

Considering a counter-reformation in the psychology of criminal conduct for women: Converging evidence, confidence intervals and consultation¹

M.A. Law²

Department of Psychology, Carleton University

When Martinson erroneously proclaimed “nothing works,” he provoked a series of unfortunate events in North American corrections as punitive policies decimating rehabilitative efforts became mainstream (see Andrews & Bonta, 2003; Martinson, 1974). Consequently, the past 30 years has seen the field of the psychology of criminal conduct (PCC) blossom in an effort to show what things do work. PCC has been based on empirical evidence showing it is possible to classify offenders, predict their behaviour and identify effective treatment programs for them (Andrews, 1982; Andrews & Bonta, 1998).

A new question is now on the horizon: Is PCC equally applicable to female offenders? The emerging evidence suggests it is (Blanchette & Brown, 2006; Law, 2004; Motiuk & Blanchette, 2000; Rettinger, 1998). It also may be time, however, to seriously question whether the application of PCC with women can or should develop in a lagging but parallel manner to that done with male offenders (Andrews & Wormith, 1989; Bonta, 1995; Cullen, 1995; Dowden & Andrews, 1999; Gendreau & Ross, 1979; Motiuk, Bonta & Andrews, 1986).

Moving forward

The current political climate is demanding increased research on female offenders, particularly on issues surrounding violent behaviours (Auditor General’s Report, 2003; Canadian Human Rights Commission, 2005; Report from Correctional Investigator, 2006). Unfortunately, researchers and arm-chair critics alike are responding with an unfocused cacophony of publications (Blanchette, 2005; Bloom, Owen & Covington, 2004; Hannah-Moffat, 2004; Hardyman & Van Voorhis, 2004). There is no continuity in the research topics, no unifying theory and no direction researchers agree upon to build an evidence-base involving women in conflict with the law (Bloom, 2003; Hannah-Moffat & Shaw, 2000). It may be time to begin the messy business of proposing *practical* steps that may be helpful when exploring PCC with women.

1. Converging evidence: Look for it!

There is evidence that even in the current arena of quantitative research on women offenders, where little has been done in terms of well-designed empirical studies, there are areas of convergence between polarized theoretical camps.

A recent meta-analysis studying predictors of recidivism and institutional misconduct with women (Law, Sullivan & Goggin, 2006) identified substantial overlap between the two leading theories used to explain female criminal conduct: social learning theory and feminist theory. While many predictors are claimed only by social learning theory (e.g., past criminal behaviour and attitudes) or only by feminist theory (e.g., victimization and mental health), many predictors are considered foundational in both perspectives, such as education, employment, associates and family history. The importance of some of these constructs in predicting general and violent recidivism was confirmed. For example, education [$r=.27$, 95% confidence interval (CI) = .21, .33] and employment ($r=.13$, 95% CI=.07, .19) predicted general recidivism. The convergence of these two theoretical camps that often claim to be diametrically opposed to each other is noteworthy.

2. Confidence intervals: Don’t be afraid!

Another avenue to potentially deepen our understanding of women’s criminal conduct is by examining confidence intervals and not just p -values in our quantitative work.

Confidence intervals have begun to garner attention in forensic psychological research, yet despite the calls by important organizations such as the American Psychological Association (Finch, Thomason & Cumming, 2002) and peer-review psychology journals such as the *Journal of Consulting and Clinical Psychology* (Fidler et al., 2005), researchers have been hesitant to fully embrace confidence intervals.

The ability to automatically generate confidence intervals using statistical software has led to

Confidence intervals provide an intuitive and quick glimpse at the variability associated with an estimate, clearly pointing out . . . the range within which one can be certain the true value of the correlation lies.

researchers reporting but not understanding them, and to readers glancing at but generally ignoring them (Fidler, Thomason, Cumming, Finch & Leeman, 2004). One possible reason for this is a poor understanding of confidence intervals and how they can enhance the reporting and interpretation of results (Belia, Fidler, Williams & Cumming, 2005; Cumming & Finch, 2005).

Although both p -values and confidence intervals provide information about the statistical significance of an estimate – for example, how strong the relationship is between a variable like age and an outcome like serious institutional misconduct – confidence intervals have the added advantage of providing information on how certain we are that this is the correct value for the estimate. In criminal psychology, this point estimate is usually the Pearson correlation coefficient r . Confidence intervals provide an intuitive and quick glimpse at the variability associated with an estimate, clearly pointing out the likely lower and upper limits of the point estimate, i.e., the range within which one can be certain the *true value* of the correlation lies. If the value, $r=0$, does not lie within the range of a 95% confidence interval, the result is statistically significant at $p < .05$.

Practical steps to interpreting and understanding 95% Confidence Intervals for correlation coefficients

- 1) Look to see if 0 (i.e., a Pearson correlation coefficient indicating no association between two variables) is between the lower and upper limit of the 95% confidence interval.
 - If 0 is not within the two limits of a 95% confidence interval, then the correlation is statistically significant at $p < .05$.
 - If 0 is within the two limits of a 95% confidence interval, then it is not possible to say with certainty that the value of the correlation is not 0 (i.e., that there is no association between the two variables) and

the correlation is not statistically significant at $p < .05$.

- 2) Determine the width of the confidence interval by calculating the difference between the lower and upper confidence limits. (A wide confidence interval often reflects a diverse population or a small sample size).
 - If the width of the interval is greater than .1, then there is excess variability associated with the correlation and it is unreliable (personal communication, Paul Gendreau).
 - If the width of the interval is .1 or less, the variability associated with estimating the correlation is acceptable.
- 3) Examine the upper and lower limits of the confidence interval to identify the minimum and maximum values that the correlation could be likely to take on. Since we know that the point estimate is only an estimate and not the true correlation, this gives the possible range of values that the true correlation could take on.

For example. . .

In a recent study (Law, 2004), data was collected on 640 federally sentenced women in the community examining seven criminogenic need domains from the Community Intervention Scale (CIS), formerly the Community Risk Needs Management Scale (CRNMS). These need domains are: associates, attitudes, community functioning, employment, marital/family, personal/emotional and substance abuse. Measures of community adjustment were re-offences coded from Canadian Police Information Center (CPIC) files providing official recidivism data. The analysis and discussion to follow is also relevant to other small subpopulations such as aboriginal, violent and sex offenders.

Analysis

The correlation between type of re-offence and the subscale score on the CIS for each criminogenic need domain was calculated, and p -values were reported for each correlation. In addition, 95% confidence intervals were reported for each correlation. An alpha level of 0.05 was considered statistically significant. An estimate was considered stable if the width of the confidence interval was less than or equal to .1. Analyses were done using SAS 9.1.3 (SAS Institute Inc., Cary, North Carolina).

Results

Table 1

Correlation between Domains of the CIS and Any Type of Re-offence by Female Offenders after Release (N=640)							
CIS domain	Pearson's correlation coefficient	p-value	Statistically significant (p < 0.05)	95% CI	Statistically significant (95% CI includes 0)	Width of 95% CI	Acceptable variability (width of CI <= 0.1)
	(r)						
Employment	0.26	< 0.0001	YES	0.18, 0.33	YES	0.15	NO
Family	0.13	0.001	YES	0.05, 0.20	YES	0.15	NO
Associates	0.19	< 0.001	YES	0.12, 0.27	YES	0.15	NO
Substance abuse	0.14	0.0005	YES	0.06, 0.21	YES	0.15	NO
Community	0.17	< 0.0001	YES	0.10, 0.25	YES	0.15	NO
Personal/Emotional	0.14	0.0006	YES	0.06, 0.21	YES	0.15	NO
Attitudes	0.16	< 0.0001	YES	0.09, 0.24	YES	0.15	NO

Discussion

The correlation coefficients observed between recidivism and criminogenic needs with female offenders were comparable to similar research on male offenders (Motiuk & Brown, 1997). Examining the correlation coefficient (column 2) across CIS domains, employment emerges as the most important predictor for any type of re-offence ($r=.26$, $CI .18$ to $.33$). In a field where correlations greater than .1 are noteworthy, coupled with the fact that women are a heterogeneous population, these are correlations to cause excitement. This in itself raises concern, especially when one considers that the data in Table 1 represents not simply a sample drawn from the population but was in fact the entire targeted population at the time of the study.

... a statistically significant result ... of small magnitude may have little clinical or practical significance for women offenders in terms of policy or programming.

Any interpretation of quantitative data should consider the magnitude (r), direction (+/-) and statistical significance (p -value) of the result. Nonetheless, many researchers seem to emphasize only the statistical significance (p -value) of results (Finch, Cumming & Thomason, 2001). Thus, a statistically significant result (p -value) of small magnitude (r) may have little clinical or practical

significance for women offenders in terms of policy or programming.

It is time to ask difficult self-critical questions like: Are we *asking* the right question? Are we asking it in the right manner? Why is it such a small relationship? What may be missing?

While many of the correlations between re-offence and criminogenic needs were statistically significant at $p < .05$, examination of the 95% confidence intervals reveals that there is substantial variability associated with these correlations. Looking at Table 1, none of the confidence intervals are narrower than .1, the present standard in criminal justice literature (personal communication, Paul Gendreau). This is disconcerting. One might then pose the question: Is .1 an appropriate standard to use when conducting research on women offenders? While this standard might be appropriate when conducting research on male offenders, who are a relatively large and homogeneous population, it may not be appropriate for the small and diverse population of women offenders.

3. Consult! Consult! Consult!

Rather than pursuing conventional deconstructionist thinking that would cease efforts to predict women offender behaviour, this may be a muted call to operationalize established constructs from PCC in a truly gender-informed manner. An example of this would be employment domains that reflect women's histories as wives, single-mothers and/or having parenting responsibilities (Blanchette & Brown, 2006). As the constructs become more "informed"

It may be time to return to grassroots, exploratory, qualitative research in order to appropriately inform and direct future quantitative research

and “specific,” the quality of quantitative relationships (r) should increase.

It may be time to return to grassroots, exploratory, qualitative research in order to appropriately inform and direct future quantitative research, rather than blindly mapping on to the progression of PCC in recent years by researchers with male populations (see Andrews & Bonta, 2003). This can be accomplished by: 1) engaging in meaningful consultation to deconstruct women’s experiences along pathways into, through and out of the penal system; and 2) considering environmental factors intertwined in these correlations, thereby responding to feminist critics’ unceasing calls for the contextualization of women’s experiences (Bloom, 2003; Hannah-Moffat, 2000).

There is a high probability that such a return to qualitative research would continue to uncover compatibilities and converging validities between social learning theory and feminist theory on pertinent factors (e.g., relationships/associates). This would provide a blended and stronger theoretical reorientation from the heated debate that has evolved between feminist and social learning perspectives.

It is time to halt the polarization of theorizing on female offenders . . . and to improve dialogue as more converging evidence becomes apparent.

Conclusion

While confidence intervals clearly add value to the interpretation of data on women offenders by providing information on the variability of the estimates, careful scrutiny of these data reveal notable latent issues. It is time to consider not just the statistical significance of results . . . and to start scrutinizing more closely the magnitude of correlations.

It is also time to begin merging historically incompatible lexicons, to recognize, for example, that “contextualization” for feminist theorists (Chesney-Lind, 2001) may be congruent with the “community reinforcement” component of the Personal, Interpersonal, Community Reinforcement (PIC-R) theory (Andrews, 1982). It is time to halt the polarization of theorizing on female offenders . . . and to improve dialogue as more converging evidence becomes apparent. It is time to stop assuming that risk/needs constructs developed with male offenders have identical underlying structures for female offenders . . . and return to exploratory qualitative research in the hopes of informing the operationalization of these constructs for future quantitative research.

Remember that statistical significance is not clinical significance.

The game plan

- Embrace the transition from null hypothesis testing (p -values) to reporting of confidence intervals. Studies conducted in small offender sub-populations – e.g., violent offenders, female offenders, sex offenders – should always report confidence intervals. This will allow readers to easily see the variability associated with these estimates and potential maximum values that correlations may take on.
- Recognize that variability may be greater in small heterogeneous sub-populations such as women offenders, aboriginals, violent offenders and sex offenders. Obtaining a confidence interval width of .1 may be unrealistic at this stage. Bands of .15 appear to be a more realistic standard to set, at least for women offenders.
- Remember that statistical significance is not clinical significance. Concrete efforts need to be made to increase the magnitude of correlations being obtained with women offenders by pursuing gender-informed constructs, exploring situational variables, and validating the applied nature of these relationships.
- Assume nothing: return to full consultation with the women being served in the correctional system with the vision of conjoining qualitative and quantitative methods. ■

- ¹ This article is partially based on the work of Law, M.A. (2004). *A longitudinal follow-up of federally sentenced women in the community: Assessing the predictive validity of the dynamic characteristics of the Community Intervention Scale*. Unpublished doctoral dissertation. Ottawa, ON: Carleton University. Advisor: D.A. Andrews.
- ² Veritas Research Institute, 2746 Rothesay Road, Rothesay N.B., E2H 2L2, e-mail: moiralaw@nbnet.nb.ca.

References:

- Andrews, D.A. (1982). *A personal, interpersonal, and community reinforcement (PIC-R) perspective on deviant behaviour*. Toronto: Ontario Ministry of Corrections Services.
- Andrews, D.A., & Bonta, J. (1998). *The psychology of criminal conduct* (2nd ed). Cincinnati, Ohio: Anderson Publishing Company.
- Andrews D.A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed). Cincinnati, Ohio: Anderson Publishing Company.
- Andrews, D.A., & Wormith, S. (1989). Personality and crime: Knowledge destruction and construction in criminology. *Justice Quarterly*, 6, 289-309.
- Belia, S., Fidler, F., Williams, J., & Cumming, G. (2005). Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods*, 10(4), 389-396.
- Blanchette, K. (2005). *Field-test of a gender informed security reclassification scale for female offenders*. Unpublished doctoral dissertation, Carleton University, Ottawa, Ontario, Canada.
- Blanchette, K., & Brown, S.L. (2006). *Assessment and treatment of women offenders: An integrative perspective*. Wiley Series in Forensic Clinical Psychology, Chichester, England.
- Bloom, B.E. (2003). A new vision: Gender responsive principles, policy, and practice. In B.E. Bloom (Ed.), *Gendered justice: Addressing female offenders* (pp.267-288). Durham, NC: Carolina Academic Press.
- Bloom, B., Owen, B., & Covington, S. (2004). Women offenders and the gendered effects of public policy. *Review of Policy Research*, 2(1), 31-48.
- Bonta, J. (1995). The responsibility principle and offender rehabilitation. *Forum on Corrections Research*, 7, 34-37.
- Brennan, T. (1987). *Classification for control of jails and prison*. In D.M. Gottfredson and M. Tony (Eds.), *Prediction and classification: Criminal justice decision making* (pp.323-366). Chicago: University of Chicago Press.
- Canada, Auditor General. (2003). *Correctional Services Canada – Reintegration of women offenders*, Chapter Four. [http://www.oag_bvg.gc.ca/domino/reports.nsf/html/20030404ce.html/\\$file/20030404ce.pdf](http://www.oag_bvg.gc.ca/domino/reports.nsf/html/20030404ce.html/$file/20030404ce.pdf).
- Canadian Human Rights Commission. (2003). *Protecting their rights: A systemic review of human rights in correctional services for federally sentenced women*. http://www.chrc_ccdp.ca/legislation_policies/consultation_report_en.asp.
- Chesney-Lind, M. (2001). Contextualizing women's violence and aggression: Beyond denial and demoralization. *Behavioural and Brain Sciences*, 22, 222-223.
- Cullen, F.T. (1995). Assessing the penal harm movement. *Journal of Research in Crime and Delinquency*, 32, 338-358.
- Cumming, G., & Finch, S. (2005). Inference by eye. Confidence intervals and how to read pictures of data. *American Psychologist*, 60(2), 170-180.
- Dowden, C., & Andrews, D.A. (1999). What works for female offenders: A meta-analytic review. *Crime and Delinquency*, 45(4), 438-452.
- Fidler, F., Cumming, G., Thomason, N., Pannuzzo, D., Smith, J., Fyffe, P., Edmonds, H., Harrington, C., & Schmitt, R. (2005). Toward improved statistical reporting in the Journal of Consulting and Clinical Psychology. *Journal of Consulting and Clinical Psychology*, 73(1), 136-143.
- Fidler, F., Thomason, N., Cumming, G., Finch, S., & Leeman, J. (2004). Editors can lead researchers to confidence intervals, but can't make them think. Statistical reform lessons from medicine. *Psychological Science*, 15(2), 119-126.
- Finch, S., Cumming, G., & Thomason, N. (2001). Reporting of statistical inference in the Journal of Applied Psychology: Little evidence of reform. *Educational and Psychological Measurement*, 61, 181-210.
- Finch, S., Thomason, N., & Cumming, G. (2002). Past and future APA guidelines for statistical practice. *Theory and Psychology*, 12, 825-853.
- Gendreau, P., & Ross, R. (1979). Effective correctional treatment: Bibliotherapy for cynics. *Crime and Delinquency*, 25, 463-489.
- Hannah-Moffat, K. (2004). Gendering risk at what cost: Negotiations of gender and risk in Canadian women's prisons. *Feminism and Psychology*, 14(2), 243-249.
- Hannah-Moffat, K. (2000). Re-forming the prison: Rethinking our ideals. In K. Hannah-Moffat & M. Shaw (Eds.), *An ideal prison? Critical essays on women's imprisonment in Canada* (pp.30-40). Halifax, Nova Scotia: Fernwood Publishing.
- Hannah-Moffat, K., & Shaw, M. (2000). (Eds.). *An ideal prison: Critical essays on women's imprisonment in Canada*. Halifax, NS, Canada: Fernwood Publishing.
- Hardyman, P., & VanVoorhis, P. (2004). *Developing gender-specific classification systems for women offenders*. National Institute of Correction (NIC Accession Number 018931).
- Harer, M., & Langen, N. (2001). Gender differences in predictors of prison violence. Assessing the predictive validity of a risk classification system. *Crime and Delinquency*, 47, 513-536.
- Law, M.A. (2004). *A longitudinal follow-up of federally sentenced women in the community: Assessing the predictive validity of the dynamic characteristics of the Community Intervention Scale*. Unpublished doctoral dissertation, Carleton University, Ottawa, Canada.
- Law, M.A. (2007). *Beyond a risk-needs paradigm: Initial security classification protocols with federal women offenders. Applications of feminist methodologies*. Unpublished transitory record submitted to Women Offender Sector, Correctional Service Canada, Ottawa, Canada.
- Law, M.A., Sullivan, S., & Goggin, C. (2006). *Security classification measures of female offenders and predictors of female criminal conduct: A literature review*. Unpublished transitory record submitted to Women Offender Sector, Correctional Service Canada.
- Martinson, R. (1974). What works? – Questions and answers about prison reform. *The Public Interest*, 35, 22-54.
- Motiuk, L.L., & Brown, S.L. (1993). *The validity of offender needs identification and analysis in community corrections*. Research Report R-34. Ottawa, ON: Correctional Service Canada.
- Motiuk, L.L., & Blanchette, K. (2000). Assessing female offenders: What works. In M. McMahon (Ed.), *Assessment to assistance: Programs for women in community corrections* (pp. 235-266). Arlington, VA: American Correctional Association.
- Motiuk, L.L., Bonta, J., & Andrews, D.A. (1986). Classification in correctional half way houses: The relative and incremental predictive criterion validities of the Megargee-MMPI and LSI systems. *Criminal Justice and Behaviour*, 13, 33-46.
- Office of the Correctional Investigator. (2006). *Response to the Canadian Human Rights Commission's Consultation Paper for the Special Report on the Situation of Federally Sentenced Women*. http://www.oci-bec.gc.ca/reports/OCIResponse_CHRC_e.asp.
- Rettinger, J. (1998). *A recidivism follow-up study investigating risk and need within a sample of provincially sentenced women*. Unpublished doctoral dissertation. Carleton University, Ottawa, Ontario, Canada.
- Walters, G.D. (2006). Risk-appraisal versus self-report in the prediction of criminal justice outcomes: A meta-analysis. *Criminal Justice and Behaviour*, 33(3), 279-304.