The Thematic Apperception Test as a Stimulus for Experimental Research on Lie-Detection in a Cross-Cultural Setting

	THONNEY-PRUNELL JAIME
	18
	238
URL	http://hdl.handle.net/10097/37026

トン ニ　プルネル　ジェイミ
THONNEY-PRUNELL JAIME

学 位 の 種 類	博 士 (文 学)
学 位 記 番 号	文博第　238　号
学位授与年月日	平成19年3月8日
学位授与の要件	学位規則第4条第1項該当
研 究 科 ・ 専 攻	東北大学大学院文学研究科 (博士課程後期3年の課程) 人間科学専攻
学 位 論 文 題 目	The Thematic Apperception Test as a Stimulus for Experimental Research on Lie-Detection in a Cross-Cultural Setting (虚偽検出実験刺激としての主題統覚テスト (TAT) の 異文化設定における利用)
論 文 審 査 委 員	(主査) 教　授　行　場　次　朗　　教　授　大　渕　憲　一 教　授　仁　平　義　明 教　授　海　野　道　郎 助教授　阿　部　恒　之

論 文 内 容 の 要 旨

FOREWORD

The present research addresses the difficulty in reproducing in an experimental context some of the emotional states present during a polygraph interrogation in a real investigation. It is based on earlier research on the methodological issues of the polygraph test done during the author's MA course (2001-2003). Mixed research findings were detected and original experiments were conducted to cast new light on the mechanisms through which polygraph tests work, and offer a valid alternative to the average laboratory lie-detection paradigm. The present research included the analysis of the participants' autonomic responses, overt behavior, and provided new finding about areas of interest to cross-cultural psychology.

The polygraph test - methodological issues: emotional realism and ground truth.

There are two major problems concerning the validity of polygraph techniques. The first one is the difficulty of corroborating their findings by criteria independent of the technique itself (Bradley & Cullen, 1993; Vrij, 2000). In field polygraph testing, many innocent suspects confess to reduce their sentence (Vrij, 2000), duress or an induced sense of guilt (Conti, 1999), and the real accuracy of the test cannot be ascertained. In laboratory studies

this problem is avoided: the subjects' responses can be compared with a *ground truth* because the researchers assign the subjects to conditions they control. However, a second problem arises: the artifice of the laboratory situation. Experimental studies are usually based on the guilty subject paradigm, thus including a staged, 'mock' crime (Bradley & Cullen, 1993) of which the participants are 'innocent' or 'guilty'. Sometimes subjects have to role-play a theft (U.S. Congress, 1983); sometimes they have to imagine committing a crime (Boucsein, 1992). To increase their motivation and facilitate detection, they are offered a reward if they 'beat the test' (U.S. Congress, 1983). Thus, mock crime studies have been criticized (Bradley & Cullen, 1993; Andreassi, 1995; Vrij, 2000) because of their lack of realism: the emotional state of their participants is hardly the same as that of suspects of an actual crime. Lykken (as cited by Andreassi, 1995) has criticized such experiments on the basis that they were viewed as games by the participants; according to Vrij (2000), their arousal can be caused by sheer delight. Iacono (as cited in Bradley & Cullen, 1993) argues that using subjects who have 'committed' a 'mock crime' to establish the polygraph's accuracy is like using subjects who pretend to be ill, instead of diseased people, to validate data in the field of mental health. Both problems, the orchestrated nature of mock crimes and the subjects' emotional states, are interconnected and they limit the extent to which the conclusions drawn from laboratory studies can be extrapolated to real-life situations.

Bradley and Cullen (1993) avoided these problems by introducing the use of embarrassing incidents as stimuli: stories narrated by the participants about true events in their own lives that had an emotional impact on them. Thus, they created a situation that was emotionally charged and had a ground truth to evaluate the test's accuracy. The polygraph was significantly more accurate when the stimuli were emotionally arousing, yet, for ethical reasons, the stories were non-criminal, and the emotional state they elicited was exclusively that of embarrassment. The chosen stimuli for this experiment were thought capable of arousing feelings of shame *and* guilt without asking the participants to tell the experimenter of true events, thus avoiding ethical problems.

Earlier experiments induced shame and guilt by asking their subjects to write stories about true experiences in which they blamed themselves (Dickerson, Kemeny, Aziz, Kim & Fahey, 2004) or exposing their subjects to pictures (Ritz, Claussen & Dahme, 2001) and found significant bodily reactions. Cramer (2003) studied the relationship between defense mechanisms and physiological reactivity to stress though the application of the Thematic Apperception Test (TAT), finding a blood pressure increase. From a clinical viewpoint, the TAT is as a tool to expose underlying inhibited tendencies which people are not willing to admit, or cannot admit because they are unconscious (Murray, 1943). The TAT is projective test; we define 'projection' as a mechanism of defense, the psychological tendency of people to interpret ambiguous stimuli in conformity with their past experiences and present needs, conscious or not (Murstein, 1963). The arousal of needs and affects by the pictures and their associations may be quite apparent in the patient's behavior and tone of voice (Murray, 1943). *We posited that the exposure to the TAT situation would induce emotional states evident through the psychophysiological arousal they would provoke. That, and the story telling required by the TAT instructions, would guarantee a situation more emotionally truthful than a mock crime, inasmuch as those stories would be based on the subjects' own emotional experience.*

The targeted emotions, guilt and shame; a cross-cultural approach

In experiments that attempt to induce emotional states, the target emotions ought to be those that appear during the field practice. Special care was employed to ensure this design aroused the desired emotional states;

it required stimuli capable of arousing guilt as well as shame, and since 1946 there has been a controversy about the definition and prevalence of those feelings in among Japanese people. Benedict (1946) made the distinction between shame cultures as opposite to guilt cultures, and placed Japan amongst the former. She defined guilt cultures as inculcating absolute standards of morality which become internalized, whereas shame cultures rely on external sanctions, requiring an audience or a man's fantasy of an audience. Various criticisms of Benedict's work have been advanced. Lebra (1983) used the TAT as a research tool and criticized Benedict's characterization. Vogel (1989) stressed that Benedict's research was not conducted in Japan. Gardner (1977) granted that although remorse may be more common in Western cultures, it is present in Japan. Najima (1977) emphasized that for the Japanese shame has a crippling effect similar to guilt for Westerners. The definitions of shame and remorse also vary. Devos (1974), who also used the TAT, related remorse to the Freudian theorization of the Superego and shame to the Ego Ideal. De Rivera (1989) defined remorse as an emotion in which we take responsibility for our actions and shame as the emotion when we realize we are who we do not want to be. When faced by a similar problem of definitions in cross-cultural research, Hermans (1999) concluded that a consensus exists: shame implies an external sanction, whereas guilt occurs when internalized rules have been broken, but *both feelings are interrelated.* We followed these definitions here.

Experimental research: the TAT as a stimulus for lie-detection techniques in laboratory settings

Hypothesis

The Guilty Knowledge Test (GKT) would provide better accuracy rates if the subjects were replying about stimuli that might stir up emotions of guilt and/or shame.

Method: participants and apparatus

Forty-eight Tohoku University undergraduate and post-graduate students volunteered to participate in the experiment; their ages ranged from 18 to 25 years old, the mean being 20.91 and the SD 1.36. Eighteen of them were male and thirty female.

After the completion of a questionnaire about their personal data, they entered an electrically shielded room where a skin conductance response meter was attached to them. Their Galvanic Skin Reflex (GSR) responses were amplified outside the room and shown on a computer monitor. A Windows 'PowerLab' model was used to record and analyze the data; the participants' facial expressions and their TAT stories were recorded for further analysis.

Stimuli

The Western TAT was deemed a better choice than the Japanese TAT since the latter depicts people dressed in traditional Japanese clothes nowadays worn mostly by the elderly, and the participants of our experiment were young; also, the clothes worn by the Western TAT characters are the kind of clothes worn in everyday life in Japan nowadays. Since there were no close-ups of the characters, they could also be seen as Asian, allowing the participants to project and identify with them. The Western TAT was also judged appropriate because of the findings provided by previous studies: to determine the validity of the TAT for Japanese subjects, Gray (1999) compared the TAT themes provided by Japanese subjects who took the Western TAT with the American themes reported by Stein (1948); seven out of the ten applied cards shared the same themes for both groups.

A special emphasis was placed on the TAT card selection. First, those cards in which the human figures were

not 'facing' the observer, but looking away, were selected. The criterion used was the linguistic/psychological implications of the idioms: *mentsu ga tsubureru* ('to lose face', connected to feelings of one's own embarrassment) *kao ga tatanai* (to embarrass other people) and *kao ni doro o nuru* ('to throw mud in somebody's face', with connotations of embarrassment). All these expressions include words that mean 'face' (*mentsu, kao*). Also in European languages there are several expressions linking the emotions of shame and guilt to the facial zone. As Hermans (1999) suggests, the existence of guilt and shame as well as a relationship between the two emotions seems to be universal. A second criterion to select the TAT cards was the range of options given to the subjects. The four chosen TAT cards depict a human figure alone (3BM), a human figure with other people who can be partially seen partially (18MB), two human figures (13MF), and a human figure with other people in the background (17GF). Finally, as a third criterion to select the stimuli, we allowed the participants different possible sexual identifications; the chosen stimuli depicted a male, a female, a couple, and a sexually ambiguous stimulus. Following previous studies (Lebra, 1983; Finn 1986) the TAT cards were complemented with verbal cues to make sure they elicited the desired emotional responses: the words *hanzai* and *tsumi* (crime), *koukai* (guilt), *haji* (shame) and *douyo* (to be in shock for what one has done) were included as part of the instructions to make the story.

We used four cards selected from a set of slides used by Buck, Caul and Miller (1974) to study communication of affects as control stimuli since (a) they did not include human figures and the mechanism of projection is facilitated by human stimuli, and (b) they were neither pleasant nor unpleasant. They included a lake and a pier, trees against the skyline, a meadow and a fence, and a small river. A verbal cue similar to that of the TAT series was used for the neutral cards to equalize both sets of stimuli.

Procedure

We conducted the experiment under a guilty information paradigm. The non-arousing stimuli were one 'buffer' card + 4 emotionally neutral cards; the emotionally arousing stimuli were one 'buffer' card + 4 TAT cards. In both cases, the subjects were asked to elaborate a story about the card they selected: the most emotionally neutral picture in the first case, the picture that best expressed a situation connected to crime, guilt and shame in the second. Then a standard GKT was applied to the subjects.

Data Evaluation: Criteria and Definition of Detection Rate

Since there is no universally accepted criterion for scoring polygraph charts (Vrij, 2000), we used a clearly defined criterion of an exclusively quantitative nature. The GSR amplitude was the chosen autonomic response because electrodermal responses have been found to be the most accurate single index in experimental settings (Andreassi, 1995; Matte, 1996; Waid & Orne, 1981); they provide clear-cut short-term changes that increase as a function of stimulus significance (Boucsein, 1992); they have been related to experimentally induced emotional states (Khalfa, Isabelle, Jean-Pierre & Manon, 2002); the original analyzing system proposed by Lykken for his GKT was solely based on the measurement of skin response amplitudes (Honts & Amato, 2000); the measurement of the GSR amplitude constitutes an average real-life practice (Matte, 1996).

Each subject completed three test trials and three control trials, six in total. Detection rate was defined as the percentage of times the card that the subject had chosen evoked the greatest autonomic response when compared with the other stimuli of the trial. Each subject was given two GKT scores based on that rate, one for the control cards and another for the treatment cards, e.g.: 0/3, 1/3, 2/3, or 3/3: higher GKT scores meant higher accuracy.

After the completion of the experiment, the GKT scores for TAT and neutral stimuli were compared.

GKT: statistical analysis and results

The Wilcoxon's matched-pairs signed-ranks test was applied to determine the statistical significance of the different GKT scores obtained through the GSR analysis for the two groups considered: neutral stimuli treatment and emotionally arousing stimuli treatment. The GKT detection rate based on these scores was significantly higher ($z = -1.752$, $p < .05$, one-sided) when the participants replied to guilt/shame-related stimuli. The table shows the number of participants that obtained each one of the possible scores (0/3, 1/3, 2/3, or 3/3) for both treatments.

		Neutral Condition				
		0/3	1/3	2/3	3/3	
TAT condition	0/3	4	3	2	0	9
	1/3	0	4	2	0	6
	2/3	4	6	3	0	13
	3/3	0	2	3	1	6
		8	15	10	1	34

This table shows the mean of the GKT scores for the control group (neutral stimuli condition) and the mean of the GKT scores for the treatment group (TAT condition) and their SD.

Stimuli	Mean	SD	N
Neutral	1.12	0.81	34
TAT	1.47	1.08	34

A total habituation of responses, which resulted in the absence of skin conductance reactions, occurred for fourteen of the subjects. No comparison was possible for those subjects and their charts were labeled 'inconclusive' and excluded from the analysis.

Discussion: experimentally induced emotions; the TAT in a lie-detection experimental context.

The use of TAT cards as emotionally arousing stimuli rendered significantly higher GKT scores than the use of neutral stimuli: in principle the induction of an emotional state by the TAT cards can be assumed since the scoring of the GKT is based on autonomic disturbance, and although emotion may not be simply identified with arousal, it is directly correlated with it (Hebb, 1965). Since the stimuli (visual and verbal) were structured on the basis of their connection to the emotions of guilt and shame, the preliminary assumption that those were the two emotions aroused deserves consideration. These findings offer experimental evidence that substantiate the existence of the non-cognitive component in the GKT to which Suzuki, Nakayama and Furedy (2004) referred.

Experimental research: guilt and shame in Japan: TAT data

Method: criteria of evaluation

We defined *guilt* as the presence of mourning, expiation, retribution and penitence in relation to the crime committed by the character; *shame* was defined as the presence of social punishment (e.g. disapproval, isolation or exclusion from group of peers, neighborhood, and family) to the character of the story.

TAT results

Twenty subjects (41 % of the sample) chose TAT card 13MF as the one that better expressed a situation in which a crime had been committed and the responsible person felt guilt and/or shame. Twenty-five participants (53 % of the sample) selected TAT card 3BM. Only three participants (6 %) chose TAT card 18MB and none chose TAT card 17GF. Fifteen subjects (31 % of the sample) chose neutral card I as one that was neither pleasant nor unpleasant for them. Nine participants (19 % of the sample), fifteen participants (31 %), and nine participants (19 %) chose neutral cards II, III, and IV respectively. Thirty subjects (64 % of the sample) made TAT stories that revolved around the emotion of guilt; five subjects (10 % of the sample) told stories in which the key emotional state was *shame*, ten participants (20 % of the sample) included both emotions in their stories and three subjects (6 % of the sample) narrated stories that included a crime with no signs of emotional repercussion or consequences for the characters, whose actions they regarded justified. The following chart shows the emotional content of the total of the stories in percentages
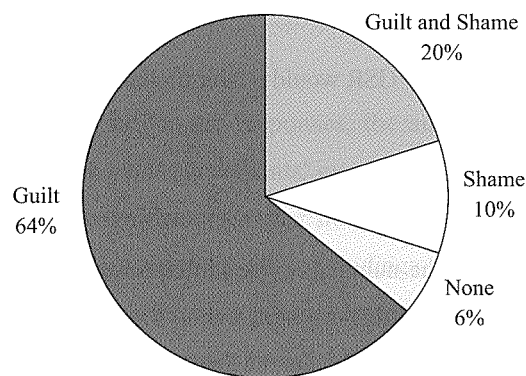
The TAT themes provided by these subjects were compared with the American themes reported by Stein (1948), like Gray (1999) did. Of the themes not reported by Stein but present in our sample, the majority (three themes out of five) were themes presented by single subject.

Discussion

The subjects' selection of neutral stimuli was evenly distributed among the four cards, an expected result since they were emotionally unexciting and homogeneous. The choices for the TAT cards, on the other hand, showed a strong tendency towards two of the four stimuli. The two preferred cards were those in which there is no social context. In card 3BM there is only one human figure. For card 13MF, all subjects but one told a story in which the female figure was dead and the male character was alone with the evidence of his crime. This argues *against* the prevalence of feelings of shame over guilt among Japanese people. An analysis of the themes related to guilt and shame strengthens this supposition: they show more similarities than differences between the USA and the Japan samples. Although Benedict acknowledged the limitations imposed by the lack of fieldwork in Japan, she did not seem to recognize that the circumstances in which her interviews took place might have changed her findings. The same kind of criticism made to Benedict's work may be pertinent for our results. The instructions we gave the participants -they were specifically asked to imagine the circumstances of a crime- might have biased the content of their stories. In a more general way, the main limitation for the extrapolation of our TAT-related findings is that they are only secondary to the main objective of our experiment.

Eye blink as an index for lie-detection - Antecedents

There is mixed evidence about the use of eye blinking as a lie detection index. Some researchers (Fukuda, 2001; Cutrow, 2000) have statistically proven it useful as an additional index, second to GSR responses. Others (Cody, Lee, & Chao, 1989; Vrij, 2000; Zuckerman, 2004) did not find any definite correlation between deception and eye blinking. The findings of Mann, Vrij and Bull (2002) contradict the popular belief that liars blink more because they behave nervously.

Experimental research: eye blink as a lie-detection index

Hypothesis (1)

We hypothesized that the GKT would also provide better accuracy rates if the subjects were replying about emotionally charged stimuli and that the exposure to the TAT situation would induce emotional states evident through the subjects' eye blinking.

Hypothesis (2)

However, we expected that the subjects' GSR would render higher accuracy rates than the use of eye blinking, that is *it would be a better index.*

Method

One of the experimenters counted the number of eye blinks for each subject during ten seconds after the stimulus onset. To determine the reliability of this eye blink count, a second experimenter performed the same task for 5 subjects, equivalent to 10, 45 % of the sample, and the Pearson Correlation Coefficient was calculated.

Data Evaluation: Criteria and Definition of Detection Rate; Statistical analysis and results

Those trials where the largest numbers of eye blinks corresponded with the subjects' chosen card were defined as a 'successful detection of deception'.

The Pearson Correlation Coefficient was calculated as .977, indicating a high correlation and therefore a high reliability.

Hypothesis (1) The Wilcoxon's matched-pairs signed-ranks test was then applied. The GKT eye blinking detection rate based on these scores was significantly higher ($z = 1.759$, $p < .05$, one-sided) when the participants replied to guilt/shame-related stimuli, which substantiates our main hypothesis. This table shows the number of subjects and their accuracy scores (0/3, 1/3, 2/3, 3/3) for both condition:

		Neutral condition				
		0/3	1/3	2/3	3/3	
TAT condition	0/3	12	1	1	2	16
	1/3	7	8	4	0	19
	2/3	9	1	2	0	12
	3/3	0	0	1	0	1
		28	10	8	2	48

Hypothesis (2) The Wilcoxon's matched-pairs signed-ranks test was applied; the GKT detection rate based on GSR scores was significantly higher ($z = 2.593$, $p < .01$, one-sided) than the accuracy scores provided by the subjects' eye blinks, corroborating the long-standing use of skin responses in lie-detection. This table shows the numbers of subjects and their accuracy scores (0/6, 1/6, 2/6, 3/6, 4/6, 5/6, 6/6) for both indexes.

		Eye-blink scores							
		0/6	1/6	2/6	3/6	4/6	5/6	6/6	
	0/6	1	0	2	1	0	0	0	4
	1/6	1	1	0	0	1	0	0	3
GSR	2/6	1	3	5	1	0	0	0	10
scores	3/6	3	2	2	0	0	1	0	8
	4/6	2	0	2	1	0	0	0	5
	5/6	0	1	2	0	0	0	0	3
	6/6	0	0	1	0	0	0	0	1
		8	7	14	3	1	1	0	34

Discussion

Like the analysis of the subjects' autonomic responses, the analysis of their overt behavior provided statistically significant results: by using the participants' eye blinking as an index the use of TAT cards as emotionally arousing stimuli rendered significantly higher lie-detection scores than the use of neutral stimuli. These results strengthen the importance of the GSR-related findings, and the way we interpreted them. These results arise two questions: (a) Are GSR and eye blink reliable indexes, significantly better than chance? (b) Why is there mixed evidence with respect to eye blinking as an index?

GSR and eye blink as indexes compared to chance rate

Finally, we aimed at investigating whether the accuracy rates obtained in this experimental design were significantly better than the accuracy expected according to sheer chance.

Hypothesis (1)

Eye blinking used as an index for the GKT would provide better accuracy rates than sheer chance.

Hypothesis (2)

GSRs used as an index for the GKT would provide better accuracy rates that sheer chance.

Statistical analysis and results

The Binomial Test was applied to determine the significance of the difference between the GKT scores provided by the subjects' a) eye blinks and b) GSRs, and the accuracy expected by sheer chance. The GKT detection rate based on eye blinks was significantly higher ($p < .05$, one-sided) than the accuracy expected by pure chance for the TAT condition, but not for the neutral stimuli The GKT detection rate based on GSR amplitudes was significantly higher ($p < .01$, one-sided) than the accuracy expected by pure chance for both the control and the treatment group.

	TAT condition = Emotionally arousing stimuli	Neutral condition = Emotionally neutral stimuli
Eye blinking as a lie detection index	Lie detection accuracy **significantly higher than chance** (p < .05, one-sided)	Lie detection accuracy **NOT significantly higher than chance** (p > .05, one sided)
GSR as a lie detection index	Lie detection accuracy **significantly higher than chance** (p < .01, one-sided)	Lie detection accuracy **significantly higher than chance** (p < .01, one-sided)

Discussion

Skin responses were found to be a useful lie-detection index regardless of the presence or absence of emotionally arousing stimuli, which is consistent with its long-standing use a lie-detection index. Eye blinking was found to be a useful index only for the condition that included emotionally arousing stimuli, which seems to indicate it cannot be used regardless of the context in which the test is applied and quality of its stimuli. These findings might be used as the basis for future research on eye blinking as a lie-detection index.

Final Discussion - summary

1) The present research has provided an experimental antecedent that might help answer one of the most important methodological issues concerning the lie-detection field: how experimenters can preserve the emotional quality of a lie-detection interrogation and at the same time have confident access to a ground truth to validate their data.

2) It has provided a new paradigm to do research on lie-detection in laboratory settings, a paradigm that has provided statistically significant results using autonomic indexes and overt behavior.

3) It constitutes a new approach, emphasizing emotional phenomena, to a field where the key processes have recently been thought to be cognitive.

4) Since the chosen stimuli are part of a clinical test, it constitutes an antecedent to future research concerning the manifest gap between clinical and experimental psychology.

5) Further research using this paradigm may cast more light on the reasons why there is such mixed evidence regarding the usefulness of eye blinking as a lie-detection index.

6) It has provided new data, based on a theoretically clinical framework but which were experimentally obtained, about a polemical topic of interest to cross-cultural psychology: is shame -as opposed to guilt- really prevalent in Japan?

7) To sum it all up: it has provided two sets of statistically significant data obtained experimentally plus an extra set of results of interest to clinical and cross-cultural psychology, constitutes an important antecedent to at least two different areas of knowledge, and last but not least, this research is, to the extent of my knowledge, 100% original.

論 文 審 査 結 果 の 要 旨

　本論文は、皮膚電気反応と瞬目を指標とした虚偽検出実験において、投影法の心理検査の一つである主題統覚テスト（Thematic Apperception Test, TAT）の図版を刺激として用いることの有効性を見出し、従来の虚偽検出実験の弱点を補い、その応用可能性を高めようとしたものである。本論文は全9章からなる。

　第1章では、自律性反応計測を利用した従来の虚偽検出研究の概観と、理論的問題点の指摘がなされた。実際の虚偽検出場面では、被検査者のリアルな情動的反応が得られると考えられるが、検査者にとって刺激内容のうち何が真実で何が虚偽なのか根本的には不明であるため（absence of ground truth）、測定結果と独立した明確な基準をもつことができず、虚偽検出結果の信頼性と妥当性の評価が困難である問題点があった（Vrij, 2000）。一方、実験的研究では、測定者はあらかじめ刺激内容の真偽を決定できるが、今度は、実験参加者の情動反応がリアルなものであるか不明になる問題が生じる。例えば、実験参加者に模擬犯罪（mock crime）のロールプレイをさせ、虚偽検出を行う課題では、参加者に虚偽にともなうストレス的な情動反応が生起するのではなく、実験者を欺くのに成功すると、むしろ快反応が生じる場合さえある。そこで本論文の提出者は、実験場面の情動反応をより実際場面に近づけるために、投影法の心理検査の一つである主題統覚テスト（Thematic Apperception Test, TAT）の図版を刺激として用いることを考案した。

　第2章では、TAT が典型的に誘発する罪（guilt）と恥（shame）の感情の違いについて、異文化比較的な研究が概観された。本研究で展開された実験では、日本版 TAT ではなく、Murray（1943）の Western TAT の図版が使用された。その理由は、日本版 TAT では和服を着た人物が登場したりして、今日の日本の日常にはマッチしないだけでなく、Western TAT を用いた先行研究のほうが圧倒的に多く、結果の比較検討が容易であるからであった。日本では恥の感情が罪の感情よりも優位になるという先行研究があるが、TAT を用いた後続研究では、必ずしもそのような傾向が明確に見出されていないことなどが指摘された。

　第3章では、罪意識テスト（Guilty Knowledge Test, GKT）パラダイムを使って、TAT 刺激が虚偽検出実験により適したものであることが、自然風景刺激を使った実験を統制条件として、実証的に示された。実験参加者は4枚の TAT 図版から一つ選び、どのような感情が誘発されたかを回答するとともに、自分なりの物語を作成することを求められた。実験では、バッファ刺激とともに、4枚の TAT 図版が提示され、それが実験参加者が選んだものであるかどうか質問され、すべてに、「いいえ」と答えることが求められ、その間の皮膚電気反応が皮膚伝導計によって測定された。自然風景刺激についても同様の手続きで GKT が実施された。その結果、TAT 図版を用いたほうが、自然風景刺激よりも有意に高い虚偽検出率が得られることが明らかにされた。これは TAT 図版が実験参加者に投影的な感情喚起を促し、より大きな自律性反応が生起したためであると考察できる。

　第4章では、上記の Western TAT を用いた実験で、アメリカの学生を対象としたデータと比較しながら、日本人の実験参加者が実際にどのような感情を抱いたかに関する質的考察がなされた。その結果、罪や恥の感情に関わらず、喚起された感情の種類と頻度は Stein（1948）がリストしたものとほぼ同様の内容であったが、病気や貧困、安楽死などのユニークなテーマをあげる参加者が若干見られることがわかった。

　第5章では、表出行動（overt behavior）の一つである瞬目を指標に用いた虚偽検出実験に関して、従来の研究の概観と理論的問題点が考察された。先行研究では、瞬目の虚偽検出の有効性について相反する

データが提出されている。これは、瞬目の頻度が虚偽意識だけでなく、種々の要因によって複雑に変化することに起因する問題であり、現状では、補助的な意味でしか虚偽検出に利用できないことが示された。

第6章では、上記の先行研究の相反するデータを踏まえながらも、TAT 図版を刺激として用いた場合には、瞬目が虚偽検出に有効がどうか、第3章で記述された GKT パラダイムを用いて実験的な検証がなされた。その結果、瞬目を指標に用いた場合でも、検出率は皮膚電気反応を用いた場合よりも低いものの、TAT 図版を用いたほうが、自然風景刺激よりも有意に高い値が得られることが明らかにされた。

第7章では、虚偽検出における偶然的正答率（チャンスレベル, chance level）の問題が議論された。第1章で述べたように、実際場面では、真の事実（ground truth）が不確定のため、チャンスレベルを正確に見積もることは不可能である。しかし、第3章で紹介した罪意識テスト（GKT）パラダイムを用いた実験では、選択肢の数を実験者が任意に設定することが可能なので、チャンスレベルを正確に導出することが可能である。虚偽検出手法が有効であるためには、いうまでもなく、チャンスレベルを上回る検出率を達成しなければならない。

そこで第8章では、皮膚電気反応と瞬目を指標に用いた GKT による虚偽検出率がチャンスレベルを上回るかどうか、TAT 図版と自然風景画像を刺激に用いた場合について、それぞれの組み合わせについて統計的に分析された。その結果、TAT 図版を刺激として用いると、皮膚電気反応と瞬目を指標に用いた虚偽検出のどちらも場合もチャンスレベルを有意に上回ったのに対し、自然風景画像を刺激にした場合には、皮膚電気反応を指標としたときにはチャンスレベルよりも有意に高い検出率が得られたが、瞬目を指標とした場合には、チャンスレベルを有意に超えることはできないことがわかった。

第9章では、これまで得られた知見を総合した議論が展開された。虚偽検出を実験室場面で行うときには、自律性反応を指標とする場合でも、瞬目のような表出行動を指標とする場合でも、TAT 図版のように、実験参加者に投影的な感情喚起を促し、より大きな潜在的・顕在的反応が生起するような工夫を施すことが極めて大切である。先行研究で、瞬目の虚偽検出の有効性について相反するデータが提出されている点についても、このような工夫がなされたか否かに大きく依存するので、結果を慎重に解釈する必要がある。また、Western TAT の図版は、日本人の実験参加者に提示した場合でも、投影的な感情喚起を文化差なく促すので、幅広い利用が可能である。

以上のように、本論文は、皮膚電気反応と瞬目を指標とした虚偽検出実験において、投影法の心理検査の一つである主題統覚テスト（TAT）の図版を刺激として用いることを考案し、その有効性を実験的に検証したものである。従来の虚偽検出実験の弱点を補うことに成功したといえる。これらの点から、本論文は、理論的にもまた応用的にも有意義な成果をもたらし、この分野の研究発展に貢献するところが大きい。よって、本論文の提出者は博士（文学）の学位を授与されるに十分な資格を有するものと認められる。