

Using the Empathy Assessment Index (EAI) to measure long-term Empathy changes as a result of Basic AVP Workshop participation: Phase 1

Don Nason: nason@cox.net
November 2017

Summary

This report summarizes the results and conclusions to date from a preliminary (Phase 1) study of changes in individual empathy as a result of participation in a Basic level Alternatives to Violence Project (AVP) workshop. The tool used is the Empathy Assessment Index (EAI) questionnaire, administered confidentially to inmates in Delta yard, a Level IV Protective Custody (SNY) yard at the Substance Abuse Treatment Facility (SATF) prison in Corcoran CA. Results for an AVP Basic workshop from a group of fifteen participants, using a before-after, matched groups design, were significant at lower than a 5% level, which means minimal Type I error. However, the Power = 0.6 indicates a higher than desirable Type II error. A necessary concern is the small actual empathy change which, though statistically significant, may not be large enough to be considered of practical importance. The results are meaningful enough that further studies are underway (Phase 2) using larger numbers of participants, tighter testing protocol, and data based on multiple levels of AVP workshops.

Background

The motivation for this study was the desire to identify some personal characteristics that are expected to be enhanced by participation in Alternatives to Violence Project (AVP) workshops, and then find a way to measure that effect. The goals are to help understand what makes the AVP program successful and to ultimately attract significant financial and institutional support for AVP. There are several human qualities that suggest themselves as candidates for study. A prominent one is *anger*, which has been investigated for example using the STAXI (State-Trait Anger Expression Inventory) by Kayser, Roberts, Shuford and Michaelis (2014). Another one is *community building*. Another one that is frequently noticed by both AVP facilitators and participants is *empathy*. So far the evidence linking empathy changes to AVP participation, though convincing to AVP adherents, is anecdotal, mainly based on the recollections of participants who believe they were impacted.

It is recognized that in order to be persuasive in the Human Relationships Community or Corrections Community beyond AVP, the evidence ultimately needs to be obtained by independent evaluators using established statistical methods of analysis and testing protocols which may involve automated data handling to assure full confidentiality of the participants. Although such procedures may indeed be needed for full professional credibility of results, it is clear that valuable preliminary work could be done in-house by AVP itself, based on our convenient access to the data, before the effort of a full analysis is undertaken. This preliminary work could help answer questions such as whether the Empathy Assessment Index (EAI) is an appropriate tool for prison inmates from a diverse range of backgrounds, educational levels and prison experiences. Also, if there is a significant empathy effect of AVP, is it large enough to be of practical significance? This report deals with Phase 1, a preliminary study which begins to answer these questions and determine if the work should proceed further. If so, Phase 2 will refine the techniques and design of the study, with the expectation that together the results will indicate if a full scale, professionally run and analyzed, research project should be undertaken to identify and measure the effect of AVP workshop participation on empathy.

In discussions with AVP colleagues, *empathy* was identified as an AVP-related quality, and AVP colleague John Shuford suggested the EAI, a tool for measuring empathy (see Appendix 1) using a self-administered questionnaire. The EAI questionnaire was developed, tested and validated at Arizona State University (Gerdes, Segal & Leitz, 2012). More information about the questionnaire and background for its development are included in Appendix I. The questionnaire itself is shown in Appendix 2 and proposes that empathy has five components---affective response, affective mentalizing, self-other awareness, perspective taking, and emotion regulation. Twenty-two statements, Q1 to Q22, each cover one of these five components, and responders select their response on a scale from “1” (never) to “6” (always), which covers a full range between opposite extremes of response. This defines an *empathy* scale and enables a numerical evaluation.

The EAI has a built-in feature to test for self-consistency of responses. Q2 is a reverse sense statement of Q10, and Q17 is a reverse sense statement of Q5. For example perfect consistency in responses would mean that if a Q5 response was “4”, then Q17 from the same individual should be “2”, which reverses to a “4”: If the response for Q10 was “1”, then consistency requires that their Q2 response should be “6”, which reverses to “1”. The degree to which consistency is lacking could indicate careless or random responses. Treated statistically for many responders, this is useful in assessing the appropriateness of the test procedure and the attentiveness of the responders as they complete the questionnaire.

Procedure

The first step in conducting a pilot study using a questionnaire such as the EAI is obtaining the required approvals. The EAI and proposed study were presented as a tool intended to evaluate the AVP program effectiveness. The work is not intended to critique or research individuals or the CA prison system. Group average, not individual, responses are the objective, and AVP already has the permission of the CA prison system to assess individual workshops using post-workshop participant questionnaires. The project was approved by the CSATF Warden.

Some initial EAI data was collected at SATF and in community AVP workshops to establish a testing protocol, which was refined somewhat as experience was gained and in consultation with AVP colleague Pat Hardy. In essence, when all the workshop participants were first seated and ready, they were given the same instructions. The participants were assured that the interest was in the group average responses. No-one's personal responses would become public or part of any workshop nor affect their opportunity to participate in future AVP workshops. To this end, respondents were asked to code their questionnaire with their birth-date, not their name, each time they did the EAI questionnaire. Thus, any changes in responses could be tracked from workshop-to-workshop using only their coded identify. Many people chose to mark their questionnaire with their name anyway. Since no-one questioned the coding, and nearly all the distributed blank questionnaires were completed and returned, it appears there was no significant issue of confidentiality of responses. The completed questionnaires were kept in a private location outside the prison.

This study involved inmates doing Basic and Advanced AVP workshops in the SATF California prison, Level IV (highest level) SNY yard. Although this prison was initially named a Substance Abuse Treatment Facility, it apparently has never functioned that way, and the inmates receive nothing more than treatment available in most other CA prisons. The designation SNY means Special Needs Yard. Inmates can apply for transfer to do their time in such yards if they or prison management believe that they are in danger and need special protection from other inmates, usually based on their intentional withdrawal from prison gangs or the nature of their offense. Typically, there are several SNY yards in a CA prison.

Participants in the study completed the questionnaire two times, called *before* and *after*. (See Appendix 2 for a copy of the questionnaire used.) First, participants in Basic workshops completed the 22-question EAI questionnaire just as they sat down to begin their Basic workshop (*before*). Inmates repeated the questionnaire two-to-four months later, just as they began their Advanced workshop (*after*), if in fact they went on to do an Advanced. Inmates with coded names that were in both these *before* and *after* groups constitute the tested group for this Phase 1 study. The time delay between *before* and *after* results was intentional, in order to best capture the lasting effects of the AVP Basic workshop and to avoid transient effects like workshop fatigue and/or the euphoria which may pertain to a just-completed workshop.

With the coded name identification, the *before* and *after* responses can be compared for each respondent on each of twenty-two statements related to Empathy, Q1 to Q22. This enables the analysis to be done as a matched-pair statistical analysis.

The analysis did not seek to distinguish between the five components of empathy, but bunched all five empathy components together. There were fifteen test subjects in the *before-after* calculations and thirty-one subjects in the scoring/reverse scoring calculation that tests for validity. These comparatively small numbers occurred because earlier rounds of questionnaires were disregarded because of intentional or unintentional variations in the test protocol as it was being refined, and because the questionnaire itself was updated from four to five empathy components and from twenty to twenty-two questions. The analysis is based on standard elementary statistics, as presented for instance, in Cohen and Lea (2004) or Gravetter and Wallnau (2011) or many other sources. Specifically, this is a **repeated measurements *t* test**. It involved simple arithmetic, aided by a hand calculator which has a statistical package. Subsequent equation numbers refer to Cohen and Lea (2004).

Results

Several aspects of the results are of interest. The earliest results showed that the **testing protocol needed to be better controlled and refined**. Trying various protocols, it became clear that the moment in the workshop when the questionnaire was done appeared to make a big difference. Doing the questionnaire at the beginning of a workshop took typically less than five minutes, and nearly everyone participated. Early was better than late, as respondent focus sometimes lagged at the end as indicated by skipped questions and seemingly repetitive answers. Using a two page questionnaire format was better than a two-sided, one-page format, where the whole backside was sometimes skipped.

One result from the raw data is that the average EAI index before Basic is “4.38”. Since “4” represents *Frequently*, and “5” represents *Almost Always*, the average responder apparently already considers himself to be already quite empathic even before AVP.

Another result related to validity is the self-consistency of the results from each individual. The questionnaire was designed to enable consistency to be checked, in that two questions, Q5 and Q10, each have a reverse-sense counterpart, as described above. A Null hypothesis analysis showed that there was practically no difference, when averaged over thirty-three responders, between a response and its reverse-sense counterpart. This encourages confidence in the general attention, effort and comprehension of the participants during this Phase 1 study.

The main result obtained is the change in the EAI index *before* and *after* the Basic workshop. For each responder, the empathy index value for each of the twenty-two statements *before* the Basic workshop was subtracted from each value long *after* that Basic. These twenty-two values were then averaged to give an averaged empathy change for that responder. These latter values were then

averaged over the fifteen responders to give the average empathy change for the group participants and produced $diff=0.39$, with standard deviation $sigma=0.53$.

The results can be interpreted with standard **Null hypothesis testing** to compute the **Type I error**, which relates the probability $alpha$ of concluding that there is a real empathy difference, when more thorough sampling would show that in fact there is not. This result using eqn. (3-12) gave $t = alpha = 2.8$, which translates (Table A-2) to only 1.5% probability that the empathy difference can be attributed to chance and not to a real difference. The usual criterion for acceptable Type I error is $< 5 \%$. Our result would satisfy an even more rigorous criteria of 2%. Thus, **the Null hypothesis can be rejected and our measured empathy change is concluded to be significant.**

A further step in **Null hypothesis testing** is a **Power Analysis** to compute the **Type II error**. This is the probability $beta$ of concluding that there is no significant empathy difference when in fact there is. The numerical complement of $beta$, $1.0 - beta$, is the **Power**. This analysis yields (interpolating in Table A-5) $Power = 0.6$. By common convention, $Power > 0.8$ is taken as the criterion for deciding the observed empathy effect is significant. The present results fall short by the $Power$ criterion. However, $Power$ can be increased by having a larger number of responders. In this situation it appears that increasing the respondent number by a factor of three could allow the $Power$ result to rise into the desired range, other factors being equal, so this is a promising direction for a Phase 2 study.

A very important practical aspect of any empathy gain is whether it is large enough to make any practical difference. Presumably, observers of behavior would be able to distinguish between adjacent steps on the empathy scale, just as questionnaire responders must do as they complete the questionnaire, understanding for example the difference between *Rarely* and *Sometimes*, or between *Frequently* and *Almost Always*, and so forth. Although the average empathy gain $diff = 0.39$ observed here is statistically significant, does this small empathy gain make a practical significance? Would it be noticeable in behavior? If the smallest distinguishable increment on the EAI scale is estimated as one-half the unit step size, or 0.5, meaning that any *empathy* difference smaller than this might be unnoticeable. Our $diff$ result is only comparable to this presumed distinguishability limit. Thus, **these Phase 1 results neither affirm nor disaffirm that empathy changes as measured by the EAI are large enough to be of practical behavioral importance.**

The results and concerns of the above work constitute Phase 1. of this study. Thanks are extended to Dr. Laura Roberts, a Statistics Educator, and to Graeme Stuart for comments on the manuscript.

Outlook/Plans for Phase 2

Phase 2 is underway, making use of results and changes suggested by Phase 1.

A study in another prison (Williams, 2012) dealing with inmate behavior, before and after AVP, concluded that the inmates most likely to be strongly affected by AVP were those who had the least previous programming in prison. With this in mind, the tested population in Phase 2 will come from a “mainline” facility at SATF. This Level IV prison yard has inmates with the most restricted movement and programming and who have not opted for reclassification to Protected Custody status in a Special Needs Yard (SNY). In contrast, Phase 1 used inmates from a protected custody Level IV yard, where inmates in general had more prior opportunity to do prison programming. There is a good possibility that the $diff$ values will be larger in Phase 2.

Important adjustments are being made in the testing protocol. Previous experience showed that it is important to make sure that the mood while the questionnaire is filled out is calm and unhurried. There is a temptation to ask for the participants' cooperation “for a few moments to fill out a quick questionnaire” before the real workshop begins. This encourages hurry and is now strictly avoided.

Instead, the message is to take whatever time is needed to provide thoughtful responses to all twenty-two questions. Additionally, assurances need to continue to be given that the results are for AVP and are not at any time seen or evaluated by prison staff.

Another change for Phase 2 will be increasing the size of the participant groups. The fifteen participants submitting questionnaires used in the *before-after* analysis in Phase I were the remnant remaining after eliminating results obtained when the protocol was developing. Though useful for the analysis, these data were clearly not normally distributed in the statistical sense. Although there is no theoretical reason to assume that the mean empathy difference *diff* will be different for a larger sample, the analysis tends toward greater accuracy overall when larger number of participants are involved because larger numbers are more likely to follow the “normal distribution” that is the basis of the theory. The goal in Phase 2 is to have fifty or more responders, wherever possible. Other things being equal, the *Power* resulting from this larger number of participants would exceed the target value of 0.8.

In Phase 2 an expansion into further stages of AVP is expected. If the empathy effect is significant or nearly significant for the Basic-to-Advanced step as in Phase 1, data taking could be extended to cover further stages in AVP training such as the Training for Facilitators (T4F) or even Apprentice Facilitator stages. Here, a marginally significant result may become significant, and this could illuminate the question of how much AVP workshop exposure is necessary for empathy changes that are large enough to be of practical importance. The analysis in these cases could still employ a two-group, matched groups mode as in Phase 1. Alternatively, if a more conventional control group is desired, a two-group, independent groups mode could be used, though this would result in lower *Power* values. Additionally, it may become worthwhile to expand the scope of the project by separating out the data for the specific components of empathy (five in the EAI case) and analyze them separately for greater insight into an AVP effect. These goals should all be possible by operating with in-house AVP resources as in Phase 1.

With the progress expected in Phase 2, coupled with the promising start of Phase 1, it will hopefully become clear whether or not empathy enhancement is an important result of AVP workshops. The results could provide a solid basis for a full-scale study by an independent research agency, with the attendant costs and potential gains in AVP credibility and financial support. If not, it may point to the need for research on other human characteristic which anecdotal evidence has shown are so beneficially affected by participation in AVP workshops.

References

- Cohen, B. H., & Lea, R. B. (2004). *Essentials of statistics for the social and behavioral sciences*. Hoboken, NJ: Wiley.
- Gerdes, K. E., Geiger, J. M., Lietz, C. A., Wagaman, M. A., & Segal, E. A. (2012). Examination of known-groups validity for the empathy assessment index (EAI): Differences in EAI scores between social service providers and service recipients. *Journal of the Society for Social Work and Research*, 3(2), 94-112. doi: 10.5243/jsswr.2012.7 Available from <http://www.jstor.org/stable/10.5243/jsswr.2012.7>
- Gerdes, K. E., & Segal, E. A. (2009). A social work model of empathy. *Advances in Social Work*, 10(2), 114-127.
- Gravetter, F. J., & Wallnau, L. B. (2011). *Essentials of statistics for the behavioral sciences* (7th ed.). Belmont, CA: Wadsworth/Cengage Learning.
- Kayser, T., Roberts, L., Shuford, J., & Michaelis, J. (2014). Minnesota AVP anger study. *International Journal of Trauma Research and Practice*, 1(1), 2-13. Available from <http://traumapro.net/wp-content/uploads/2014/06/IJTRP-V1-I1.pdf>

- Lietz, C. A., Gerdes, K. E., Fei, S., Geiger, J. M., Wagaman, M. A., & Segal, E. A. (2011). The Empathy Assessment Index (EAI): A confirmatory factor analysis of a multidimensional model of empathy. *Journal of the Society for Social Work & Research*, 2(2), 104-124. doi: 10.5243/jsswr.2011.6
- Segal, E. A., Gerdes, K. E., Lietz, C. A., Wagaman, M. A., & Geiger, J. M. (2017). *Assessing empathy*. New York: Columbia University Press.
- Williams, K. G. (2012). *Using a causal-comparative ex post facto design to examine the effects of the alternatives to violence project on the reduction of incidents of violence and misconduct on inmates housed at California state prison-Solano in Vacaville, California*. (Unpublished Doctor of Education dissertation), University of La Verne, California.

Appendix I:

The Empathy Assessment Index (EAI)

The Empathy Assessment Index (EAI) is a 22 item questionnaire which underwent extensive development over a four year period with over 3500 participants (Gerdes, Geiger, Lietz, Wagaman, & Segal, 2012; Gerdes & Segal, 2009; Lietz et al., 2011; Segal, Gerdes, Lietz, Wagaman, & Geiger, 2017).

It is built on five components of interpersonal empathy as identified by Segal et al. (2017):

(a) **affective response** – *the brain includes neurological pathways that are capable of physiologically simulating the experiences of others. Often referred to as “mirroring” in the literature, this ability is unconscious, automatic and involuntary. For example, if a person starts crying in front of you, even if you do not understand why, you too will feel like crying. Affective response can run through all types of emotions (happy, sad) as well as physical sensations (feeling pain when watching another person being physically hurt). Humans appear to be hard-wired to mimic one another, setting the stage for experientially connecting to another person.*

(b) **affective mentalizing** – *not all physiological reactions or mirroring come from the actual viewing of an event or experience. Often we are exposed to stories or explanations of events, and as we are hearing the information, our mind develops a picture of the events. This allows us to develop perceptions of another’s experiences. It may also trigger an affective or physiological response. When this occurs, we are “mentalizing” or imagining the event and potentially experiencing it as if it is happening to us as well.*

(c) **self-other awareness** – *once the affective response occurs, we need to recognize the difference between the experiences of another person from our own. We may feel like crying (as in the example above) but it is the other person’s experience and not our own. This moves empathic response into a cognitive or conscious arena.*

(d) **perspective-taking** – *assuming that one successfully mirrors and then processes the affective response to understand that it belongs to the other person, it becomes possible to cognitively process what it might be like to personally experience the experiences of another. This is what we commonly refer to as “stepping into the shoes of another.”*

(e) **emotion regulation** – *the last component helps us to move through these affective and cognitive processes without becoming overwhelmed or swept up into someone else’s emotions. This is the ability to sense another’s feelings without becoming overwhelmed by the intensity of the other person’s experience. (p. 121).*

More background to the EAI can be found in an extract from *Assessing Empathy* (Segal et al., 2017) available through Google Books at <http://bit.ly/2gB0uvR>. Their book is available for ordering from <https://cup.columbia.edu/book/assessing-empathy/9780231181914>

Appendix 2: Empathy Assessment Questionnaire

© Karen E. Gerdes, Elizabeth A. Segal, & Cynthia A. Lietz (2012)
Arizona State University

For use with permission of authors

Please respond to the following questions by selecting the choice that most closely reflects your feelings or beliefs:

	Never	Rarely	Sometimes	Frequently	Almost always	Always				
	1	2	3		4	5	6			
1)	When I see someone receive a gift that makes them happy, I feel happy myself. [AR]				1	2	3	4	5	6
2)	Emotional stability describes me well. [ER]				1	2	3	4	5	6
3)	I am good at understanding other people's emotions. [AM]				1	2	3	4	5	6
4)	I can consider my point of view and another person's point of view at the same time. [PT]				1	2	3	4	5	6
5)	When I get angry, I need a lot of time to get over it. [ER] R				1	2	3	4	5	6
6)	I can imagine what the character is feeling in a good movie. [PT]				1	2	3	4	5	6
7)	When I see someone being publicly embarrassed I cringe a little. [AR]				1	2	3	4	5	6
8)	I can tell the difference between someone else's feelings and my own. [SOA]				1	2	3	4	5	6
9)	When I see a person experiencing a strong emotion I can accurately assess what that person is feeling. [AM]				1	2	3	4	5	6
10)	Friends view me as a moody person. [ER] R				1	2	3	4	5	6
11)	When I see someone accidentally hit his or her thumb with a hammer, I feel a flash of pain myself. [AR]				1	2	3	4	5	6

- 12) When I see a person experiencing a strong emotion, I can describe what the person is feeling to someone else. [AM] 1 2 3 4 5 6
- 13) I can imagine what it's like to be in someone else's shoes. [PT] 1 2 3 4 5 6
- 14) I can tell the difference between my friend's feelings and my own. [SOA] 1 2 3 4 5 6
- 15) I consider other people's points of view in discussions. [PT] 1 2 3 4 5 6
- 16) When I am with someone who gets sad news, I feel sad for a moment too. [AR] 1 2 3 4 5 6
- 17) When I am upset or unhappy, I get over it quickly. [ER] 1 2 3 4 5 6
- 18) I can explain to others how I am feeling. [SOA] 1 2 3 4 5 6
- 19) I can agree to disagree with other people. [PT] 1 2 3 4 5 6
- 20) I am aware of what other people think of me. [SOA] 1 2 3 4 5 6
- 21) Hearing laughter makes me smile. [AR] 1 2 3 4 5 6
- 22) I am aware of other people's emotions. [AM] 1 2 3 4 5 6

Contains 5 components: Affective Response [AR], Affective Mentalizing [AM], Self-Other Awareness [SOA], Perspective-Taking [PT], and Emotion Regulation [ER]. AR = 5 items, AM = 4 items, SOA = 4 items, PT = 5 items, and ER = 4 items. Reverse scoring indicated by **R**
 Note: The scores for Questions 5 and 10 are reversed.

More details of the questionnaire and scoring are available from *Assessing Empathy* (Segal et al., 2017).

The authors can be contacted directly:

Elizabeth A. Segal, Ph.D.
 School of Social Work
 Arizona State University
esegal@asu.edu

Karen E. Gerdes, Ph.D.
 School of Social Work
 Arizona State University
kegerdes@asu.edu

Cynthia A. Lietz, Ph.D.
 School of Social Work
 Arizona State University
clietz@asu.edu