Criminal Justice, 5: 9-24.

Lum, C. (2009). 'Translating Police Research into Practice.' Ideas in American Policing Lecture Series. Washington, DC: Police Foundation.

Lum, C., Koper, C. and Telep, C. (2011). 'The Evidence-Based Policing Matrix.' Journal of Experimental Criminology, 7: 3-26.

Lum, C. and Koper, C. (2015). 'Evidence-Based Policing.' In Dunham, R. and Alpert, G. (eds), Critical Issues in Policing. Longrove, IL: Waveland Press, pp. 260-274.

Lum, C., and Koper, C. S. (2017). Evidence-Based Policing: Translating Research into Practice. Oxford: Clarendon.

Lum, C., Telep, C., Koper, C. and Grieco, J. (2012). 'Receptivity to Research in Policing.' Justice Research and Policy, 14(1): 61-95.

Lumsden, K. (2016). 'Police Officer and Civilian Staff Receptivity to Research and Evidence-Based Policing in the UK: Providing a Contextual Understanding Through Qualitative Interviews.' Policing, 11(2): 157-167.

Lumsden, K. and Goode, J. (2018). 'Policing Research and the Rise of the 'Evidence-Base': Police Officer and Staff Understandings of Research, tts Implementation and 'What Works'.' Sociology, 52(4): 813-829.

Palmer, I., Kirby, S. and Phythian, R. (2019). 'Assessing the Appetite for Evidence Based Policing: A UK Based Study.' International Journal of Police Science and Management, 21(2): 91-100.

Schulenberg, J.L. (2007). 'Analysing Police Decision-Making: Assessing the Application of a Mixed-Method/Mixed-Model Research Design.' International Journal of Social Research Methodology, 10(2): 99-119.

Sheldon, K., Davies, J. and Howells, K. (2012). Research in Practice for Forensic Professionals. London: Routledge. DOI: 10.4324/9780203805329

Sherman, L. (1998). 'Evidence-Based Policing.' Ideas in American Policing Lecture Series. Washington, DC: Police Foundation.

Sherman, L. (2013). 'The Rise of Evidence-Based Policing: Targeting, Testing, and Tracking.' In Tonry, M. (ed), Crime and Justice: A Review of Research, Chicago, IL: University of Chicago Press, pp. 377–451.

Sherman, L. (2015). 'A Tipping Point for 'Totally Evidenced Policing': Ten Ideas for Building an Evidence-Based Police Agency.' International Criminal Justice Review, 25(1): 11-29.

Sherman, L. et al. (2002). Evidence Based Crime Prevention. New York: Routledge.

Sherman, L. and Murray, A. (2015). Evidence-Based Policing. Los Angeles, CA: SAGE Publications.

The Scottish Institute for Policing Research (2017). 'Policing 2026 Evidence Review.' Available at: http://www.sipr.ac.uk/Plugin/Publications/assets/files/Policing_2026_Evidence_Review.pdf (accessed 20 June 2019).

Snilstveit, B. et al. (2016). 'Evidence and Gap Maps: A Tool for Promoting Evidence Informed Policy and Strategic Research Agendas.' Journal of Clinical Epidemiology, 79: 120-129.

Telep, C.W. and Lum, C. (2014). 'The Receptivity of Officers to Empirical Research and Evidence Based Policing: An Examination of Survey Data from Three Agencies.' Police Quarterly, 1-27.

Telep, C.W. and Winegar, S. (2016). 'Police Executive Receptivity to Research: A Survey of Chiefs and Sheriffs in Oregon.' Policing: A Journal of Policy and Practice, 10(3): 241-249.

Tilley, N. and Laycock, G. (2017). 'The Why, What, When and How of Evidence Based Policing.' In Knutsson, J. and Tompson, L. (eds), Advances in Evidence-Based Policing. Abingdon, UK: Routledge, pp. 10–26.

Weisburd, D. and Neyroud, P. (2011). 'Police Science: Toward a New Paradigm.' New Perspectives in Policing. Cambridge, MA: Kennedy School of Government.

Willis, J.J. and Mastrofski, S.D. (2014). 'Pulling Together: Integrating Craft and Science.' Policing, 8(4): 321-329.

Wood, J., Fleming, J. and Marks, M. (2008). 'Building the Capacity of Police Change Agents: The Nexus Policing Project.' Policing and Society, 18(1): 72-87.

Figure 1: Policing Research Evaluation Framework (PREF)

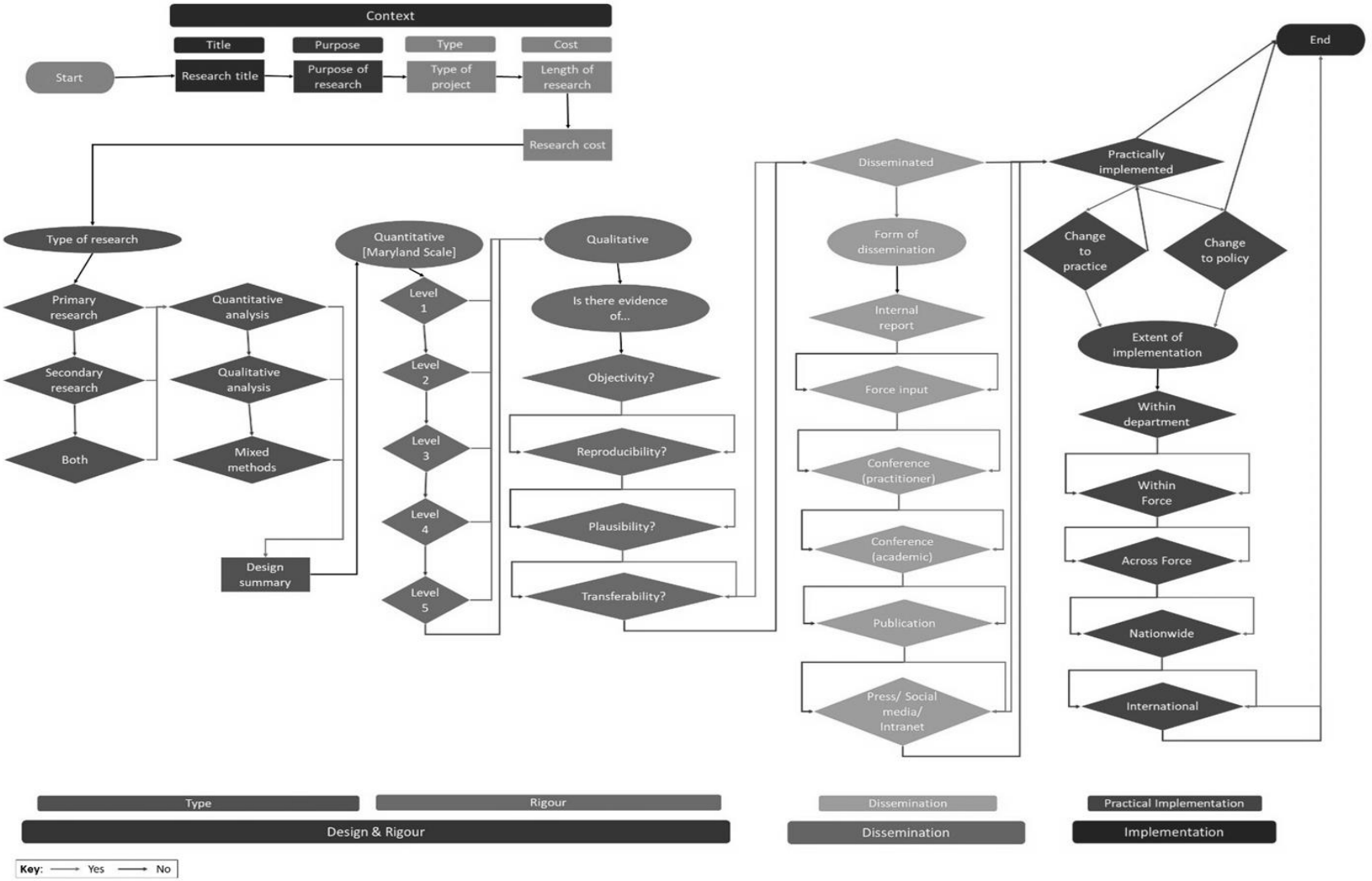

Table 1. Descriptive overview

	n	*%*		*n*	*%*
Design and rigour:			**Dissemination:**		
Data type:			Internal report	27	93
Secondary	16	55	Force/stakeholder input	16	55
Primary	7	24	Practitioner conference	13	45
Both	6	21	Academic conference	7	24
Analysis approach:			Press/social media/intranet	6	21
Quantitative	13	45	Publication	4	14
Mixed	10	34	**Implementation:**		
Qualitative	6	21	Findings implemented	15	52
Maryland scale:			*Types:*		
Level 1	18	78	Change to practice	14	93
Level 2	4	17	Within department	14	93
Level 3	1	4	Within force	14	93
Level 4	0	0	Across forces	4	27
Level 5	0	0	Change to policy	2	13
Research validity:			Nationwide	1	7
Plausibility	29	100	Internationally	0	0
Reproducibility	18	62	*Implementation failure reason:*		
Transferability	15	52	Implementation	11	79
Objectivity	14	48	Measurement	3	21

Table 2. Significant 2x2 Chi-Square Associations: Data type

	Secondary data	χ^2, OR	Both	χ^2, OR
Has objectivity	85.7%	10.208**, 16.50		
Has reproducibility	77.8%	9.805**$^+$, 15.75		
Has transferability	33.3%	5.992*$^+$, 0.136	40.0%	7.061*$^+$, -
Was implemented	33.3%	5.992*, 0.136	40.0%	7.061*$^+$, -
Changed practice	35.7%	4.144*, 0.202		
Implemented within department	35.7%	4.144*, 0.202		
Implemented within force	35.7%	4.144*, 0.202		
Disseminated at an academic conference	14.3%	6.237*$^+$, 0.078		

** p<.05, ** p<.01, *** p<.001, +Fisher's exact, - OR not computed*

Table 3. Significant 2x2 Chi-Square Associations: Analysis approach

	Mixed	χ^2, OR	Qualitative	χ^2, OR	Quantitative	χ^2, OR
Has objectivity	-	-	0.0%	7.061*$^+$, -	78.6%	12.461***, 23.83
Has reproducibility					66.7%	9.151**$^+$, 20.00
Has transferability	66.7%	14.246***$^+$, -			20.0%	7.744**, 0.10
Was disseminated			14.8%	8.235*$^+$, -		
Disseminated at a practitioner conference	61.5%	7.635*$^+$, 11.20			23.1%	4.507*, 0.18
Disseminated at an academic conference	71.4%	5.575*$^+$, 8.50			0.0%	7.497**$^+$, -
Was implemented	66.7%	14.246***$^+$, -			20.0%	7.744**, 0.10
Changed practice	64.3%	10.641**$^+$, 25.20			21.4%	5.992**, 0.136
Implemented within department	64.3%	10.641**$^+$, 25.20			21.4%	5.992**, 0.136
Implemented within force	64.3%	10.641**$^+$, 25.20			21.4%	5.992**, 0.136

** p<.05, ** p<.01, *** p<.001, +Fisher's exact, - OR not computed*

Table 4. Significant 2x2 Chi-Square Associations: Dissemination

	Force/ Stakeholder input	χ^2, OR	Practitioner conference	χ^2, OR	Publication	χ^2, OR	Press/ Intranet	χ^2, OR
Has transferability	73.3%	4.144*, 4.95	86.7%	21.992***, -			40.0%	7.061*$^{+}$, -
Implemented	73.3%	4.144*, 4.95	86.7%	21.992***, -			40.0%	7.061*$^{+}$, -
Implementation failure	18.2%	6.873*$^{+}$, -						
Approach amended							0.0%	6.667*$^{+}$, -
Changed practice	78.6%	5.992*, 7.333	92.9%	25.246***, -	28.6%	4.971*$^{+}$, -	42.9%	8.106**, -
Implemented within dept.	78.6%	5.992*, 7.333	92.9%	25.246***, -	28.6%	4.971*$^{+}$, -	42.9%	8.106**, -
Implemented within force	78.6%	5.992*, 7.333	92.9%	25.246***, -	28.6%	4.971*$^{+}$, -	42.9%	8.106**, -

** p<.05, ** p<.01, *** p<.001, +Fisher's exact, - OR not computed*

Table 5. Implementation themes and examples

Theme	Frequency	Example
Amendment to a current approach/initiative	6 (40%)	A mental health triage initiative was discontinued on the basis of an evaluation included within this sample.
Feeding forward to further research	5 (33%)	The findings from an MSc dissertation project generated calls for five other research projects which, combined, informed the development of the Force's current missing from home response.
Initiation of fuller review process	2 (13%)	Findings from a research report triggered Force to review support mechanisms for new recruits.
Offered support/initiated expansion of an effective current practice/initiative	2 (13%)	Results around the staff's view of code of ethics was highly supportive of current practice.
Tool or product development	2 (13%)	Results from a PhD thesis has since been used to develop a domestic abuse victim engagement app, which is currently being tested within a Randomised Control Trial based study.
Training development	2 (13%)	Based on the findings of an MSc dissertation project, a training package for responding to mental health related incidents has been delivered to all management staff, tactical firearms commanders and strategic firearms commanders (n > 400 staff at the time of this analysis).

Table 6. Significant 2x2 Chi-Square Associations: Implementation

Implemented		**χ^2, OR**	**Approach amended**	**χ^2, OR**	**Changed practice**	**χ^2, OR**	**Within dept.**	**χ^2, OR**	**Within Force**	**χ^2, OR**
Has objectivity	28.6%	5.81*, 0.15			28.6%	4.21*, 0.20	28.6%	4.21*, 0.20	28.6%	4.21*, 0.20
Has transferability	100.0%	29.00***, -			93.3%	25.26***, -	93.3%	25.26***, -	93.3%	25.26***, -
Has reproducibility			0.0%	8.75**$^{+}$, -						

** p<.05, ** p<.01, *** p<.001, +Fisher's exact, - OR not computed*