



College of
Policing

college.police.uk

Domestic Abuse Matters 2.0

Evaluation of first responder training

Iain Brennan (University of Hull)

Andy Myhill (College of Policing)

© College of Policing Limited (2017). 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, or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

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

This publication is available at: <http://whatworks.college.police.uk/Research/Pages/Published.aspx>.

The College of Policing will provide fair access to all readers and, to support this commitment, this document can be provided in alternative formats.

Any enquiries regarding this publication, including requests for an alternative format, should be sent to us at: contactus@college.pnn.police.uk.

Executive summary

Key findings and implications

- Findings from an evaluation of the first iteration of the Domestic Abuse Matters training showed positive impacts on some indicators of officer knowledge and understanding of coercive control, but not on others. The training was revised to be more interactive and to include established models and powerful bodyworn camera footage to help explain the dynamics and impact of coercive control.
- There was a small positive impact of the second iteration of the training on measures of police officer knowledge of coercive control and attitudes to domestic abuse.
- Positive impacts of the training on officer knowledge and attitudes were observed in both research sites – Humberside and Suffolk Police Forces – and effects were consistent across forces.
- Respondent attitudes were already in some instances positive prior to the training. The most likely explanation for the statistically small improvements in attitudes is a lack of sensitivity in the measurement tool combined with ‘ceiling’ or ‘floor’ effects (extreme baseline scores that left little scope for positive change). Further evaluation work could develop and test a more reliable measure for officer knowledge of coercive control.
- The consistency in baseline scores and pre-post training changes that were observed across the two forces suggest that this training will be relevant for other forces and may yield similar positive outcomes.

Introduction

In 2016, the College of Policing published an evaluation of training for first responding officers that focused on coercive control and responding to domestic abuse (Wire and Myhill, 2016). This training formed part of a wider programme: Domestic Abuse (DA) Matters. The evaluation found a positive impact on some indicators of officer knowledge and understanding of coercive control, but not for others; there was no impact on officer attitudes to domestic abuse. In response to the findings of the evaluation, aspects of the training content and delivery methods were revised for further piloting and evaluation.

The intervention

Revised training was delivered to first responding officers in two police forces. Key changes to the training were the inclusion of footage from bodyworn video, depicting officers responding to an incident involving a coercive and controlling perpetrator, and the restructuring of material on the dynamics of coercive control around the established and interactive ‘power and control wheel’.

Research methods

The evaluation employed a pre-post design with respondents completing the evaluation

materials immediately before and after the training. Respondents' pre- and post-surveys were matched using a unique identification number.

Change in individual survey items was tested using a series of paired t-tests. Analysis of data revealed similar baseline scores, but there were some differences in the nature and size of effects between forces.

Findings

The training was associated with a positive effect on officer understanding of coercive control and underlying attitudes to domestic abuse. Changes in mean scores, although often statistically significant, were generally small in magnitude. Improvements were, however, consistent across forces.

Measure	Impact
General attitudes towards domestic abuse	Small positive
Relevance of physical and non-physical risk factors	Very small positive
Attitudes towards victims of domestic abuse	Small positive

Conclusions and implications

The classroom training element of the DA Matters programme achieved consistent, positive improvements in frontline officer understanding of coercive control and attitudes towards domestic abuse, although effect sizes were small.

As the training material was revised in line with findings from a previous evaluation, and the delivery of the material by trainers more tightly controlled, the most likely explanation for the statistically small improvements in knowledge and understanding is a lack of sensitivity in the evaluation measurement tool. In addition, baseline scores that were close to the most desirable follow-up score probably created in some instances 'ceiling' and 'floor' effects which limited the extent to which scores could improve (ceiling or floor effects occur when the baseline score for an item is extremely low or extremely high, respectively, which limits the potential decrease or increase in average score at follow-up).

- Training for first responding officers can lead to consistent improvements in their understanding of coercive control, understanding of victim decision-making, and awareness of non-physical risk factors for future victimisation.
- Further work would be required to produce a standardised measurement tool sensitive enough to detect greater improvements in practitioner understanding of coercive control and their attitudes towards domestic abuse.
- Measures of attitudes towards domestic abuse should be complemented by measures of behavioural change, such as increases in the use of police powers on coercive control and increases in references to coercive and controlling behaviours on risk assessments, relative to forces that have not received this training.

Contents

	Page
Executive Summary	i
Acknowledgements	iv
1. Introduction	1
2. The intervention	2
3. Research methods	4
4. Findings	7
5. Conclusions and implications	14
Appendix 1. Outline of the training	16
Appendix 2. Evaluation questionnaire	17
References	21

Acknowledgements

Our thanks go to Colin Connolly and Ashley Allen at Humberside Police for coordinating the training and data entry, and to Millie Garrod and the Improvement and Evaluation Team at Suffolk Constabulary for collating and providing data.

We would also like to thank Dr Ben Bradford (University of Oxford) for an extremely helpful peer review.

1. Introduction

In 2016, the College of Policing published an evaluation of training for first responding officers that focused on coercive control and responding to domestic abuse (Wire and Myhill, 2016). The Domestic Abuse (DA) Matters programme was designed in partnership by SafeLives and the College in response to a review of training on domestic abuse undertaken by SafeLives (Morgan, 2015). This review was commissioned by the College in response to a recommendation by Her Majesty's Inspectorate of Constabulary (HMIC) following their 2014 thematic inspection of forces.

The first iteration of DA Matters was piloted in Hertfordshire Constabulary in 2015. The programme incorporated classroom training for first responders and supervisors, a peer coaching element, and a force 'health check'. A full description of the programme can be found in Wire and Myhill (2016). The College evaluated the training for first responders.

Key findings and implications from the evaluation of the first responder training were:

- Classroom-based training for first responders had positive effects for some indicators of knowledge and understanding of coercive control (immediately following the training), but no effect for others.
- There was no effect on officers' wider attitudes to domestic abuse.
- Future iterations of the training should include more interactive and self-reflective learning that simulates practice in responding to cases of domestic abuse.
- Future implementation should try to ensure that the training material is delivered consistently, to time, and in accordance with the learning objectives.
- Future evaluation of the training should have a pre- and post-test in order to establish baseline levels of knowledge and attitudes.

In response to the findings of the evaluation, the College initiated a revision of the training for first responders, with a view to further testing and evaluation of the programme.

Structure of the report

Chapter 2 details the changes to the training that were made for the second iteration of the programme, and describes implementation of the revised training in two police forces. Chapter 3 details the research methods used for the evaluation of the second iteration of the training, and Chapter 4 presents the findings. Chapter 5 presents conclusions and implications.

2. The intervention

The College contracted an associate, a former police trainer, to work alongside SafeLives and College staff to revise the first responder training both from a content and delivery perspective. Incorporating additional material on coercive control suggested by Women's Aid, and bodyworn camera footage used only in the training for coaches and supervisors¹ in the first iteration, the associate's brief was to make the delivery of the content more interactive.

The following key changes were made to the first responder training:

- Addition of material describing the typical 'stages' of coercive control and the impact on the victim. This material was adapted from a model of police intervention proposed by Kelly (1999).
- Material on the dynamics of coercive control strengthened and made more interactive by debriefing the 'power and control wheel' (<http://www.theduluthmodel.org/pdf/powerandcontrol.pdf>).
- Footage of an actual domestic abuse incident moved to later in the training in order to consolidate learning around behaviours associated with coercive control.
- Interactive debrief and group exercises relating to bodyworn camera footage to consolidate learning on perpetrators' tactics of manipulation and victims' minimisation of abuse.
- Officers encouraged to reflect on the learning by considering a personal development plan for their response to domestic abuse.

A full outline of the one-day classroom training session is presented in Appendix 1.

Implementation

The second iteration of the training was implemented in two police forces.

Humberside

Training sessions involved up to 25 first responders. The training was delivered by a consortium of trainers, coordinated by the College. Some trainers were part of the team which delivered the first iteration of the training; others were new trainers. A 'train the trainer' event was held prior to implementation.

Training took place at eight sites simultaneously over five days (at weekly intervals) between 23rd February and 22nd March 2016. Approximately 850 first responders completed the training.

¹ For the second iteration of the DA Matters programme, coaches and supervisors receive the same training input and are known as 'champions'.

Suffolk

Thirty-four training sessions were conducted between 4th May and 27th May 2016 and consisted of up to 25 first responders. Approximately 800 first responders completed the training. Of those 800, 25% were selected by trainers and by the constabulary for further training for the role of champions.

Training sessions were delivered by experienced SafeLives trainers alongside local trainers, from both police and specialist support sector backgrounds, selected and subject to a two-day pass or fail 'train the trainer' event. These trainers continue to deliver DA Matters in Suffolk under licence from the College to any remaining, untrained first responders and to newly identified champions and new recruits.

3. Research methods

The evaluation employed a pre-post design with respondents completing the evaluation materials immediately before the training ('baseline') and immediately following the training ('follow-up').

Development of the evaluation materials

As the training material was revised and implemented over a short timescale, it was not possible to test and refine a questionnaire measurement tool before the evaluation. Specifically, items and measures were developed by the evaluation team based on an evaluation of an earlier iteration of the training programme and the team's reading of the aims of the revised training. The evaluation materials were revised following implementation and prior to the analysis stage, in collaboration with the trainers (discussed in further detail below).

The initial evaluation tool contained four independent sections.

- Section A required respondents to indicate the extent to which they agreed with sixteen items relating to domestic abuse incidents and police responses to domestic abuse. Respondents indicated their agreement using a five-point scale ('Strongly disagree' to 'Strongly agree').
- Section B required respondents to indicate how relevant each of nine pieces of information about a domestic abuse incident would be in predicting future harm ('Definitely not relevant' to 'Definitely relevant'). Some of these items related to physical domestic abuse and the presence of other criminal offences, and some related to non-physical abuse. The goal of the training was to increase the perceived importance of non-physical factors in determining risk. Therefore, no change would be anticipated in the rating of physical abuse or other risk factor items, but increased importance of non-physical factors was desirable.
- Section C used the same items as Section B, but asked respondents to rank these nine pieces of information in order of their value in predicting future harm to a victim. This exercise was an alternative approach to testing the perceived value of non-physical risk factors.
- Section D required respondents to indicate the extent to which they agreed with ten reasons that might explain why victims do not leave abusive partners, in order to test further their understanding of the dynamics of coercive control.

Refining the data collection tool

The items were developed with a view to their being able to test for a desirable change in respondent knowledge and attitudes. However, given the complexity and subjectivity involved in responding to domestic abuse cases, some responses can be interpreted as either desirable or undesirable depending on context. This ambiguity can damage the validity of the item as an indicator of change. An example of this is: 'Domestic abuse is a high priority for the police'. This item could be interpreted by the respondent – as a representative of the police force – as reflecting their personal attitudes, but it could also be interpreted in the wider context of the organisation and its priorities. A respondent's endorsement of the item, 'The

victim is willing to make a statement' as an indicator of risk of future harm could be interpreted as a sign that the severity of the incident is so great that it has overcome any reluctance on the part of the victim to report the incident. Equally, however, not endorsing this item could also reflect a belief that the victim is being controlled by the perpetrator – another form of harm.

Given the potential for equally correct but conflicting interpretations of the same item, the research team sought the views of the programme trainers (who were subject matter experts from both police and voluntary/charity sector agencies) to identify those items that had the least potential for conflicting opinions. Using an online survey tool, trainers were asked to indicate the direction of change for each of the 35 questionnaire items that they would expect if the training achieved its aims. The three options were: 'increase', 'decrease' and 'no change'. Seven trainers completed the task. Using 70% agreement (i.e. 70% of trainers selecting the same response) as a threshold for inclusion in the final analysis, twenty items were retained – seven items from Section A, five items from Section B, five items from Section C and eight items from Section D. Item selection was undertaken before items were analysed to avoid the possibility of selection bias.

The items of the questionnaire were further reduced at the point of data analysis. Although the questionnaire instructions for Section C stipulated that respondents should rank the items in descending order from 1 to 9 in order of perceived value and should only use one '1', one '2' and so on, over 40% of respondents at baseline measurement did not follow these instructions correctly (they 'ranked' two or more items as, for example, priority 1). Therefore, Section C was excluded from the analyses.

To avoid confusion, the Sections described above (A, B, D) were renamed Sections 1, 2 and 3, respectively.

Data collection

Participants completed the evaluation document immediately before and immediately after the training event (within a few minutes).

Trainees entering the training room were asked to select an evaluation pack from the front of the room. These packs contained three documents: an evaluation cover sheet, Booklet One and Booklet Two (see Appendix 2). The booklets were identical except for their front page titles – 'Booklet One' and 'Booklet Two'. On the front of the booklets was a space for the respondents to indicate their unique code number, which was affixed to the front of the evaluation pack envelope. The unique codes were randomly-generated four digit numbers. These code numbers were used to maintain the anonymity of the respondents, while enabling the matching of baseline and follow-up questionnaires. No individually identifiable information was requested from the respondents. Furthermore, no demographic, organisational rank or role information was collected, and completed questionnaires were not grouped or analysed according to date or location of training.

Prior to the commencement of the training, respondents were asked to write the code number on Booklet One, to complete the questionnaire and to pass the completed questionnaire to the front of the room. At the end of the training respondents were asked to complete the follow-up questionnaire (Booklet Two), entering the same unique code number on the front of the booklet, and to then return the completed booklet to the trainer.

Data analysis

For all sections, the overarching aim of the analysis was to detect if, and to what extent, average response scores on items changed in the anticipated direction. Firstly, descriptive statistics were generated to demonstrate the pre- and post-intervention mean scores.

Change in individual items was tested using paired t-tests, which determine whether a change from baseline to follow-up is statistically significant. The study had a relatively large sample. Very large samples increase the risk of false positive results, making statistical significance a problematic indicator of effect. Consequently, Cohen's d was calculated from the t-test statistic and was employed as the preferred indicator of effect size.

Box 1: Cohen's d

Cohen's d is a commonly used effect size indicator. It is calculated by dividing the difference in average scores (item post-training average – item pre-training average) by the standard deviation of the pre-training item score (Cumming, 2013). Interpreting the Cohen's d value can require some statistical knowledge, but a crude rule-of-thumb is that a d value of 0.2 equates to a 'small' change, a value of 0.5 equates to a 'medium' change and a value of 0.8 equates to a 'large' change (Cohen, 1969). These thresholds are dependent on what is typical for the field and the context of measurement. For example, the transparent nature of this evaluation and the recency of the training would suggest effect sizes ought to be towards the higher end of the scale. Cohen used an analogy of the difference in average heights between groups of girls. A d value of 0.2 ('small') would be akin to the difference in average height between two groups of 14- and 15-year old girls; a d value of 0.5 ('medium') would be akin to the difference between the heights of 14- and 18-year old girls; and a d value of 0.8 ('large') would be akin to the difference between the heights of 13- and 18-year old girls.

The study involved a large number of statistical tests: 44 in total. As statistical significance is prone to false positives² when a large number of tests are performed, it is prudent to adjust the threshold for statistical significance using a correction known as the Holm correction (Holm, 1979) to increase our confidence that an observed result reflects a real change rather than a statistical anomaly.

Force-level differences have been observed in other police training (Miller and Alexandrou, 2016) and the delivery of the DA Matters training differed slightly between the two forces. Therefore, it was considered prudent to treat the forces as having differences and to analyse their data separately.

² A false positive is when an outcome being tested for is found mistakenly to have occurred. When a large number of statistical tests are performed, it increases the likelihood that a statistically significant result will occur by chance.

4. Findings

- A positive impact of the training on officer knowledge and attitudes was observed in both research sites.
- Baseline scores for the items were sometimes close to the maximum possible score, indicating limited potential or need for improvement in respondent attitudes pre-training.
- There were consistent improvements in general attitudes towards domestic abuse and attitudes towards victims of domestic abuse among the sample of first responders. The perceived relevance of physical and non-physical risk factors was less affected and in several of these items, no statistically significant change was observed.
- Although statistically significant, the magnitude of changes was, generally, small.

Achieved sample

The evaluation materials were completed by 1,551 first responders (802 in Humberside and 749 in Suffolk) although not all questionnaires were completed fully or correctly. As the analysis was a repeated-measures design, in cases where a respondent did not complete the two corresponding items (e.g. Item 1 of Section 1) in both the baseline and follow-up questionnaires, that item was excluded for that individual. This process is known as pairwise deletion. Following pairwise deletion, averages of 762 and 722 respondents were available for Humberside and Suffolk, respectively.

The descriptive statistics for each item, by force, are described in Tables 1 and 2. Values represent the minimum and maximum scores for each item (based on responses), the mean score, and the standard deviation (SD).

Table 3 reports the results of paired t-tests comparing differences in mean scores for pre- and post-data, by force.

Table 1 (Humberside)	Anticipated change	Pre			Post			
Section 1: Extent to which agree/disagree with each statement		Min/Max	Mean	SD		Min/Max	Mean	SD
1. A lot of domestic abuse incidents are just petty verbal arguments	Decrease	1, 5	2.90	1.01		1, 5	2.24	0.95
2. Victims of domestic abuse should just leave	Decrease	1, 5	2.61	0.90		1, 5	2.43	0.91
3. There are a lot of false reports of domestic abuse	Decrease	1, 5	2.80	0.86		1, 5	2.48	0.79
4. When speaking to police, many victims downplay the seriousness of domestic abuse incidents	Increase	1, 5	3.66	0.79		1, 5	3.86	0.85
5. It's hard to see why victims of domestic abuse don't just leave	Decrease	1, 5	2.79	0.95		1, 5	2.42	0.96
6. A victim of domestic abuse who won't make a statement to the police must have a good reason	Increase	1, 5	3.23	0.87		1, 5	3.60	0.93
7. There's not much point in completing a risk assessment after a domestic abuse incident if there has not been a criminal offence	Decrease	1, 5	1.91	0.72		1, 5	1.74	0.79
Section 2: Relevance of risk factors (Definitely not relevant to definitely relevant)								
8. There has previously been a series of 'verbal only' incidents	Increase	1, 4	3.78	0.47		1, 4	3.70	0.60
9. The victim does not want to engage with police	Increase	2, 4	3.37	0.54		1, 4	3.45	0.58
10. The victim seems genuinely frightened	Increase	1, 4	3.26	0.63		1, 4	3.33	0.62
11. The victim reports that the abuser tracks their phone	Increase	1, 4	3.38	0.77		1, 4	3.37	0.77
12. The victim has no friends to stay with	Increase	1, 4	3.73	0.51		1, 4	3.79	0.46
13. There has been physical violence	No change	1, 4	3.68	0.50		1, 4	3.69	0.51
14. A criminal offence has been committed	No change	1, 4	3.52	0.65		1, 4	3.54	0.63
Section 3: Extent to which agree/disagree with each statement								
15. People give as good as they get in domestic abuse	Decrease	1, 5	2.17	0.84		1, 5	2.02	0.77
16. They are financially dependent on the perpetrator	Increase	1, 5	3.99	0.74		1, 5	4.00	0.71
17. They would fear for their safety if they left	Increase	1, 5	4.02	0.76		1, 5	4.29	0.65
18. Victims and perpetrators are often 'as bad as each other'	Decrease	1, 5	2.48	0.84		1, 5	2.22	0.81
19. Victims feel they have nowhere else to go	Increase	1, 5	4.13	0.61		1, 5	4.24	0.55
20. The abuser is preventing them from leaving	Increase	1, 5	4.01	0.66		1, 5	4.22	0.57
21. Victims fear the unknown	Increase	1, 5	4.18	0.61		2, 5	4.30	0.56
22. Victims stay to protect children	Increase	1, 5	4.33	0.68		1, 5	4.36	0.57

Table 2 (Suffolk)	Anticipated change	Pre			Post			
Section 1: Extent to which agree/disagree with each statement		Min/Max	Mean	SD		Min/Max	Mean	SD
1. A lot of domestic abuse incidents are just petty verbal arguments	Decrease	1, 5	2.86	1.02		1, 5	2.38	0.98
2. Victims of domestic abuse should just leave	Decrease	1, 5	2.75	0.89		1, 5	2.54	0.92
3. There are a lot of false reports of domestic abuse	Decrease	1, 5	2.78	0.83		1, 5	2.50	0.82
4. When speaking to police, many victims downplay the seriousness of domestic abuse incidents	Increase	1, 5	3.82	0.72		1, 5	4.11	0.68
5. It's hard to see why victims of domestic abuse don't just leave	Decrease	1, 5	2.65	0.97		1, 5	2.39	0.99
6. A victim of domestic abuse who won't make a statement to the police must have a good reason	Increase	1, 5	3.19	0.91		1, 5	3.64	0.92
7. There's not much point in completing a risk assessment after a domestic abuse incident if there has not been a criminal offence	Decrease	1, 5	1.89	0.67		1, 5	1.67	0.65
Section 2: Relevance of risk factors (Definitely not relevant to definitely relevant)								
8. There has previously been a series of 'verbal only' incidents	Increase	1, 4	3.76	0.51		1, 4	3.77	0.49
9. The victim does not want to engage with police	Increase	1, 4	3.38	0.60		1, 4	3.48	0.55
10. The victim seems genuinely frightened	Increase	1, 4	3.23	0.63		1, 4	3.36	0.60
11. The victim reports that the abuser tracks their phone	Increase	1, 4	3.40	0.65		1, 4	3.47	0.62
12. The victim has no friends to stay with	Increase	1, 4	3.79	0.45		2, 4	3.82	0.39
13. There has been physical violence	No change	2, 4	3.73	0.46		2, 4	3.76	0.43
14. A criminal offence has been committed	No change	1, 4	3.49	0.61		1, 4	3.53	0.59
Section 3: Extent to which agree/disagree with each statement								
15. People give as good as they get in domestic abuse	Decrease	1, 5	2.00	0.73		1, 5	1.88	0.75
16. They are financially dependent on the perpetrator	Increase	1, 5	4.04	0.64		1, 5	4.11	0.62
17. They would fear for their safety if they left	Increase	1, 5	4.08	0.68		1, 5	4.41	0.56
18. Victims and perpetrators are often 'as bad as each other'	Decrease	1, 5	2.32	0.79		1, 5	2.07	0.80
19. Victims feel they have nowhere else to go	Increase	2, 5	4.16	0.53		1, 5	4.30	0.52
20. The abuser is preventing them from leaving	Increase	1, 5	4.00	0.64		2, 5	4.28	0.59
21. Victims fear the unknown	Increase	2, 5	4.13	0.57		2, 5	4.37	0.54
22. Victims stay to protect children	Increase	2, 5	4.34	0.58		2, 5	4.42	0.56

Table 3: Mean differences (post-pre); paired t-tests; Cohen's d	Humberside			Suffolk		
	Mean difference	t	d	Mean difference	t	d
Section 1: Extent to which agree with each statement						
1. A lot of domestic abuse incidents are just petty verbal arguments	-0.65	-18.24***	-0.65	-0.48	-12.35***	-0.46
2. Victims of domestic abuse should just leave	-0.17	-5.50***	-0.19	-0.21	-5.95***	-0.22
3. There are a lot of false reports of domestic abuse	-0.31	-10.31***	-0.37	-0.28	-7.93***	-0.29
4. When speaking to police, many victims downplay the seriousness of domestic abuse incidents	0.21	5.56***	0.19	0.29	10.4***	0.39
5. It's hard to see why victims of domestic abuse don't just leave	-0.37	-9.76***	-0.36	-0.26	-6.50***	-0.24
6. A victim of domestic abuse who won't make a statement to the police must have a good reason	0.37	9.44***	0.33	0.45	12.36***	0.46
7. There's not much point in completing a risk assessment after a domestic abuse incident if there has not been a criminal offence	-0.16	-4.62***	-0.17	-0.22	-7.47***	-0.28
Section 2: Relevance of risk factors						
8. There has previously been a series of 'verbal only' incidents	-0.09	-3.82**	-0.15	0.01	0.20	0.008
9. The victim does not want to engage with police	0.07	2.99	0.10	0.10	4.41***	0.16
10. The victim seems genuinely frightened	0.07	2.55	0.08	0.13	5.35***	0.20
11. The victim reports that the abuser tracks their phone	-0.01	-0.26	-0.02	0.07	2.64	0.10
12. The victim has no friends to stay with	0.06	2.78	0.08	0.03	2.14	0.08
13. There has been physical violence	0.01	0.71	0.007	0.03	1.22	0.05
14. A criminal offence has been committed	0.01	0.39	0.003	0.04	2.19	0.08
Section 3: Extent to which agree with each statement						
15. People give as good as they get in domestic abuse	-0.14	-4.33***	-0.16	-0.12	-3.57**	-0.13
16. They are financially dependent on the perpetrator	0.001	-0.04	0.01	0.07	2.06	0.08
17. They would fear for their safety if they left	0.26	8.38***	0.28	0.33	11.17***	0.42
18. Victims and perpetrators are often 'as bad as each other'	-0.25	-7.32***	-0.27	-0.25	-7.94***	-0.30
19. Victims feel they have nowhere else to go	0.11	4.50**	0.15	0.14	6.61***	0.25
20. The abuser is preventing them from leaving	0.21	8.08***	0.28	0.28	10.46***	0.40
21. Victims fear the unknown	0.12	4.84***	0.16	0.24	9.57***	0.36
22. Victims stay to protect children	0.03	0.91	0.02	0.08	3.24*	0.12

*Holm-adjusted $p < 0.05$, **Holm-adjusted $p < 0.01$, *** Holm-adjusted $p < 0.001$

Impact on underlying knowledge and attitudes

Paired t-tests identified statistically significant changes in the anticipated direction for most of the evaluation items. However, as noted, when sample sizes are large (as in this evaluation) tests of statistical significance are susceptible to inflated likelihood of false positive results. A more accurate reflection of the effect of the training is to interpret change in terms of a standardised 'effect size' (Cohen's *d*; see Box 1) that is independent of sample size.

As the baseline scores and mean differences described earlier in the section were consistent across the two forces, the results are presented together. In cases where results diverged, we note this and offer possible explanations for the divergence.

General attitudes towards domestic abuse

All items relating to general attitudes towards domestic abuse that were measured in Section 1 changed in the anticipated and desired direction. It is important to consider the generally small to medium effect sizes in this section and the mean scores for each item at baseline. The range of item scores was between 1 and 5 (the full range of possible scores), but baseline scores tended to be close to the centre. This indicates that small effect sizes were not due to ceiling or floor effects. Floor or ceiling effects occur when the average baseline score for an item is extremely low or extremely high, respectively, which limits the potential decrease or increase in average score at follow-up. In general, respondents did not have extremely desirable or extremely undesirable attitudes towards domestic abuse. As demonstrated by the pre- and post-training mean scores, the change from baseline to follow-up was relatively small, with absolute mean changes ranging between 0.16 and 0.65 of a point on the five-point scale.

Within the section, there was considerable variation in the extent of improvements in items. In particular, the training seems to have been effective in changing officer perceptions of domestic incidents as 'petty arguments'. The training was also associated with improved perceptions of victim's responses to abuse – particularly in the apparently contradictory behaviour that some victims display when the police attend incidents, such as withdrawing complaints of abuse and not leaving an abusive partner. A central goal of the training was to emphasise the conflicts and barriers faced by victims of domestic abuse, which manifests in apparently contradictory behaviour. The changes observed in respondent attitudes suggests that the training shifted attitudes towards a more understanding position.

Relevance of physical and non-physical risk factors

Only three of the five items relating to the relevance of physical and non-physical risk factors measured in Section 2 exhibited statistically significant changes in the anticipated direction³, and the pattern in these changes differed across forces. When items did change, the magnitude of change was small with effect sizes ranging between 0.003 and 0.20 and absolute mean differences ranging between 0.01 and 0.13. Importantly, it seems likely that this set of items was susceptible to ceiling effects. The mean baseline score for these seven items ranged from 3.23 to 3.78. Given that the maximum possible score for these items was 4, there was little scope for improvement as a result of training. Encouragingly, in Suffolk, all the items that the training explicitly sought to increase in relevance – those that reflected the

³ Two items, 'There has been physical violence' and 'A criminal offence has been committed' were not expected to change as they have been identified as important to officers (Robinson et al, 2016) and indeed may be indicative of increased threat of harm.

importance of non-physical abuse as a risk factor for future harm, and non-physical aspects of vulnerability, such as 'the victim has no friends to stay with' – improved despite having high baseline values (although these did not reach statistical significance). In Humberside, the item 'There has previously been a series of 'verbal only' incidents' yielded a statistically significant decrease in perceived relevance, which is the opposite to the desired direction of change. However, the baseline score for this item was 3.78 and the mean change in the two items that were not anticipated to change was 0.02, suggesting that a change of 0.08, as observed for this item, may reflect a small regression to the mean rather than any undesirable effect of the training.

Notably, most of the items that the training explicitly sought to increase in relevance – those that reflected the importance of non-physical abuse as a risk factor for future harm, and more subtle signs of coercive control, such as 'the victim has no friends to stay with' – improved despite having high baseline values. However, scores on the item 'the victim reports that the abuser tracks their phone' decreased in Humberside (although not to a statistically significant degree), suggesting that some more subtle signs of risk were not picked up by officers. Future iterations should consider elevating within the training the importance of these subtle indicators of risk.

Despite the generally small effect sizes achieved in this section, it is encouraging that high baseline scores were observed. This finding suggests that, prior to the training, officers regarded non-physical risk factors as important indicators of future risk. A new criminal offence of controlling and coercive behaviour in an intimate or family relationship was enacted in England and Wales in December 2015, which may have had some impact on officer perceptions of non-physical risk factors prior to the training. The College made a short e-learning course available to all forces at the time the offence was enacted. This e-learning was not compulsory, however, and feedback from Humberside and Suffolk suggested uptake was likely to have been low.

Attitudes towards victims of domestic abuse

Section 3 sought to measure changes in attitudes towards victims of domestic abuse, specifically to improve understanding of why victims may not leave abusive partners. A central theme of the training was to highlight the complexity of the decision faced by victims of domestic abuse and the multiple considerations that go into this process. Scores on 7 of the 8 items in this section showed statistically significant improvements following training. As with Section 2, it is important to note that 6 of the 8 items in Section 3 were within approximately one point of the minimum or maximum desirable score. These extreme baseline positions created the strong potential for floor or ceiling effects.

The perception that there is often no primary perpetrator in cases of domestic abuse shifted slightly in the desired direction; fewer respondents indicated support for the items 'People give as good as they get in domestic abuse' and 'Victims and perpetrators are often as bad as each other'. Baseline support for these items was less extreme than for the other items, suggesting a general view that both partners play some role in the occurrence of domestic abuse. The observed change, then, is indicative of a positive effect of the training in demonstrating that there is usually a dominant abuser.

Items relating to victims being psychologically and physically trapped in abusive relationships also received greater support following the training – illustrated through changes to the items 'Victims feel they have nowhere to go', 'The abuser is preventing them from leaving' and 'Victims stay to protect children' (although improvement on the latter was very small and not

statistically significant in Humberside). Respondents demonstrated increased awareness that, for victims, leaving an abusive partner may not protect them, at least initially, and in fact potentially increases the dangers they face, as demonstrated by the increased endorsement of the statement 'They would fear for their safety if they left'. The only item that did not change to a statistically significant degree was that victims do not leave because 'They are financially dependent on the perpetrator'. It is perhaps surprising that no change occurred because exercising financial control over a partner is a key tactic used by many perpetrators of coercive control. Future iterations of the training may wish to emphasise this aspect of abusive behaviour.

Comparison of effects across forces

The analysis did not initially set out to test for differences in the effectiveness of the training between forces. As the training programme was identical in the two forces, there was an expectation that any improvement in attitudes around domestic abuse would be largely consistent. This proved to be the case: very similar baseline and mean difference scores were observed across Humberside and Suffolk, which can be seen in Table 3.

This cross-force similarity in the pre-post change and in baseline scores relating to coercive control is valuable information. The similar baseline scores indicate that training developed with knowledge of a single force is likely to be reflective of the attitudes of other forces. While this conclusion is, admittedly, based on a single pair of forces, it is encouraging and suggests that the message of the training should have an impact on officers in other forces as the training is rolled out more widely. Further – again with the caveat that we are making inferences based on just two forces – the degree of change exhibited by the two forces suggests that the training will have similar effects on officers in other forces.

5. Conclusions and implications

The classroom training element of the DA Matters programme achieved small but consistent improvements in frontline officer knowledge of coercive control and attitudes towards domestic abuse. That improvements were not larger in statistical terms, especially in relation to knowledge transfer, was perhaps disappointing when considering measurement took place prior to and immediately after the training was delivered when the messages of the training would have been fresh in the mind of trainees. There are several possible explanations for these small observed effect sizes.

Training design

The content and delivery of the training was revised significantly from the first iteration of the programme. The second iteration contained more interactive elements, and anecdotal feedback from trainers was that the material flowed better and was easier to train. Though it is possible that the content of the training and the classroom mode of delivery may in part explain the small observed impacts, it is perhaps less likely, as the training was revised from the first iteration in line with evidence-based recommendations (see Wire and Myhill, 2016).

Training implementation

A key issue identified during the evaluation of the first iteration of the training was inconsistent delivery of the material (see Wire and Myhill, 2016). For the second iteration, training in one force area was delivered by a single provider who implemented stringent recruitment and monitoring processes for trainers. In the other site, consistent delivery of the course material was emphasised during a 'train the trainer' event to a pool of trainers comprising some of the trainers from the original pilot and some new recruits. The fact that similar results were observed in both sites suggests delivery of training is again less likely as an explanation for the small improvements in knowledge and attitudes.

Evaluation design

The evaluation measurement tool was also revised from the first iteration, to reflect both learning from the original evaluation and the changes to the content of the training material. Due to the timescales for revising and implementing the second iteration of the training, there was however no time to test and refine the measurement tool.

The slightly inconsistent feedback received from trainers concerning the 'direction' in which certain items should move was a cause for concern. If the trainers reflected disagreement relating to specific items, then it might be assumed that respondents also will have reflected some disagreement. This outcome may have reduced the precision of the measurement and diluted the observed impact of the training. The lack of consensus in the anticipated direction of change following training may reflect the multitude of responses that victims may have to coercive control. In turn, this complicates the generation of valid measures of attitude change. Future measures of training effectiveness should aim to identify items for which there is a higher level of consensus.

In addition, high baseline scores potentially created a 'ceiling effect' that limited the potential for large positive change. One interpretation is that existing knowledge and attitudes of coercive control were high in both forces, yet officer knowledge and attitudes were singled

out by HMIC (2014) as requiring improvement, and more recent research (Robinson et al, 2016) suggests this issues remains, despite the introduction of the coercive control law in late 2015. Another possibility is that it was easy for respondents to identify the 'desirable' response to certain items, and that high baseline scores reflected to some extent social desirability bias in responses. Once again, in the absence of comprehensive testing and refining of the evaluation measures it is not possible to rule out this possible explanation.

One issue, then, with the evaluation of the training component of the DA Matters programme is the lack of a standardised measurement tool for coercive control. Existing measurement tools, such as the Checklist of Controlling Behaviours (Lehmann et al, 2012) are relatively new and focused primarily on identifying the existence or prevalence of coercive control. Further work would be required to develop a measurement tool that can measure accurately practitioners' understanding of coercive control while accommodating the complexities of this behaviour and victims' responses to abuse.

Future research

Any future evaluation of the DA Matters programme could usefully include consideration of whether gains in knowledge are sustained weeks or months following the training. In addition, further evaluation could consider any impact of the wider programme, including the role of 'champions', on officers' wider attitudes to domestic abuse and behaviour in responding to calls for service.

Future research could also refine the measurement tool developed for this evaluation in order to make it more sensitive to detecting improvements in knowledge transfer and attitudinal change.

Key implications

- Training for first responding officers can lead to consistent improvements in their understanding of coercive control.
- Further work would be required to produce a standardised measurement tool sensitive enough to detect improvements in practitioners' understanding of coercive control and their attitudes towards domestic abuse.
- Measures of attitude towards domestic abuse could be complemented by measures of behavioural change, such as increases in the use of police powers on coercive control and increases in references to coercive tactics on risk assessments, relative to forces that have not received this training.
- The similarity in baseline scores and pre-post changes observed in this study offer some support for the idea that attitudes towards domestic abuse are consistent across forces and that future roll-out of this training in other forces can elicit consistent positive effects.

Appendix 1. Outline of the training

The training included the following sessions and elements:

Session 1: Course opening

- Introductions; aims; learning outcomes; context of 2014 HMIC inspection.
- Session asking attendees to discuss what makes their job difficult in relation to responding to domestic abuse. Themes such as 'uncooperative victims' developed to be picked up and challenged in subsequent sessions.

Session 2: Dynamics of domestic abuse

- PowerPoint input and whole class debrief of stages in coercive control, based on Kelly's (1999) model of police intervention, and the 'power and control wheel'.
- 'Leaving' video, followed by group exercise on the difficulties of leaving an abusive partner.
- Discussion of coercive control legislation and evidence gathering.

Session 3: Reality of domestic abuse and responding to it

- Exercise where victim statements are read out and learners identify questions that will help them to uncover coercive control.
- Viewing and debriefing of video footage of actual domestic abuse, captured unintentionally by the perpetrator.

Session 4: Responding to domestic abuse

- Interactive discussion and exercises to debrief real bodyworn video footage of police responding to a domestic-related incident.
- PowerPoint input and discussion of how to write up a report of a domestic-related incident, including use of language.
- Reflection on what a personal development plan might look like for individual officers.

Session 5 – Course closure

- Recap of aims; feedback and evaluation.

Appendix 2. Evaluation questionnaire

Booklet one

On the outside of your envelope, there is a label with a number between 0 and 9999. Please enter this number in the space below.

— — — —

Please turn over the page and complete the questionnaires that follow

Please read the following statements and indicate (✓) the extent to which you agree or disagree with each.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1 A lot of domestic abuse incidents are just petty verbal arguments					
2 Victims of domestic abuse should just leave					
3 There are a lot of false reports of domestic abuse					
4 Victims of domestic abuse should always make a statement about the incident					
5 In general, domestic abuse without violence is less harmful than domestic abuse with violence					
6 In policing, there is little reward in dealing with domestic abuse cases					
7 Alcohol is the cause of the majority of domestic abuse					
8 Securing a conviction for non-violent domestic abuse is likely to be very difficult					
9 When speaking to police, many victims downplay the seriousness of domestic abuse incidents					
10 It's hard to see why victims of domestic abuse don't just leave					

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
11 Victims' reports of domestic abuse are usually true					
12 A victim of domestic abuse who won't make a statement to the police must have a good reason					
13 There's not much point in completing a risk assessment after a domestic abuse incident if there has not been a criminal offence					
14 Managing domestic abuse is a high priority for the police					
15 The best way to reduce domestic abuse would be to ban alcohol					
16 Proving 'coercive control' in court will be very difficult					

At a domestic abuse incident, in your opinion, how relevant are the following pieces of information to predicting further domestic abuse harm?

	Definitely not relevant	Unlikely to be relevant	Probably relevant	Definitely relevant
A There has been physical violence				
B There has previously been a series of 'verbal only' incidents				
C The victim does not want to engage with police				
D The house is well-presented				
E A criminal offence has been committed				
F The victim seems genuinely frightened				
G The victim is willing to make a statement				
H The victim reports that the abuser tracks their phone				
I The victim has no friends to stay with				

At a domestic abuse incident, in your opinion, how strongly do you feel these items predict further harm to a victim of domestic abuse?

Please rank these items in **descending** order from 1 to 9 in terms of their ability to predict further harm (with **1** indicating the **strongest** predictor of harm and **9** indicating the **weakest** predictor of harm). Please note: There should only be one '1', one '2' etc.

	Rank
A There has been physical violence	
B There has previously been a series of 'verbal only' incidents	
C The victim does not want to engage with police	
D The house is well-presented	
E A criminal offence has been committed	
F The victim seems genuinely frightened	
G The victim is willing to make a statement	
H The victim reports that the abuser tracks their phone	
I The victim has no friends to stay with	

How much do you agree that the following factors explain why victims do not leave abusive environments?

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1 People give as good as they get in domestic abuse					
2 They are financially dependent on the perpetrator					
3 They would fear for their safety if they left					
4 Victims and perpetrators are often 'as bad as each other'					
5 Victims want to make the relationship work					
6 Victims feel they have nowhere else to go					
7 The abuser is preventing them from leaving					
8 Victims believe their abuser will change					
9 Victims fear the unknown					
10 Victims stay to protect children					

References

- Cohen J (1969) *Statistical power analysis for the behavioural sciences*. New York: Academic Press.
- Cumming G (2013) Cohen's d needs to be readily interpretable: Comment on Shieh (2013). *Behavior Research Methods*, 45(4): 968–971.
- HMIC (2014) *Everyone's business: Improving the police response to domestic abuse*. Report, Her Majesty's Inspectorate of Constabulary, UK, March.
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2): 65-70.
- Kelly L (1999) *Domestic Violence Matters*. London: Home Office.
- Lehmann P, Simmons C and Pillai V (2012) The validation of the Checklist of Controlling Behaviours (CCB): Assessing coercive control in abusive relationships. *Violence Against Women*, 18(8): 913-933.
- Miller J and Alexandrou B (2016) *College of Policing stop and search training experiment: Impact evaluation*. Ryton on Dunsmore: College of Policing.
- Morgan M (2015) *A review of domestic abuse training and development products and services provided by the College of Policing*. Unpublished report.
- Robinson A, Myhill A, Wire J, Roberts J and Tilley N (2016) *Risk-led policing of domestic abuse and the DASH risk model*. What Works Centre: Crime Reduction Research.
- Wire J and Myhill A (2016) *Domestic Abuse Matters: Evaluation of first responder training*. London: College of Policing.