


Summer 1933

## Measuring the Severity of the Third Degree

Herman C. Beyle

Spencer D. Parratt

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

---

### Recommended Citation

Herman C. Beyle, Spencer D. Parratt, Measuring the Severity of the Third Degree, 24 Am. Inst. Crim. L. & Criminology 485 (1933-1934)

This Criminology is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern University School of Law Scholarly Commons.

with  $\frac{1}{8}$  inch metal below the slot. The shackle broke at 2150 pounds, with  $\frac{1}{4}$  inch metal below; another broke at 2150 pounds. A third shackle was made of the Bohannon type with a slot; this broke at 4130 pounds. The brass showed about 66 per cent of copper.

### *Tests Upon Staples*

To know the strength of a padlock but half answers the question of security—the other half being the strength of the fitting into which it is placed. Staples screwed to birch wood with four  $\frac{3}{4}$  inch No. 8 (0.163 inch diameter) screws, required 660 pounds to pull them off. Those fastened with four  $1\frac{1}{2}$  inch No. 10 (0.189) screws required 1670 pounds. A steel staple  $\frac{3}{16}$  inch diameter broke at 1440 pounds while one  $\frac{7}{32}$  inch diameter broke at 2600 pounds. A staple of this latter size driven into birch and clinched pulled out at 1230 pounds.

If it be a question of maximum security, choose a heavy iron hasp with at least  $\frac{7}{32}$  inch case hardened staple, fastened with four 2 inch No. 12 (0.216) screws or better bolts and nuts, and the largest size iron or bronze padlock with a case hardened shackle and five or six tumblers.

The authors wish to express their obligation to the late Professor H. W. Hayward of the Mechanical Engineering Department for suggestions as to engineering details.

---

## MEASURING THE SEVERITY OF THE THIRD DEGREE

---

HERMAN C. BEYLE AND SPENCER D. PARRATT<sup>1</sup>

---

The National Commission on Law Observance and Enforcement, after making all deductions for the inherent uncertainties of the subject matter, presented the following propositions as established factors in American law enforcement procedure: (1) the inflicting of pain, physical or mental, to extract confessions or statements from witnesses or suspected criminals is widespread throughout the country; (2) physical brutality is extensively practiced; (3) the method most commonly employed by law enforcing officials in attempting to extract confessions or statements is protracted questioning, at times by relays of questioners, so protracted that the prisoner's energies are spent and his powers of resistance overcome; (4) methods of

---

<sup>1</sup>Syracuse University, Syracuse, N. Y.

intimidation adjusted to the age or mentality of the victim are used; and (5) prolonged illegal detention is a common practice.<sup>2</sup> Particularly significant is the fact that such practices are contrary to the law, with the necessary consequence that officers engaged in them are criminals.

It is beyond doubt that the practices of the so-called "third degree" are shocking in character and extent, and violative of American traditions and institutions. In fact, few fundamental principles are more firmly grounded in our legal system than those providing that no person shall be convicted of a crime through self-incrimination, and that involuntary confessions shall not be admitted as evidence in a criminal trial. Both of these rules have a long history and their final incorporation into the American system of law is a landmark in the development of personal security against arbitrary governmental interference. The use of torture to exact confessions was common in England until the middle of the seventeenth century, and no scruple existed against using such confessions as evidence in criminal trials. The traditional horrors of the Star Chamber and the trials during the inquisition echo down the corridors of history. Experience has shown that justice and violence in its execution are disharmonious in a law enforcement system. From the viewpoint of the development of free government, the practices of third degree must be considered somewhat anomalous; as an anachronism that has continued its existence or raised its head in recent times because of new factors present in modern American society.

The American Bar Association Committee on Lawless Enforcement summarizes the effect of third degree practices upon the relationship between the public and the law enforcing officials as follows:

"The use of the third degree is obnoxious because it is secret; because the prisoner is wholly unrepresented; because there is present no neutral, impartial authority to determine questions between the police and the prisoner; because there is no limit to the range of the inquisition, nor to the pressure that may put upon the prisoner. . . . Probably the third degree has been a chief factor in bringing about the present attitude of hostility on the part of a considerable portion of the population toward the police and the very general failure of a large element of the people to aid or cooperate with the police in maintaining law and order."<sup>3</sup>

As the above quotation indicates, there is, and must be, an intimate relationship between police methods and public attitudes. At present,

---

<sup>2</sup>Publication No. 11, *Report on Lawlessness in Law Enforcement*, United States Government Printing Office, Washington, D. C., 1931.

<sup>3</sup>Quoted from Publication No. 11, *ibid.*, p. 191.

one might indulge in the presumption that American public opinion does not sanction the third degree methods of the law enforcement officers. The authors of the *Report on Lawlessness in Law Enforcement* accept, without question, that American public opinion not only does not sanction the use of third degree methods, but that the very existence of such methods in our law enforcing process is indicative of public ignorance of the nature and extent of such abuses.<sup>4</sup>

A fundamental question might be raised as to whether or not the American public is so ignorant of police practices as the authors of the above report believe. Instances of notorious police practices are given in the news from time to time. Might not the conjecture be made that the American people are not entirely ignorant of such practices, but that they actually tolerate them as an integral part of the system as it exists?

An attempt to determine the status of public opinion relative to tolerance or intolerance of third degree practices presents a preliminary problem of clarifying the meaning and subject content of the term "third degree." Careful consideration of the problem indicates that the term covers a multitude of relatively specific methods of bringing coercion to bear upon suspects. Preliminary inspection of a list of situations in which coercion is applied will indicate that a range in severity or intensity of coercion exists. Third degree becomes, not a single kind of action varying in severity, but a variety of actions presenting severity of great diversity.

Whereas one observer might hesitate to condemn trick questioning and emotional appeals to "heart interests," another might consider such action despicable and beneath contempt of official police practice. Considerable variety of opinion might be found concerning the efficacy of the use of threats of harm, either physical or injury to reputation. But third degree practices shade gradually from the slightest amount of coercion, or even promise of favoritism, to the exercise of the most highly developed methods of torture that occasionally bring death to unhappy victims. Public opinion may or may not approve of all of the specific phases of third degree practices. There may be a determinable classification of specific practices which find favor, others which are tolerable, with still others beyond the pale of toleration of any sort. But before such a conclusion might be made some method needs must be devised to measure severity of practices.

A host of problems arise to confront one who seeks to determine

---

<sup>4</sup>*Ibid.*, pp. 191-192.

the status of public opinion relative to third degree practices. Is there a different toleration point among different segments of the public? Do policemen, criminals, and laymen consider the same practices equally severe? Is the toleration point flexible, depending upon the characteristics of the suspect, or the type of crime of which he is accused? Is there one standard when the suspect is Mr. John Citizen, with no distinguishing characteristics, and another standard when the suspect is a negro, a gangster, a petty thief with a repititious criminal record, or a "red?" Is a suspect facing the charge of murder to be allowed less personal protection against illegal discretion and third degree abuses than a person charged with petty larceny? Is the standard of toleration conditioned by living in large cities or in rural areas? These and other basic questions present themselves.

The potential significance of answers to the foregoing queries is particularly emphasized when one considers that they would materially aid in clarifying the apparent differential between official practice and legal standards. It might be better to have a standard fixed by law that could find acceptance in practice, than one so out of sympathy with existing methods of law enforcement that the co-operative relationship of police and public is jeopardized. Better to make the law compatible with a standard that is enforceable than to permit the present realm of illegal discretion without responsibility. Or, perhaps, an objective determination of the status of the relationship of public opinion and third degree methods will substantiate the presumption that public opinion does not sanction, nor tolerate the practices complained of in the *Report on Lawlessness in Law Enforcement*. It might indicate a need for publicity and for bringing the public to a realization of the existence of a practice of which it does not approve, and which is within its power to destroy.

The foregoing comment affords the setting, not the task of this article. Were one to attempt a campaign of fact-seeking that would answer queries as to what is public belief as to official practice and what is the point dividing approval and disapproval of practice, in different areas of the United States, in different times, among different groups, and as respects different offenses and offenders, there would be need of some means of recording observation and opinion in terms of known precision, tested ambiguity, and justifiable comparability. The usual schedules and questionnaires calling for free answer produce returns couched in general language, language that is highly ambiguous and frequently loaded with prejudice and bias. Some tested instrument is needed that would open a door

to clarification of standards. The purpose of this article, then, is that of presenting such an instrument, and its testing, in the form of a scale for the measurement of severity of third degree practices, applicable to the marking of practices either as observed or as approved. Projects for the application of this scale are under way; but the widespread interest of inquirers in the subject matter to which it can be applied would seem to justify its present release as a severely tested instrument for possible use by others. Indeed, if it be acceptable for systematic application, such application would need to be made by many. Such use of the scale is invited.

### *The Scale for Measuring Severity*

This scale presents a marked departure from previous scales similarly constructed. Previous scales, produced by psychophysical methods, as adapted to the use of scale construction by Professor Thurstone,<sup>5</sup> have been affect scales.<sup>6</sup> They have been designed to measure the degree of unfavorable-favorable attitude toward some social object by endorsement of statements of opinion evidencing some tested meaning as respects the degree of disfavor or favor. This scale uses as landmarks brief descriptions of specific behaviors—police practices in obtaining confessions, tested as respects a single characteristic—the severity of the several practices. It is not an affect scale, then, but a scale for the measurement of a characteristic of related behaviors. By manipulation of accompanying directions, however, it may be adapted to various uses. Thus, if the assistant investigator is asked to mark the behaviors he has observed, the scale becomes an instrument for recording observations as to actual third degree practice. If the subject is asked to mark the practices which he approves, the scale becomes an instrument for determining the limits of favorable and unfavorable affect. Combination of uses, moreover, would enable the investigator to assemble data in comparable terms, whether they be observations as to what practices exist, what practices are thought to exist, or what practices are approved.

The numerical continuum upon which the behavior landmarks were intended to be located extends from zero (no severity at all) to ten (the maximum of severity). Inasmuch as the raters, how-

<sup>5</sup>L. L. Thurstone and E. J. Chave, *The Measurement of Attitude*. (Chicago, 1929.)

<sup>6</sup>See "The Measurement of Social Attitudes." Series of scales edited by L. L. Thurstone, The University of Chicago Press. See also, H. C. Beyle, "A Scale for the Measurement of Attitude Toward Candidates for Elective Governmental Office," *American Political Science Review*, Vol. XXVI, No. 3 (June, 1932), pp. 527-544.

ever, considered some of the behaviors to be so lacking in severity as to evidence the opposite of that characteristic, the landmarks of the completed scale have been actually located along the numerical continuum extending from  $-0.5$  through  $0.0$  to  $+9.5$ . These landmark statements present a series of specific acts beginning with those characterized by something near no severity at all and ranging in tested progression to the highest possible severity—death.

In collecting a list of tentative statements to serve as possible landmarks on a linear scale representing the selected behavior variable, the considerable literature on law enforcement was reviewed. Particularly fruitful, however, was the report of the Wickersham Commission. From these sources, brief descriptions of distinct police practices were assembled, comprising a first tentative list of seventy-two statements. In accepting these statements for testing we were not concerned with prejudgment as to whether the described behaviors were generally used. We were only concerned that they should be brief and graphic descriptions of a police practice in obtaining confessions; that they should run the gamut of severity with as few gaps as possible; and that they should be as representative as possible of all modes of behavior characterized by more or less of severity. It may be of interest to note that the alleged practices in the recent Hyman Stark case, that of standing upon a prone suspect with one foot upon his stomach and the other upon his throat and of rocking back and forth to force a confession, was discarded as being either a practice insufficiently plausible to receive the serious consideration of raters or one likely to shock their sensibilities. The elements of that practice, however, were included. They would be statements 9, 33, and 5 of the final scale, located between  $+8.4$  and  $+9.5$  of the numerical continuum. Thus, in terms of behavior elements the list is thought to be sufficiently varied to permit of combinations which would adequately describe almost all inquisitorial situations to be found in American police practice.

The description of a police practice involved several difficulties. To be useful in a scale intended for wide application, the description must not make use of colloquial expression. It must be graphic. And it must be brief. In the case of description that is colloquial or obscure, raters will fill in the picture for themselves in different ways, just as subjects will do later, and the result will be rating evidencing extreme ambiguity. Also, in the case of lengthy description, one rater will attend to one feature in the description, another to some other feature, and again the resultant ratings will warn of ambiguity.

The first tentative list of seventy-two statements, therefore, were presented for the comparative judgments of fifty raters, and their returns were taken as indications of needed revision. Such revision yielded a second tentative list of eighty-one statements. These were the statements which were tested for inclusion in the final scale.

This tentative list of eighty-one statements was presented to two hundred and six judges for rating. Of these, fifty-six were New York State troopers from Companies D and C at Oneida and at Sydney, New York.<sup>7</sup> Fifty were prisoners in the state prison at Auburn, New York, or the Utah state penitentiary at Salt Lake city.<sup>8</sup> The remaining one hundred were citizens of normal status, "laymen" as we will call them, half of them being interviewed in Syracuse, New York,<sup>9</sup> and half being interviewed in Utah, Illinois, Ohio,<sup>10</sup> Minnesota, or Oklahoma—anywhere but in New York state.

Use of the Thurstone technique for scale building and testing does not require that one use as raters a sample which will be representative of those who may later respond to the instrument; but only that the raters be compared of groups as diverse as possible. Should diverse groups be able to agree as to the scale position and the ambiguity of statement landmarks, one would have considerable confidence that the instrument has as great stability as is possible. Certainly troopers and prisoners are diverse groups; and peculiarly diverse as respects the third degree which affects those of each group so differently and so vitally. Furthermore, the process of rating by these diverse groups is not one that raises much doubt as to the insight or candor involved. If the laymen were without actual experience with inquisitorial practices of the police, the prisoners and the troopers certainly were not. And in rating, the troopers were not asked to say what practices they used, but merely how severe they considered the stated, possible practices of other police forces to be. Also, the prisoners were not asked to say how severe the practices

---

<sup>7</sup>Acknowledgment is made of the courtesy and careful rating of the troopers, and particularly of the kind consideration of Professor Raymond Moley, Superintendent John A. Warner, Deputy Superintendent George P. Dutton, Captains Stephen McGrath and Daniel E. Fox, Lieutenant Ronan, and Corporal Ray L. Gallon, who facilitated the collection of ratings.

<sup>8</sup>Acknowledgment is made of the attention and earnest rating of the prisoners, and particularly of the kind assistance of Professor Thomas R. Fisher, Mr. D. W. Parratt, Warden Joseph H. Brophy, and Chief Clerk R. R. Westover, who aided greatly in the collection of ratings.

<sup>9</sup>Acknowledgment is made of the kindness of the members of the classes in "Law Enforcement" and in "Public Opinion" at Syracuse University and their friends who made or secured ratings.

<sup>10</sup>Acknowledgment is made of the help of Mr. Elmer E. Hilpert and others who made or secured ratings.



were they had encountered, but only how severe they considered the several assumed practices. The chief requirement is that the raters understand what it is that they are rating, and that they carefully pass comparative judgment upon all of the practices presented to them. This requirement was well met because of interest in the subject matter. The findings to follow should give considerable confidence in the carefulness of the ratings and in the stability of the results.

On the basis of this rating and subsequent testing, fifty of the eighty-one statements were found acceptable for inclusion in the final instrument. These are the statements listed below. The list comprises the tested landmarks of the third degree severity scale. The numbers preceding the statements are arbitrary, serving the purpose of ready identification. They serve also to provide a random order of the statements. In the parentheses following each statement are two figures. The first is the scale position of the statement as determined by the two hundred and six raters, and the second is the coefficient of ambiguity of the statement, or the semi-interquartile range of the ratings of the two hundred and six raters. To save confusion, the first will be designated as "S. P." and the second as "A".

1. To make him confess, a suspect is slapped in the face by a policeman. (S. P., 4.4; A., 3.8.)
2. Police ask detailed questions, tricking suspect into confessing crimes. (S. P., 0.8; A., 2.6.)
3. At police headquarters the police lock a suspect in a cell and take turns questioning him, not letting him have any rest or sleep for three consecutive days and nights. (S. P., 7.4; A., 3.3.)
4. Police question a suspect at police headquarters for an hour. (S. P., 0.3; A., 2.0.)
5. Police kill suspect by brutal treatment when he refuses to answer questions. (S. P., 9.5; A., 1.0.)
6. Policeman shakes his doubled fist under a suspect's nose and tells him to confess. (S. P., 1.7; A., 2.7.)
7. Police beat a suspect for five minutes with a rubber hose to make him confess. (S. P., 7.1; A., 3.4.)
8. Police lock a suspect in jail all one night without permitting him to get in touch with his friends or lawyer to try to obtain a writ of habeas corpus. (S. P., 2.9; A., 3.5.)
9. Police kick a suspect in the stomach to cause him to confess. (S. P., 8.4; A., 2.4.)
10. Police hold a suspect in a miserably cold room for one night to make him confess. (S. P., 5.2; A., 3.6.)
11. Police trick suspect into confessing by pretending to be friendly. (S. P., 0.5; A., 2.2.)

12. Police knock suspect unconscious when he refuses to answer questions. (S. P., 8.7; A., 2.0.)
13. Police strike a suspect with a heavy book to make him confess. (S. P., 5.9; A., 3.5.)
14. When two suspected accomplices in the same crime are brought to police headquarters they are placed in separate cells, the police telling each suspect that the other has confessed and implicated him, without regard to the truth of such statements. (S. P., 1.9; A., 4.0.)
15. Police twist a suspect's arms and legs to make him confess. (S. P., 7.6; A., 3.0.)
16. Police place a suspect in a back room at police headquarters and brandish whips and clubs before suspect, threatening to strike him if he will not confess. (S. P., 3.9; A., 4.0.)
17. Police push lighted cigars against the skin of a suspect to make him confess. (S. P., 8.0; A., 2.3.)
18. Police pretend they know more than they do about a crime to cause a suspect to tell his story and incriminate himself. (S. P., 0.7; A., 2.6.)
19. Police strike a suspect in the body with fists and tell him that he will get more if he refuses to confess. (S. P., 6.6; A., 3.6.)
20. Police tell a suspect that he shall be deprived of food until he is willing to confess. (S. P., 2.3; A., 3.7.)
21. Police beat a suspect for an hour with a rubber hose to make him confess. (S. P., 8.5; A., 2.2.)
22. When the police question a suspect at police headquarters they first tell the suspect that his failure to answer questions will show that he is guilty. (S. P., 1.4; A., 2.5.)
23. Police lock a suspect in jail all day without food and then tell him he can have food as soon as he will confess. (S.P., 4.7; A. 3.7.)
24. Police lock a suspect in a dark damp cell without any place to sit or lie down for three days to make him confess. (S. P., 7.9; A., 2.6.)
25. Police are careful to explain to suspect that whatever he says may be used against him. (S. P., —0.3; A., 1.4.)
26. Police place a suspect in a back room at police headquarters with a bright blinding light before his eyes and force him to sit facing the light for an hour while they attempt to brow-beat him into confessing. (S. P., 6.5; A., 3.8.)
27. Police beat a suspect until he requires medical treatment and then refuse medical attention until the suspect confesses. (S. P., 9.0; A., 1.7.)
28. Police give suspect a light tap with a club at time of arrest so that he will know what will happen if he refuses to talk. (S. P., 3.5; A., 3.7.)
29. To make a suspect confess, the police deprive him of all food, except water, for three consecutive days. (S. P., 7.4; A., 3.2.)
30. Police promise leniency to a suspect to cause him to tell all he knows about suspected crimes. (S. P., 0.8; A., 2.4.)
31. To wear a suspect's resistance down and cause confession, the police

- lock him up for the night, but wake him every hour to ask him insulting and accusing questions, causing the suspect to answer before he can collect his thoughts. (S. P., 5.1; A., 4.0.)
32. Police give a suspect a cigar to make him talk about crimes that he is suspected of having been implicated in. (S. P., 0.0; A., 1.6.)
  33. The police strike a suspect in the throat until blood spurts from his mouth to make him confess. (S. P., 8.8; A., 1.6.)
  34. Police tell a suspect that they will arrest his pals unless he makes a confession. (S. P., 1.1; A., 2.6.)
  35. Police lock a suspect in a miserably cold room and keep him there for three days to make him confess. (S. P., 7.4; A., 3.2.)
  36. Police ask a suspect leading questions to trick him into confessing crimes. (S. P., 0.6; A., 2.6.)
  37. In examining a suspect at police headquarters, the police tell the suspect dreadful stories of what happens to persons who refuse to confess. (S. P., 2.0; A., 3.4.)
  38. Police strip a suspect of his clothes at police headquarters and strike him one smart blow with a rubber hose, threatening more unless suspect confesses. (S. P., 6.6; A., 3.8.)
  39. Police trick suspect into confessing by threatening to put him under an examination by means of a machine to detect lies. (S. P., 1.3; A., 2.9.)
  40. During an examination at police headquarters a suspect is told that unless he confesses the police will tell his friends and relatives that he is a common jailbird. (S. P., 1.9; A., 3.3.)
  41. To make him confess, police strike a suspect in the face hard enough to cause blood to flow from the nose. (S. P., 7.5; A., 2.9.)
  42. Police swear at suspect. (S. P., 1.2; A., 2.4.)
  43. During an examination at police headquarters a suspect is told that unless he answers all questions, the police will see that his employer knows he is a crook. (S. P., 2.2; A., 3.5.)
  44. To make him confess, police strike a suspect in the mouth hard enough to loosen teeth. (S. P., 7.8; A., 2.8.)
  45. Police are careful not to frighten suspect. (S. P., -0.5; A., 1.1.)
  46. Police deprive a suspect of food for one meal to force him into confessing. (S. P., 1.8; A., 3.0.)
  47. Police refuse to give a suspect, who is known to be a drug addict, any drugs until he is almost crazy, telling him that he can have drugs when he confesses. (S. P., 7.3; A., 3.3.)
  48. Police tell a suspect that they know he is guilty and that it will be for his own good to confess. (S. P., 0.8; A., 2.5.)
  49. The police strike a suspect a blow in the face with a doubled fist and tell him he will get more if he refuses to confess. (S. P., 6.5; A., 3.7.)
  50. When an arrest is made the arresting officer tells the suspect that he will come to no harm if he will answer all questions asked, but that he will get hurt if he refuses to talk. (S. P., 1.3; A., 2.4.)

Nothing would be added to the meaning of the scale if arbitrary points at which one passes from moderate to extreme severity were

pointed out. Indeed the difference is one of degree rather than one of kind. The statements themselves give meaning to the numerical description of the degree of severity. Of course, the several statements have characteristics or attributes other than that of varying severity. They describe various representative modes of more or less severe treatment, such as pressure without personal physical contact, or pressure through personal application of physical force. Some investigators may be interested in such qualitative designation of the patterns of police practices which may be described by means of the scale landmarks; and that use of the instrument is legitimate. But one should note that the numerical description of police practices by means of the scale relates to the single behavior variable, the more or less of severity.

### *Use of the Scale*

As has been said, the scale may be used in various ways by alteration of directions for marking. In the footnotes is a set of directions illustrative of what might be devised to accompany the scale.<sup>11</sup> Use of this set of directions would make the scale one for

<sup>11</sup>The directions might be somewhat as follows:

**DIRECTIONS:** Beginning on the next page is a list of statements. Each statement describes a particular police practice now in use somewhere in the United States. The popular name for these practices is the "third degree."

Since the various practices represent different degrees of severity ranging from no severity at all to extreme severity, it is conceivable that some may be approved while others may be disapproved. It is conceivable, too, that whether a practice is approved or not depends on the nature of the offense and the sort of person suspected of the offense. *You are asked to indicate which practices you approve and which you disapprove for use in dealing with certain offenses and offenders.* At the right of each statement are squares in which you may mark with a cross, thus X, to indicate your approval. Please do so in the manner indicated in the following example of possible marking as to a police practice.

Thus, if you approve of the following police practice in the case of an average citizen suspected either of murder or burglary, in the case of a gangster suspected of any one of the five specified offenses, and in the cases of a negro, a communist, or a foreigner suspected by any one of the offenses except that of bootlegging, you would so indicate by the following marking:

#### (EXAMPLE)

	Average Citizen	Gangster	Negro	Communist	Foreigner	
51. Police force a suspect to watch another suspect given a beating with a rubber hose to frighten him into confessing.	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Murder
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Burglary
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Embezzlement
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Assault and Battery
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Bootlegging

*Put a cross in a square to indicate your approval. Leave a square blank to indicate your disapproval.*

The statements describing police practices are not listed in order of their

the measurement of approval-disapproval of third degree police practices. By alteration of the directions, the scale could be converted into an instrument for recording belief as to the use of more or less severe third degree practices. The alteration needed would be such as would secure conformity with the changed request: "Put a cross in a square to indicate that you think a practice is used by the police of your locality. Leave a square blank to indicate that you think a practice is not used." By another change, the scale could be made an instrument for recording the observations of cooperating investigators as to actual third degree practice. The altered request to which the directions would need to conform for such use might be: "Put the appropriate check mark<sup>12</sup> in a square to indicate that you have the designated sort of evidence that a practice was used by the police of ..... during ..... Leave a square blank to indicate that you have no evidence that a practice was used during the period." Following any variety of directions, the statements of the scale could be set up for marking in some manner as illustrated in the accompanying footnote.<sup>13</sup> Needless to say, the directions and

severity, but are presented in random order. Do not alter any statement. If a statement is altered it will be ignored in scoring the result. The approval of an altered statement would not be comparable with approvals by others, and consequently would be valueless for scoring.

*Do not sign your name on this paper.* We wish your marking to be strictly secret. The information gained from the collection of markings will be used for scientific purposes only. Furthermore, this is *not a testing of your opinion*. People naturally differ as to what should and what should not be approved. Your judgment is sought because you have practical knowledge about police work and its needs.

<sup>12</sup>Different check marks could be employed to indicate different sorts of evidence, such as court records, reports in newspapers, admissions of police interviewed, and the like.

<sup>13</sup>Following any variety of directions, the statements of the scale could be set up in the following manner:

	Average	Citizen			
	Gangster	Negro	Communist	Foreigner	
1. To make him confess, a suspect is slapped in the face by a policeman.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Murder
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Burglary
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Embezzlement
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Assault and Battery
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Bootlegging
2. Police ask detailed questions, tricking suspect into confessing crimes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Murder
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Burglary
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Embezzlement
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Assault and Battery
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Bootlegging

the form of presentation of the statements may be further varied to meet the needs of other objects of inquiry. The obvious qualification, of course, is that the inquiry would need to involve in some manner the more or less of severity of third degree police practices.

In treating returns from the marked instrument, two distinct procedures may be followed. Mention is made of this matter for the reason that the second of the suggested procedures is one which has not usually been followed in the treatment of data gained from use of this sort of scale, and that it is one which permits of ready handling of a large number of returns. The usual procedure is to summarize the return from each individual subject in terms either of the median of the endorsed landmarks, or the range of endorsed landmarks. One could then summarize the returns for classes and sub-classes of individual subjects in terms of the distribution of individual scoring. In case an investigation contemplated thousands of returns instead of hundreds, another procedure which lends itself to aid from sorting machines may be used. By this method, the marked statement would be the item tabulated. Thus by classes or sub-classes of subjects, a distribution could be made in terms of the number or percent of a class marking each of the landmarks. The first of these procedures would probably give for summarization, curves which are symmetrical or only moderately asymmetrical in form. The second would probably yield curves which are variants of the asymmetrical form of distribution. Significance should not be attached to the form of resulting curve as being the criterion of institutional data as distinguished from "natural" data, as is suggested in some quarters, for it is method of treatment as much as nature of the data that determines the result.

In case the usual procedure is used of first determining the individual scoring, and then the scoring of classes of individuals, it is suggested that the upper limit of approved or observed landmarks be the principal scoring noted. In the case of this scale, it is limits of tolerance which are seemingly most important. However, in the event that the purposes of some investigation would be served by record of central tendencies of approvals by individuals, it is suggested that the mid-point of the range of approved or observed landmarks be employed rather than the median approval or observation. This

- |  |   |  |
|--|---|--|
| 3. At police headquarters the police lock a suspect in a cell and take turns questioning him, not letting him have any rest for three consecutive days and nights. | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> | Murder<br>Burglary<br>Embezzlement<br>Assault and Battery<br>Bootlegging |
|--|---|--|

suggestion arises from the nature of the distribution of the landmarks upon the continuum of the scale. The list of fifty statements contains a large proportion that distribute rather evenly over the scale. But it contains a few more than were needed to secure such representative distribution. It seemed desirable to provide a list sufficiently rich for purposes of description to enable one, by a combination of marking, to describe most inquisitorial situations likely to be encountered. To this end, needed statements were added to the list; and since there seems to be two main varieties of the third degree, pressure without personal physical contact, and pressure through personal application of physical force, the extra statements needed for rich description mass at two portions of the continuum, between  $+0.5$  and  $+2.3$  and between  $+7.1$  and  $+9.0$ . Use of the median marked statement would be affected by this massing of landmarks. Use of the mid-point of the range of marking would not.

### *Testing the Scale*

Naturally those who might be inclined to make use of this scale will wish to know more of the testing to which the list of landmarks was submitted. Probably no scale built with the Thurstone technique has been put to so severe a test as this one. Reliance upon the ratings of such diverse groups as troopers and prisoners, particularly with respect to a matter which would be likely to be viewed so differently by them, would appear to be as rigorous a test as could be devised. However, severe testing is particularly desirable in this instance. This is the first scale measuring a characteristic of related behaviors. Should it prove workable, it would point the way to similar provision of instruments for the development of inquiry respecting the many characteristics of the many behaviors that constitute government. As such a pioneering attempt, it affords a test upon a most difficult ground. Few political and governmental behaviors involve participants in as vital a clash as does the third degree. Many involve less. On one side in the third degree is the interest of freedom, possibly life, if the inquisition develops incriminating leads; while on the other is success or failure in a dangerous profession. It is important, therefore, to understand thoroughly just what the stability and reliability of the scale are.

One way of presenting the degree of stability of the statement scale position which was discovered in the treatment of the data secured from the ratings made by the troopers, the prisoners, and the laymen, is through the following Pearson correlation coefficients. The

items involved in each of those correlations are the scale positions of the fifty statements of the final instrument. The two series of values correlated in each instance are the scale positions assigned these statements by each of the two designated groups of raters, all series being classified by one-tenth of a degree intervals, and the continuum of the scale being considered to extend over the eleven degrees from  $-1.0$  to  $+10.0$ . The correlation between the scale positions assigned by the troopers ratings and those assigned by the prisoners ratings was  $+0.9773 \pm 0.0043$ .<sup>14</sup> The similar correlation showing the relation of troopers' assignments of scale positions and the assignment of scale positions by the laymen was  $+0.9968 \pm 0.0008$ . For the assignments by prisoners as compared with those by the laymen it was  $+0.9808 \pm 0.0037$ . For the assignments by the fifty laymen from New York state as compared with those by the fifty laymen from elsewhere than New York it was  $+0.9763 \pm 0.0045$ . Finally, the similar correlation of the assignments made by the laymen as compared with the assignments made by the troopers and prisoners combined was  $+0.9974 \pm 0.0005$ .

It is needless to point out that these correlations are unusually high for social data. They probably mean that the statements of the scale are descriptive of institutional behaviors which have rather stable meaning within the cultural area of which the New York state troopers, the New York and Utah prisoners, the New York laymen, and laymen selected elsewhere are a part. The laymen particularly, may not have known the exact description of the institution as it actually exists in a given locality, but they appeared to understand behavior elements of the institution when described to them. The remarkably high coefficients of correlation reported should give one considerable confidence in the stability of the tested landmarks of the third degree severity scale; at least, for use within the area of the American police system.

The correlations are surprisingly high. Probably very much higher than one might have dared to expect. That astonishment, however, must not be permitted to blind one to the limitations of the scale as needs to be shown in other ways. The probable error of the scale value of statement landmarks employed in the scales that Professor Thurstone has reported in full is 0.06 scale units in an instance where a scale continuum was considered to extend over

<sup>14</sup>The coefficients of these correlations are carried out to four places only for the reason that the coefficients are so high and the probable errors so small that it appears necessary to do so in order that one may note what the effect of subtracting the probable error from the coefficient might be.



eleven units.<sup>15</sup> The similar figure for the statements of the candidate scale was 0.10 scale units. As has been pointed out elsewhere, there may not be as great a difference in these two figures as might appear, since there was no elimination of so called "erratic" raters in the building of the latter scale.<sup>16</sup> Rather, the full possibility of ambiguity was sought to be revealed. Now in the case of the third degree severity scale, one encounters the possibility not only of erratic raters, but of vitally interested raters in the case of troopers and prisoners, and of raters who have had no direct experience with the rated practices in the case of the laymen. A greater probable error is to be expected. The figure is 0.33 scale units.

This comparison affords necessary warning as to what differentials secured by application of the instrument are to be considered significant. Stated in another way, such differentials begin to have significance when greater than 0.66 scale units. They would appear to be particularly significant when greater than 0.99 scale units. This is not as precise as might be wished. But doubtless it is the truth to be found in social data. The scale may be practical, if not absolutely ideal. It would seem, however, that many of the differentials which are to be found in data yielded through application of the scale are sufficiently great to be significant. One should note, furthermore, that the failure of absolute reliability, just warned of, is very much less than that which would otherwise be encountered through continued reliance upon description in off hand free verbalization, language often loaded with prejudice and bias. Use of the instrument requires description sharpened by the necessity of record in terms of specific behaviors rather than impressionistic verbal pictures of whole complexes of behaviors, pictures often blurred by rationalization, bias, and wishful thinking. For comparison of observations and opinions of a number of investigators or subjects, the scale may serve as a common descriptive language, the degree of precision of which is known and can be accounted for. If absolute precision is unattainable in dealing with social data, that which is then to be desired is as great precision as is possible and some close check upon the amount of the lack of precision. That end is served by the scale.

Continuing the comparison of tested scales, one may point out that in the case of Professor Thurstone's reported scale the mean discrepancy in the scale values of the landmarks as rated by 150

<sup>15</sup>L. L. Thurstone and E. J. Chave, *The Measurement of Attitudes*, p. 42.

<sup>16</sup>See the *American Political Science Review*, Vol. XXVI, pp. 543-544. See also, pp. 531 and 538-539.

raters and as rated by 300 raters, all being selected on the same basis, was 0.056 scale units.<sup>17</sup> In the case of the candidate scale, where the raters were not selected on the same basis, but were 100 adherents of major parties as contrasted with 100 non-adherents of major parties, selected from different areas, the mean discrepancy in the scale values of the landmarks as rated by the first group and as rated by the entire two hundred was 0.084 scale units, and the similar discrepancy in the ratings of the second group as compared with the ratings of the entire two hundred was 0.088 scale units. Now when similar figures are considered for the different groups rating the third degree statements as compared with the entire body of raters, the figures are found to be in the case of the one hundred laymen, 0.338 scale units; and in the case of the one hundred and six troopers and prisoners, 0.396.

The statements of this scale, furthermore, have a little greater ambiguity than those of the candidate scale.<sup>18</sup> That is to say, the meanings which will probably be read into them by different subjects will present a little greater variation. With troopers and prisoners rating these statements, there was a more vital clash of interests than with opposed partisans ratings statements indicating favor or disfavor toward a candidate for elective governmental office. Probably it is not so much a matter of indefinite language as variation of response to language symbols. Such a matter is understandable to those accustomed to deal with social data. But particularly important is it to know the extent to which such a factor enters into one's descriptive record. Thus, the average of the semi-interquartile ranges of the ratings of the candidate scale statements was 2.0 scale units; while the average of the similar ranges of the ratings of the statements of the third degree severity scale is 2.9 scale units. Even so, the scale position assigned each statement is tested to be the best approximation possible. It will come nearer to representing the majority of reactions than any other possible assignment.

In reporting upon the testing of the scale, it should be added that Professor Thurstone's tests for irrelevancy of statements to the selected variable are inapplicable. In building affect scales such a test is needed, for in that case one is using as landmarks opinions which may in many subtle ways indicate some variable other than that selected to be measured. In the case of a scale for the measurement of some selected characteristic of related behaviors, however,

<sup>17</sup>L. L. Thurstone and E. J. Chave, *The Measurement of Attitudes*, pp. 42-44.

<sup>18</sup>See *American Political Science Review*, Vol. XXVI, pp. 539-541, 543-544.

one actually starts with the very behaviors of interest. It becomes a matter of discarding those descriptions which are not clear and which have meaning that can not be fixed definitely upon the continuum. It is not a matter of guarding against inclusion of faulty indications of the matter in question. Logically, too, professor Thurstone's tests

are inapplicable. Whether one uses the formula  $\frac{\frac{n}{n} ab^{19}}{b}$  or what

amounts to the formula  $\frac{\frac{n}{n} ab^{20}}{a}$  the theory of the test is that, of those

who endorse a landmark, fewer and fewer will endorse other landmarks located increasing distances away in *either* direction on the scale. This works in the case of data involved in issues where more than one mode may be expected. It does not work, however, in the case of institutional data where those who endorse a given landmark may be expected to accept most landmarks located on the scale in *one* of the directions from the given landmark, but fall away in their endorsement of landmarks located at points farther and farther away in the *other* direction. The data collected for the employment of this test show a gradual decline in the frequency of endorsement as one considers statements scaled at higher and higher degrees of severity. This should afford considerable confidence that the statements are not inconsistent in their relation to the selected behavior variable. As further application of the scale yields additional data, continued probing may be expected to develop some index for testing what would be comparable to Professor Thurstone's test for irrelevance, if such appears to be needed.

The significance of this scale for inquiry respecting the third degree has been suggested. It remains to suggest an even greater significance of this sort of scale. Such significance attaches from the fact that it points the way to further development of scales which would measure some characteristic of related behaviors. Development along this line would seem to offer great promise for the introduction throughout the entire field of inquiry called political science of greater precision of observation, and of notation upon observation,

<sup>19</sup>L. L. Thurstone and E. J. Chave, *The Measurement of Attitude*, pp. 46-55.

<sup>20</sup>L. L. Thurstone, "Attitudes Can Be Measured," *American Journal of Sociology*, Vol. XXXIII (1927-28), pp. 549-551.