

Trait Aggression in the ANES 2013 Internet Recontact Study

Attributes of New Items & Recommendations for Use

Nathan P. Kalmoe
Monmouth College

Introduction

The ANES 2013 Internet Recontact Study included 4 new questions designed to measure trait aggression – a personality factor indicating the relative tendency for people to behave aggressively in everyday life. The four questions proposed in the Online Commons were drawn from a questionnaire used in hundreds of psychological studies, and the four items were selected based on nationally-representative tests to identify the strongest of the 12 items (Kalmoe 2012). However, administrative changes to question wording and response options raise questions about measurement reliability and validity in their new ANES forms.

In this report, I present measurement tests comparing the new items to the original measures in nationally-representative survey data for between-subjects tests and in a bridging study with a convenience sample that included both versions for within-subjects comparisons.

Three of the four new ANES items pass most of these tests. A fourth item measuring the physical component of trait aggression falls short because of seemingly confused responses to an ambiguous inverted question. I ultimately recommend dropping this item from the index, and I suggest replacement wording for that item in future studies. Although content validity is reduced by losing a dimension of the construct in the 2013 data, remaining items are strongly correlated with the original 4-item index and demonstrate sufficient reliability and validity for ANES 2012-13 analyses.

I conclude with examples illustrating the utility of trait aggression in political behavior by using it to predict several forms of argumentative political participation, external efficacy, and political trust.

Trait Aggression & its Measures

Trait aggression is an important personality factor that figures prominently in aggression research as a predictor of aggressive attitudes and behaviors and as a moderator for situational factors that spur aggression (e.g. Anderson & Bushman 2002; Marshall & Brown 2006). Trait aggression was recently introduced in political behavior research to predict support for various forms of politicized violence and to moderate the effects of mild violent cues in political communication (Kalmoe 2013, 2015). Given the roles of hostility, anger, and aggression in many dimensions of political conflict, trait aggression holds great promise as an individual-level predictor of aggression dynamics in politics.

Trait aggression is most commonly measured with the 29-item Buss-Perry Aggression Questionnaire (BPAQ, Buss & Perry 1992), which consolidated the popular but longer Buss-Durkee measures

(Buss & Durkee 1957). More recent measurement efforts reduced the BPAQ to 12 items while retaining its strong psychometric properties (BPAQ-SF, Bryant & Smith 2001).

In 2013, I proposed a 4-item version of the BPAQ-SF for the ANES EGSS4 to investigate aggressive personality's role in political attitudes and behaviors (Kalmoe 2012). I presented analysis from two nationally representative surveys showing that the proposed items were the most robust from each of subscale of the Bryant and Smith battery (i.e. physical aggression, verbal aggression, anger, and hostility). Table 1 presents the text of the proposed items.

Table 1: Proposed Items from the BPAQ-SF

Indicate whether each of the following statements is true or false for you.
1. Given enough provocation, I may hit a person.
2. My friends say I am somewhat argumentative.
3. I have trouble controlling my temper.
4. At times I feel I have gotten a raw deal out of life.
Response scale (6 pt.): Completely true for me, Mostly true for me, Slightly true for me, Slightly false for me, Mostly false for me, Completely false for me

Table 2: Items in the ANES 2013 Internet Recontact Study

<p>1. How much provocation would it take for you to hit someone? [C5_C1] A great deal A lot A moderate amount A little bit None at all</p> <p>2. How argumentative are you – that is, how much do you argue with others? [C5_C2] A great deal A lot A moderate amount A little bit None at all</p> <p>3. How hard is it for you to control your temper? [C5_C3] Extremely hard Very hard Moderately hard A little bit hard Not hard at all</p> <p>4. How often do you feel like you have gotten a bad deal out of life? [C5_C4] Always Most of the time About half the time Some of the time Never</p>
Response options were randomized to swap the order of presentation, first to last or vice versa.

Survey designers will notice these items are somewhat awkwardly worded and they point in the same direction, posing some challenges for interpretation. Administrators reworded these items for the ANES 2013 Internet Recontact Study in a face-valid way that removed some of their awkwardness

but inadvertently discarded the extensive validation behind the proposed items in the process. Table 2 presents text of the ANES items as administered in 2013.

These questions correspond closely with the originals, and for the last three items, the new wording seems clearer. For these three, answering “a great deal,” “extremely,” and “always” indicates more aggressiveness. But responding “a great deal” to the first item indicates the *lowest* level of aggression – it would take a lot to make you hit a person. It is a subtle change from the original, but the wording raises the risk of confusion if respondents don’t read carefully or if they struggle with the mental gymnastics of inverting their answer to a question about aggression in the negative. (Notice also that 6-point scales were reduced to 5.)

In the remainder of this report, I test whether these items perform as well as the originals, with a particular eye on the first item. I find that the last three items do fairly well across the tests, but that the first item has major problems that cast doubt on its utility. Thankfully, the remaining three items mostly prove to be a reasonable stand-in for the four-item measure, with one important exception.

Comparison Studies

To test the new ANES 2013 items, I compare their performance to measures in two other studies. For direct within-subject comparisons of trait aggression items, I included both versions of the questions in an online survey in February 2015 with a nationally diverse set of participants recruited through Mechanical Turk (MTurk) without probability sampling from a population.¹ The questions were identical to those in Tables 1 and 2. The original items were asked at the start of the survey as part of the 12-item BPAQ-SF battery, and the 4 ANES items were asked at the end several minutes later. Participant demographics, socioeconomics, and political views were similar to those reported elsewhere (Berinsky, Huber, & Lenz 2012). Although a nationally-representative U.S. adult sample with both question versions would be ideal, these participants provide strong evidence regarding the comparability of the two trait aggression versions in a diverse national group of participants. Individual responses for personality traits are unlikely to change substantially over the course of a short survey, and so any differences can be plausibly attributed to measurement differences.

For between-subjects comparisons, I fielded two nationally representative U.S. adult surveys with Knowledge Networks (now GfK) in 2010 (KN ’10).² They were completed within one month of each other and both included the original 12-item BPAQ-SF battery at the start. Since the surveys are identical in these respects, I combine them to maximize sample size. These data enable nationally representative inferences about the four BPAQ-SF trait aggression items, so the population is essentially the same sampled in the ANES 2012-13 study two years later. (As a personality trait, aggregate aggression dynamics tested here are unlikely to shift in the population over this time.) The weakness of this data is that it relies on between-subject comparisons. Each comparative study was self-administered online, so mode is identical to the ANES ’13. I assume response process is similar.

Within-Subjects Item Pair Correlations

¹ This data collection was supported by Monmouth College Faculty Research & Development funds.

² This project was supported by the Gerald R. Ford Fellowship and the Marsh Research Fellowship at the University of Michigan. Data for Study 2 were collected by Time-sharing Experiments for the Social Sciences, NSF Grant 0818839, Jeremy Freese and Penny Visser, Principal Investigators.

MTurk data provides the most direct test for how old and new items correspond using within-subjects comparisons. The correlations between matched items are .24 for ‘Hit’, .57 for ‘Argue,’ .59 for ‘Temper,’ and .63 for ‘Bad deal.’ So three of the pairs are strongly correlated and the ‘Hit’ pair is only weakly correlated. Put in terms of scale reliability, ‘Hit’ item pairs have a Cronbach’s alpha of .38 while the other three are at .70, .71, and .77, respectively. These tests raise the first empirical red flag about the new physical aggression item, but the other three items pass the test. Old and new versions of those three items appear relatively equivalent.

How do people respond to the two versions of the ‘Hit’ question? Table 3 shows the percentage of response placements on the new item for each response level of the BPAQ-SF item.

Table 3: Within-Subjects Responses to the Two ‘Hit’ Items in the MTurk Study

BPAQ-SF Version “Given enough provocation, I may hit a person”	ANES Version “How much provocation would it take for you to hit someone?”					Total
	A great deal	A lot	A moderate amount	A little bit	None at all	
Completely false for me	93.08	2.31	0.86	1.15	2.59	100
Mostly false for me	75.73	13.11	1.94	2.91	6.31	100
Slightly false for me	62.07	21.84	5.75	8.05	2.30	100
Slightly true for me	53.80	31.01	8.23	4.43	2.53	100
Mostly true for me	42.03	42.03	11.59	4.35	0.00	100
Completely true for me	55.26	23.68	10.53	0.00	10.53	100

N = 905. Cells indicate percentages by row.

The comparison shows enormous confusion at the end-points, especially the thin high end. *Nearly all* people who placed themselves on the high side of aggression on the BPAQ-SF item put themselves on the low side for the new ANES version: 79% to 85%. Confusion was reduced but still present at the other end of the scale: 4% to 10% who placed themselves at the low end of aggression on the BPAQ-SF item placed themselves on the upper end. No wonder the correlation between them is so low. A pale version of this pattern also appears for other pairs (see Appendix Table A1), but not nearly to this extent, which explains the much higher correlations for those other items.

Comparing Descriptive Statistics in ANES’13 & Similar Samples (Between-Subjects)

Next, I investigate the distributional properties of the four items and their reliability as an index. I create an additive index of the four items and a 3-item index that excludes the problematic ‘Hit’ item. The five-point responses are rescaled so that 1 is most aggressive and 0 is least. The indices retain the 0 to 1 range. No answer, unit non-response, and breakoff are coded as missing and excluded from analysis. Cumulatively, these represent less than 2% of the ANES sample. All

nationally representative tests in this report include sample weighting [c5_weight for ANES data], and the KN combined studies are similarly weighted. Table 4 presents the descriptive statistics.

Table 4: Nationally Representative Descriptive Statistics for 2 Versions (Between-Subjects)

	<i>Mdn</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>Ku</i>	α	<i>N</i>
ANES '13							
Hit	.00	.17	.26	1.75	5.59	(.55) (--)	1,555
Argue	.25	.31	.21	0.74	3.84	(.36) (.44)	1,558
Temper	.00	.17	.22	1.51	5.34	(.26) (.34)	1,557
Bad deal	.25	.23	.24	1.37	4.87	(.34) (.55)	1,553
4 item Index	.19	.22	.14	0.95	4.27	.45	1,551
3 item Index	.25	.24	.17	1.23	5.53	.55	1,553
KN '10 BPAQ-SF							
Hit	.00	.21	.28	1.10	2.99	(.74) (--)	907
Argue	.20	.28	.28	0.69	2.50	(.69) (.64)	904
Temper	.20	.25	.26	0.90	3.03	(.67) (.59)	906
Raw deal	.20	.34	.30	0.47	2.16	(.74) (.73)	905
4 item Index	.25	.18	.23	1.32	4.08	.77	883
3 item Index	.26	.29	.23	0.64	2.88	.74	890

Note. *Mdn* = median, *M* = mean, *SD* = standard deviation, *Sk* = Skewness, *Ku* = Kurtosis, Cronbach's α . Subscale α 's with item excluded in parentheses, for the 4-item and then 3-item comparisons.

Distributions for individual items and indices are similar between the ANES '13 and KN '10 versions. When combining the items in an index, however, the four BPAQ-SF items in KN '10 hold together much better than the new ANES items. 'Hit' in the ANES drags down index reliability, suggesting it does not belong as part of the construct and may suffer the same confusion problems seen in the MTurk study. Dropping the new item from the index substantially increases index reliability in the ANES but dropping the equivalent BPAQ-SF item in the KN '10 study does not affect reliability that was already high. The item-rest correlation for ANES 'Hit' is .11 (.29 to .38 for the other three). The same for 'Hit' in the KN surveys is .51 (.51 to .64 for others).

Table 5 shows within-subjects descriptive statistics in the MTurk data. Results are akin to Table 4: distributions are similar across paired items and indices (perhaps warning against reading too much into this test). But the new ANES 'Hit' item is weakly related to the others with a .20 item-rest correlation (.36 to .48 others), and the BPAQ-SF 'Hit' item does far better: .50 (.47 to .63 others). Figure A1 in the Appendix shows distributions for each of the four 'Hit' measures.

In sum, the ANES '13 'Hit' item fails reliability tests much like the ANES version in the MTurk study. Given within-subjects problems seen in the MTurk study, these results suggest the ANES '13 'Hit' item is similarly compromised. Dropping the ANES 'Hit' item in the MTurk study increases reliability by about 6 points; doing the same in ANES '13 increases reliability about 10 points.³

³ Another piece of evidence in that direction is found in the means. The MTurk study showed that high-aggression individuals were much more likely to change their response to low-aggression on the ANES version than the reverse. This should lead to substantially lower means for the ANES version. We see a 17-point gap between 'Hit' items for the MTurk study where we know the problem exists, and we see a 4-point gap in the ANES '13 study. The latter could indicate that the lopsidedness of mistakes was less or that the

Overall, these comparisons suggest the new ANES ‘Hit’ item diverges from the trait aggression construct as a whole, reinforcing inconsistencies observed for the item in the MTurk study.

Table 5: Descriptive Statistics for Two Versions in the MTurk Study (Within-Subjects)

MTurk	<i>Mdn</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>Ku</i>	α	<i>N</i>
ANES Version							
Hit	.00	.12	.24	2.37	8.03	(.62) (–)	906
Argue	.25	.27	.21	0.71	3.59	(.47) (.52)	905
Temper	.00	.15	.20	1.33	4.77	(.39) (.39)	906
Bad deal	.25	.26	.23	1.01	3.80	(.47) (.63)	900
4 item Index	.19	.20	.15	0.95	3.88	.56	899
3 item Index	.17	.23	.16	0.91	4.08	.62	899
BPAQ-SF							
Hit	.20	.29	.31	0.72	2.32	(.71) (–)	920
Argue	.20	.30	.28	0.60	2.29	(.67) (.59)	919
Temper	.20	.26	.27	0.88	2.89	(.64) (.56)	919
Raw deal	.40	.41	.29	0.20	2.13	(.73) (.72)	920
4 item Index	.30	.32	.22	.63	3.05	.75	899
3 item Index	.27	.32	.22	0.57	2.98	.71	899

Note. *Mdn* = median, *M* = mean, *SD* = standard deviation, *Sk* = Skewness, *Ku* = Kurtosis, Cronbach’s α . Subscale α ’s with item excluded in parentheses, for the 4-item and then 3-item comparisons.⁴

Tests of Predictive Validity – State Violence & Political Violence

What about predictive validity for each item? Past research finds links between trait aggression and support for state violence (Kalmoe 2013). One KN study included questions about the death penalty, bombing Iran, and torture of terror suspects, as did the ANES 2012 time-series linked to ANES ’13.⁵ In the KN study, 4-point response items are coded 0 to 1, with 1 indicating strong support for violent options. ANES’13 included a 5-point death penalty item, a 3-point Iran response, and a 7-point torture item. I create additive indices for each (KN α = .69, ANES α = .52). Table 6 presents bivariate OLS models of support for state violence predicted by each trait aggression item or index. Results for each policy are shown in Table A2 of the Appendix.

In the KN study, each of the items has a relatively similar positive relationship with state violence attitudes. But in the ANES study, the problematic ‘Hit’ item is the only one that shows a clear relationship, and then it is less than half the size seen for the original item.⁶ This evidence indicates that while ‘Hit’ may not fit well in an index with other items, it may actually reflect part of the

numbers of mistakes overall was less. Other items also show lower means with the new ANES versions, suggesting another possible explanation.

⁴ In MTurk data, correlations between the BPAQ-SF 4-item index and the 3- and 4-item ANES indices in the MTurk data are .63 for both. However, the problems found with the ‘Hit’ item make inclusion unpalatable.

⁵ The ANES item asked “Do you favor or oppose the death penalty for persons convicted of murder?” and then followed up with strongly or not strongly. The KN study asked “Do you favor the death penalty for persons convicted of murder, or do you favor life imprisonment without parole instead?” with a 4-point response scale indicating strength. The KN wording probably reduces average death penalty support by explicitly suggesting an alternative, but it is not clear how the mean difference would affect correlations with trait aggression presented here.

⁶ Failure of the other items could also arise from slight question wording and response differences.

construct better than the rest, at least in predicting violent attitudes. This is especially surprising given its measurement problems. The other ANES aggression items perform poorly in comparison.

Table 6: Bivariate OLS Models Predicting State Violence Support (Between-Subjects)

Death Penalty	<i>b</i>	<i>s.e</i>	R^2	<i>N</i>
ANES '13				
Hit	.09	.04	.009	1,553
Argue	.04	.05	.001	1,556
Temper	-.01	.04	.000	1,555
Bad deal	-.03	.04	.001	1,551
4 item Index	.06	.06	.001	1,549
3 item Index	-.01	.05	.000	1,551
KN '10 BPAQ-SF				
Hit	.21	.06	.032	399
Argue	.13	.07	.013	399
Temper	.19	.06	.030	400
Raw deal	.16	.06	.022	400
4 item Index	.21	.07	.032	383
3 item Index	.19	.06	.026	387

Political Violence against Authorities (Within-Subjects)

The MTurk study did not ask state violence items, but it did have two items from a political violence against authorities scale (Kalmoe 2014), one of which was “Agree or Disagree? Some of the problems citizens have with our government could be solved with a few well-aimed bullets.” This euphemism for assassinating leaders was assessed on a 5-point agree/disagree scale. Table 7 provides a predictive within-subjects test with the ‘Bullets’ item.

Table 7: Bivariate OLS Models Predicting “Well-Aimed Bullets” Support in MTurk Study

MTurk	<i>b</i>	<i>s.e</i>	R^2	<i>N</i>
ANES Version				
Hit	.28	.03	.08	905
Argue	.27	.04	.06	904
Temper	.32	.04	.07	905
Bad deal	.26	.03	.07	899
4 item Index	.65	.05	.16	898
3 item Index	.50	.05	.12	898
BPAQ-SF				
Hit	.22	.02	.08	905
Argue	.19	.03	.05	904
Temper	.22	.03	.06	903
Raw deal	.15	.03	.04	905
4 item Index	.34	.03	.10	899
3 item Index	.29	.03	.08	900

Surprisingly, the results show all of the new ANES versions perform at least as well as the original BPAQ-SF items, individually and as indices, including the ‘Hit’ item. This is reassuring evidence for the predictive validity of the ANES version of these items, especially after the poor showing on state violence. Overall, these tests add inconsistent evidence about the predictive validity of new items.

Tests of Convergent & Discriminant Validity (Between-Subjects)

Finally, I present evidence testing convergent validity with sex, age, and education, and discriminant validity with partisanship. In general, women are slightly less aggressive than men, older people are less aggressive than young adults, educated people are less aggressive than the less-educated, and there is no relationship between aggression and partisanship (Kalmoe 2015). Table 8 shows Pearson’s correlations for items and indices in both nationally representative studies.

Table 8: Pearson’s Correlations for Convergent & Discriminant Validity (Between-Subjects)

	Male	Age (yrs.)	Education (4 cat.)	7-pt. Partisanship (Rep.)
ANES ‘13				
Hit	.09	-.04	-.17	-.03
Argue	.10	-.09	.05	-.06
Temper	.03	-.11	.00	-.05
Bad deal	.02	-.17	-.13	-.12
4 item Index	.09	-.16	-.12	-.10
3 item Index	.06	-.17	-.04	-.10
KN ‘10 BPAQ-SF				
Hit	.15	-.28	-.22	-.05
Argue	.16	-.22	-.08	-.01
Temper	.07	-.17	-.24	.00
Raw deal	.05	-.21	-.20	-.06
4 item Index	.13	-.26	-.24	-.03
3 item Index	.12	-.23	-.21	-.03

In most tests where relationships are expected for convergent validity, correlations appear similar but slightly stronger for the BPAQ-SF items than the ANES items. Notably, trait aggression does show a small but significant correlation with partisanship in the ANES data, indicating Republicans are slightly less aggressive. This holds up even after adding demographic controls (not shown), but is not replicated in the other data. Here, the ‘Hit’ ANES item generally looks like its KN counterpart (but weaker) except for the relationship with age that almost vanishes entirely.

Table 9 presents the within-subject convergent and discriminant validity in the diverse but non-probabilistic MTurk sample for sex, age, education, and partisanship. Results appear substantially different for convergent validity by sex, age, and perhaps education for the ‘Hit’ item. In contrast, the other items perform similarly in both versions. Overall, the ANES ‘Hit’ item continues to be distinctive, this time in tests for convergent validity, though differences are more pronounced in the less representative data than in the ANES’13 comparison.

Table 9: Correlations for Convergent & Discriminant Validity in MTurk (Within-Subjects)

MTurk	Male	Age (6 cat.)	Education (4 cat.)	7-pt. Partisanship (Rep.)
ANES Version				
Hit	.06	-.08	-.06	-.06
Argue	.08	-.12	-.04	-.06
Temper	.02	-.09	-.02	.02
Bad deal	.00	-.15	-.13	-.07
4 item Index	.04	-.17	-.10	-.07
3 item Index	.02	-.16	-.09	-.05
BPAQ-SF				
Hit	.21	-.21	-.11	.01
Argue	.09	-.17	-.03	-.10
Temper	.05	-.13	-.03	-.02
Raw deal	.00	-.13	-.11	-.05
4 item Index	.12	-.21	-.09	-.05
3 item Index	.06	-.18	-.07	-.07

Revising the Physical Aggression Item

Given the apparent problems in the physical aggression item, how might it be improved to provide a better functioning 4-item battery? Here is one sound candidate, suggested by Ted Brader:

“Given enough provocation, how likely are you to hit someone? Very likely, somewhat likely, somewhat unlikely, or very unlikely?”

This ‘Hit’ item doesn’t require the mental somersaults of the ANES’13 version, and it adheres more closely to the language of the original BPAQ-SF item. Future surveys administering the ANES four-item battery should strongly consider this revision.

What Does Trait Aggression Predict in the ANES?

The 2013 ANES trait aggression battery was meant to show its role in predicting a wide variety of political attitudes and behaviors. So what does it do? Here I present a brief summary of initial tests. In each case, I estimate OLS models including the 3-item ANES trait aggression index and controls for partisanship, race, education, age, sex. The latter three are important demographic predictors of trait aggression as well, and all five are generally predictive of outcomes assessed here.⁷

People high in trait aggression are especially argumentative with others in everyday life. The verbal aggression item is meant to tap argumentativeness, but the whole trait aggression construct is implicated. Politics provides many opportunities for people to argue, often in uncivil ways, so aggression people may be more likely to engage in argumentative political participation. And they do. Trait aggression independently and significantly predicts a wide range of these expressions,

⁷ In each case, trait aggression in 2013 is used to “predict” outcomes in 2012. This seems backwards. The reader might be somewhat reassured by the fact that trait aggression is a stable personality trait across lifespans and is therefore unlikely to be easily moved by factors in the particular domain of politics.

including persuading others how to vote, signing petitions, and calling in to political radio or TV programs on politics. Results are even stronger for the verbal aggression item, including posting about politics on social media platforms and contacting members of Congress to voice an opinion. These behaviors are vital components of democratic participation and government responsiveness.

Trait aggression also indicates a general inclination to see oneself as continually aggrieved by a hostile and treacherous world. Indeed, an ANES item on social trust is strongly related to the trait aggression index and its hostility subcomponent. Similarly, I find aggressive individuals are significantly more likely to say they can't trust Washington to do what is right and that many people in government are corrupt. Likewise, aggressive people are more likely to say they have no say in government and public officials don't care about them. In short, trait aggression predicts negative evaluations of external efficacy and trust in government, both important for attitudes and behaviors.

Of course, the null results for state violence attitudes are unexpected and worrisome given consistent findings in the KN study and several psychology studies with convenient samples. Other plausible relationships are statistically null in ANES data, including expressions of anger toward candidates, negative retrospective evaluations, and dislike for parties and candidates. It is unclear whether these indicate no relationship, or whether tests are weakened by the new ANES items.

These findings are important because they identify a political role for a deep-rooted factor driving human behavior, not just another political correlate. Trait aggression is stable personality factor grounded in biology and long-term socialization. It routinely governs our daily orientations and behaviors, which is why citizens continue to employ it when it comes to politics. These relationships are novel and consequential.⁸

Conclusion

Here, I compared several measurement tests for new trait aggression items in the ANES 2013 Recontact Study to the original validated items in two similar representative samples and in within-subjects tests with both versions in a diverse convenience sample. I also described promising evidence of trait aggression's role in a broad set of important political attitudes and behaviors.

The results leave us with a puzzle. On one hand, the new physical aggression item in the ANES fails to correlate with its original wording in within-subjects comparisons, it fails to relate closely with the other items in the construct, and it shows weaker relations with convergent variables. Most damningly, the new item is wildly inconsistent with the BPAQ-SF item in a within-subjects comparison. In contrast, the other three new ANES trait aggression items pass each of these tests.

However, the new physical aggression item performs better than the others in predictive validity tests related to violent political attitudes. The other three items correlate with one form of political violence, as expected, but not with the other, diverging from the BPAQ-SF items' performance.

⁸ Trait aggression is distinct from Big Five personality traits, correlating with several of those dimensions to about the same degree that those independent dimensions correspond with each other. The 2013 ANES included TIPI items measuring the Big Five. Adding those factors as controls does little to diminish the size of effects described in this section, though the added co-linearity reduces statistical precision.

Ultimately, I recommend against including the new ANES ‘Hit’ item in an index of trait aggression. Future uses should consider the recommended revision to ‘Hit’. The remaining three items correspond well with their original counterparts, they are similar in distribution to the originals, they show solid convergent and discriminant validity, and they form a reliable index. The one note of concern comes from their mixed performance on predictive validity for violent political attitudes.

Of course, the measurement choice could vary for other scholars depending on their purposes and their preferences for weighting the mixed evidence here.

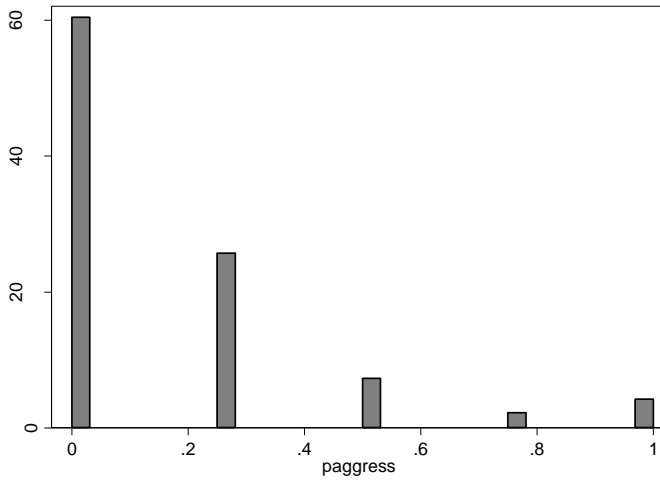
References

- Anderson, C. A., & Bushman, B. J. 2002. Human aggression. *Annual Review of Psychology*, 53, 27-51.
- Berinsky, A., Huber, G., & Lenz, G. 2012. Evaluating online labor markets for experimental research: Amazon’s Mechanical Turk. *Political Analysis*, 20, 351-368.
- Buss, A. H., & Durkee, A. 1957. An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology*, 21, 343–349.
- Buss, A. H., & Perry, M. 1992. The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452-59.
- Bryant, F. B., & Smith, B. D. 2001. Refining the architecture of aggression: A measurement model for the Buss-Perry Aggression Questionnaire. *Journal of Research in Personality*, 35, 138-167.
- Kalmoe, N.P. 2012. Aggressive personality & politics. ANES Online Commons proposal, EGSS4.
- Kalmoe, N.P. 2013. From fistfights to firefights: Trait aggression and support for state violence. *Political Behavior*, 35(2), 311-330.
- Kalmoe, N.P. 2014. Fueling the fire: Violent metaphors, trait aggression, and support for political violence. *Political Communication*, 31(4), 545-563.
- Kalmoe, N.P. 2015. Trait aggression in two representative U.S. surveys: Testing the generalizability of college samples. *Aggressive Behavior*, 41(2), 171-188.
- Marshall, M. A., & Brown, J. D. 2006. Trait aggressiveness and situational provocation: A test of the “Traits as Situational Sensitivities” (TASS) model. *Personality and Social Psychology Bulletin*, 32, 1100-1113.

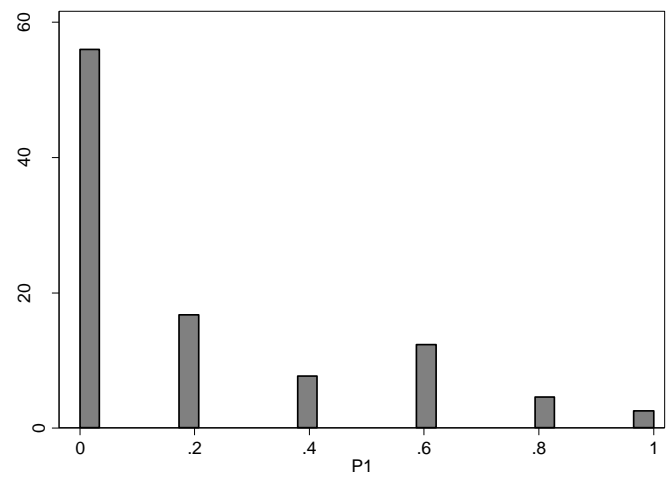
Appendix

Figure A1: Distributions of the 'Hit' Item

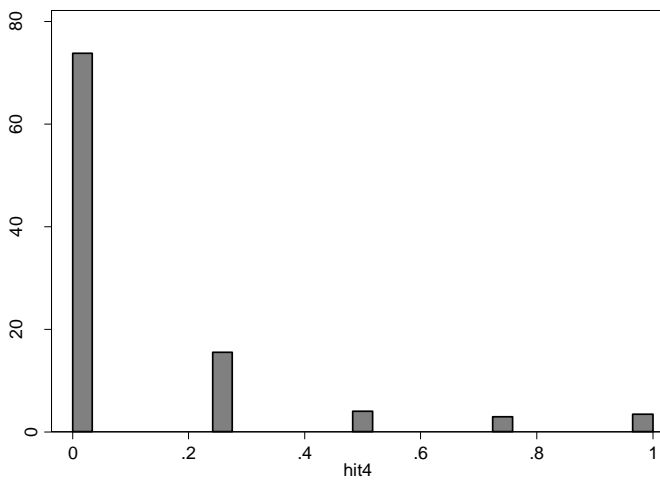
ANES: ANES version 'Hit'



KN: BPAQ-SF version 'Hit'



MTurk: ANES version 'Hit'



MTurk: BPAQ-SF version 'Hit'

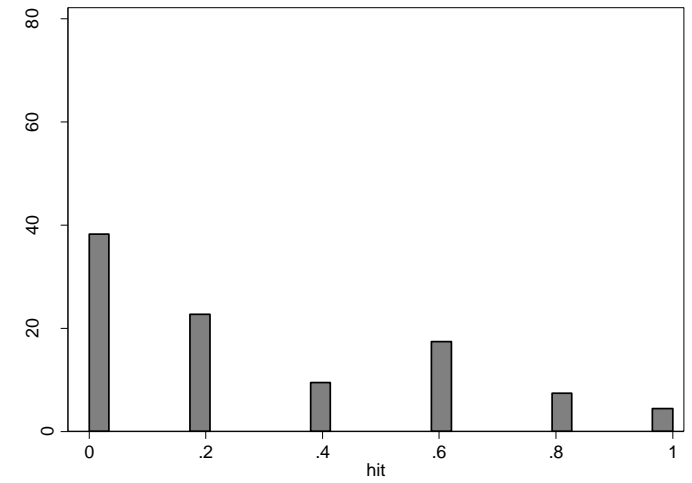


Table A1: Within-Subjects Responses to the Item Pairs in the MTurk Study

Raw deal BPAQ-SF	ANES version					Total
	0	0.25	0.5	0.75	1	
0	73.29	26.09	0.62	0	0	100
0.2	43.94	52.02	3.03	0.51	0.51	100
0.4	12.1	71.97	12.1	3.18	0.64	100
0.6	2.87	74.59	15.16	6.97	0.41	100
0.8	1.18	28.24	29.41	38.82	2.35	100
1	20.37	9.26	9.26	44.44	16.67	100

Temper BPAQ-SF	ANES version					Total
	0	0.25	0.5	0.75	1	
0	90.55	7.93	0.3	0.61	0.61	100
0.2	57.41	37.64	4.56	0.38	0	100
0.4	20	68.8	11.2	0	0	100
0.6	5.93	54.24	37.29	2.54	0	100
0.8	16.67	18.75	52.08	12.5	0	100
1	66.67	0	9.52	4.76	19.05	100

Argue BPAQ-SF	ANES version					Total
	0	0.25	0.5	0.75	1	
0	51.85	43.77	3.03	1.35	0	100
0.2	20.27	70.72	6.76	2.25	0	100
0.4	1.48	64.44	30.37	3.7	0	100
0.6	2.48	39.13	50.31	6.83	1.24	100
0.8	7.58	27.27	39.39	24.24	1.52	100
1	22.73	27.27	9.09	18.18	22.73	100

Table A2: Bivariate OLS Models Predicting Death Penalty Support

Death Penalty	<i>b</i>	<i>s.e</i>	R^2	<i>N</i>
ANES '13				
Hit	.13	.05	.009	1,546
Argue	.07	.06	.001	1,549
Temper	-.02	.06	.000	1,548
Bad deal	.02	.05	.000	1,544
4 item Index	.15	.08	.004	1,542
3 item Index	.04	.08	.000	1,544
KN '10 BPAQ-SF				
Hit	.17	.07	.016	401
Argue	.12	.08	.009	399
Temper	.16	.08	.013	403
Raw deal	.10	.07	.007	399
4 item Index	.23	.09	.020	388
3 item Index	.21	.09	.017	392

Bivariate OLS Models Predicting Support for Bombing Iran

Bomb Iran	<i>b</i>	<i>s.e</i>	R^2	<i>N</i>
ANES '13				
Hit	.02	.05	.000	1,492
Argue	.06	.07	.001	1,493
Temper	.02	.06	.000	1,492
Bad deal	.01	.06	.000	1,489
4 item Index	.06	.10	.000	1,489
3 item Index	.05	.09	.000	1,489
KN '10 BPAQ-SF				
Hit	.11	.07	.009	398
Argue	.15	.06	.017	369
Temper	.13	.07	.011	400
Raw deal	.03	.06	.001	396
4 item Index	.17	.08	.013	385
3 item Index	.16	.07	.013	389

Bivariate OLS Models Predicting Support for Torture

Torture	<i>b</i>	<i>s.e</i>	R^2	<i>N</i>
ANES '13				
Hit	.09	.04	.005	1,496
Argue	.03	.06	.001	1,499
Temper	-.01	.05	.000	1,498
Bad deal	-.10	.04	.007	1,494
4 item Index	.00	.07	.000	1,492
3 item Index	-.06	.06	.001	1,494
KN '10 BPAQ-SF				
Hit	.32	.08	.048	402
Argue	.15	.09	.010	402
Temper	.17	.08	.014	403
Raw deal	.20	.07	.022	403
4 item Index	.27	.09	.030	386
3 item Index	.21	.08	.020	390