Graham Davies, Sally Lloyd-Bostock,
Mary McMurran, Clare Wilson (Eds.)

Psychology, Law, and Criminal Justice

de Gruyter

International Developments in Research and Practice

Psychology, Law, and Criminal Justice

International Developments in Research and Practice

Edited by
Graham Davies
Sally Lloyd-Bostock
Mary McMurran
Clare Wilson

Walter de Gruyter · Berlin · New York 1996

Contents

Part VIII: Prison Psychology

Part IX: European Perspectives on Law and Psychology

Communication Channels and Gender Differences in Decoding and Integration of Cues in Legal Decision-Making[1]

Gerhard Kette and Vladimir J. Konečni

Introduction

It has been frequently asserted that judges and juries integrate non-evidentiary and extra-legal elements into their verdict decisions. In civil-law cases, the verdicts - perhaps of necessity - often involve ambiguous personal and psychological data about the litigants, so that such cases may be especially vulnerable to extra-legal biasing influences.

Gender of legal participants, as a particularly controversial extra-legal factor, has been the subject of an especially large amount of researchers' attention. In this paper, we consider the gender of the "judge" as an extra-legal factor in civil-law (small-claims) cases. However, we do so in the context of examining the extent to which the decoding and decision-making strategies dealing with emotion-laden information are differentially affected by gender. We shall first consider the potential biases induced by the sex of judge, and culprit, as well as any that may be induced by the differential weighing and processing of emotional information.

The experimental data in which the *sex of the juror* served as an independent variable established few consistent differences between the decisions of the male and female jurors. The bulk of the data suggest that men are more active in jury deliberations than women and thus may be more influential in determining verdicts (for a summary see Kette, 1994). Most of the actual trial data relating to gender were based upon civil suits. Jurors favored plaintiffs of their own sex in verdicts, but consistently awarded higher damages to men than to women. When the sex of the juror does significantly predict decision behavior, the results sometimes go in one direction (e.g. females acquit more than males) and sometimes in the opposite one.

Studies Examining Gender-Related Issues in Judges' Decision Making

In reviewing the experimental jury research, few studies have examined the relationship between the sex of the judge and his or her sentencing behavior (most of the latter research was conducted with archival data). This reliance on archival methodology reflects, in part, the absence of information about the sentencing behavior of female judges. In most jurisdictions the majority of judges is male and thus relatively few offenders are sentenced by female judges. Most of the results of these archival analyses must therefore be considered tentative. In the past, researchers have mainly focused on interactions between the sex of the judge and three case

characteristics: the offense of which the defendants are convicted and the sex and race of defendants.

In an archival study using more than 27,000 cases, Myers and Talarico (1987) found that contrary to expectation, the prison sentences (of both black and white offenders) tended to be longer in courts where female judges presided. They also considered the issue of whether female judges treated certain types of offenders differently than males judges. The general expectation was that, with the exception of rape, female judges would sentence more leniently than male judges. In actuality, they found considerable similarity in the sentences imposed by male and females judges. Where significant differences existed, they suggested that female judges may be more severe than male judges toward both violent and property-crime offenders.

Despite small sample sizes of female judges, their data also provided some support for the expectation that female judges sentenced rape offenders more harshly than male judges. Rape offenders would more likely be imprisoned and receive longer prison sentences if sentenced by female judges. Surprisingly, however, they received more lenient split sentences (time in jail, followed by probation) from female judges. Finally, they found that male judges did not treat female offenders more leniently than females judges. In fact, male judges appeared to sentence female offenders significantly more harshly than did female judges, and there was also some evidence that male judges sentence male offenders more leniently than female judges. Again, samples were small, and these differences are suggestive rather than definitive.

The most detailed information about sex differences in sentencing behavior is based on archival research by Kritzer and Uhlmann (1977). They developed competing expectation about the sentencing behavior of female judges, arguing, on the one hand, that female judges may be more threatened than male judges by normative violations and for this reason be more punitive, and, on the other, that because of differences in sex-role socialization, females judges may be more lenient. Previous evidence of "chivalrous" or paternalistic attitudes toward female offenders led Kritzer and Uhlmann to expect that male judges would punish female offenders more leniently than would their female colleagues.

Virtually no support was found for these hypotheses. In general, the sentencing behavior of male and female judges was more alike than different. Most notably, there were no significant differences in the way male and female judges treated female (or male) offenders. Male judges did not appear to be paternalistic toward female offenders, nor did females judges appear to be more empathetic. Although female judges treated rape offenders more harshly than did male judges, the differences were not significant. Only one difference in sentencing behavior was noteworthy: female judges punished larceny offenders more harshly than did male judges. Other differences in sentencing behavior existed, but could be traced to legally relevant facts.

More recent work with archival data (Gruhl, Spohn, & Welch, 1981; Myers & Talarico, 1987) has demonstrated a slight (but statistically non-significant) tendency for female judges to use the prison option more frequently than do male judges. Although sentencing behavior differed depending on the type of offender, there was no consistent tendency for female judges to be lenient or male judges to be more

severe. Contrary to expectation, female judges were less likely to imprison rapists. Male judges sentenced female offenders more leniently than did female judges. They were less likely to send female offenders to prison, and imposed slightly less severe sentences. In contrast, female judges treated male and female offenders similarly.

Thus, the literature on sex biases in judges' decision-making provides a quite inconsistent picture about the behavior of male and female judges, especially when the gender of the defendant is also considered. The gender differences appear only occasionally, are of varying intensity, though they emerge in civil- as well as criminal-law decision-making. Quite unexplained are the strength and direction of these effects and their causes.

Furthermore, the vast majority of these studies failed to consider the gender of the probation officers who recommend sentences to judges in criminal cases. Yet, in archival research with over 1200 cases, Konečni and Ebbesen (1982) found the probation officers' recommendations and the judges' sentencing decisions to be identical 87% of the time and there was considerable evidence that the causal arrow went from probation officers to judges. This suggests the need to understand a 2 (female vs. male) × 3 (probation officer vs. judge vs. defendant) pattern and nothing is at present known about this (not to mention further possible interactions of this 2 × 3 pattern with the type of crime, strength of evidence, prior criminal record of the defendant, and his/her in-jail vs. out-on-bail status - which factors were identified by Konečni and Ebbesen (1982) as causally affecting the probation officers' recommendation (and only *indirectly* the judges' sentences)).

Studies Examining Gender-Related Issues in Decoding Non-Verbally-Communicated Information

The research on sex differences in decoding non-verbal behavior suggests that females are better at it than males (cf. Hall's (1978) meta-analysis of 75 studies). Furthermore, Rosenthal and DePaulo (1979) supported the hypothesis that non-verbally women are more interpersonally accommodating than men. They are superior decoders and encoders, principally when reacting spontaneously (Fujita, Harper, & Wiens, 1980) and particularly with regard to positive messages (Noller, 1980). A longitudinal and cross-sectional examination of the acquisition of females' superiority in decoding non-verbal cues, indicated what with age, females lost more of their advantage for the less controllable, or more covert channels, but gain more advantage for the less "leaky" channels (Blanck, 1981). In the context of an interview procedure, Christensen and Rosenthal (1982) illustrated the importance of gender and non-verbal decoding skills as determinants of interpersonal expectancy effects in communication. Here, males were more biased by their expectations and produced stronger behavioral confirmation in their communication partners (interviewees) than females, with female interviewees more readily showing "behavior confirmation". More recent studies have attempted to clarify this phenomenon by adding new variables such as measures of personality (Kombos & Fournet, 1985; Schneider & Schneider-Duker, 1984), factors of socialization and attitudes (Noller, 1986), or by methodological criticism (Eisenberg & Lennon, 1983).

In the courtroom, non-verbal communication (facial expression, gestures, body movement, and paralinguistic elements such as variation of voice pitch, speech rate, and loudness) may subtly affect the proceedings and outcome of a trial, yet the participants may be unaware of this.

Hypotheses

The present study examined the relevance of gender in the decoding, integrative and "verdict" decision-making activities by simulated judges to whom civil-law case information was communicated in four different ways: Video and Sound; Video Only; Sound Only; and Transcript Only. The focus was especially on the role of extra-legal emotion-laden information. Specifically, we tentatively hypothesized, on the basis of the literature review, that female decoding and decision-making processes may be disproportionately influenced by intense emotional information, especially when the content and emotionally-colored information are inconsistent (hence the four-communication-channel experimental design).

In part, these hypotheses were based on a study by Kette and Brandstätter (1990) which demonstrated - using the multiple-time-series methodology - that pure emotionally-laden information (entirely excluding legal facts) influenced verdicts and that for females non-verbal cues were a stronger predictor of verdicts than for males.

Our specific hypotheses were: (a) women are better than men at *decoding* non-verbally-communicated information of a predominantly emotional nature in a legal context; (b) therefore, women are more accurate in *predicting* the outcomes of civil-law cases in the Video Only condition; (c) women tend to put more emphasis on non-verbal cues than men do, which leads them to weigh that kind of information to greater extend in arriving at verdicts.

Method

Research participants The research participants were 72 male and 72 female graduating (social science) students from the University of California, San Diego, ranging in age from 22 to 31, with a mean of 25 years. All of them participated as part of their introductory psychology course requirements. The subjects were tested in groups of two to five, with each subject in a group randomly assigned to experimental condition. (Data from a further nine subjects were not analyzed because these individuals had previously viewed parts of the case materials.) There were 18 subjects in each condition of the basic 2 (gender) $\times$ 4 (mode of presentation) design.

Stimulus case materials The stimulus materials were 15 different videotaped civil-court cases which were presented in the long-running TV series "The People's Court". Each "episode" consists of a short introduction (giving the title of the case and a description of the plaintiff's complaints), followed by the presentation of the evidence (testimony by the plaintiff, defendant, and witnesses), and the judge's verdict. The judge was male and in all the cases the same person. In the episodes, the camera typically focused on the courtroom participant who was talking. Therefore, accurate observation of non-verbal cues (both facial and body) could be made by the subjects.

It should be noted that the judge is a real (retired) judge, not an actor, and that the litigants are similarly not actors, but persons involved in a real-life litigation. Finally, the judge's decision is binding, in that a contract to that effect is previously signed by the litigants. In the 15 cases that were used, the range of amounts sued for by the plaintiffs was $24.47 to $1,500.00 (the upper limit being fixed by law for small-claims courts), with a mean of $877.36.

The introduction to a case and a portion of the evidence presentation were used as stimulus materials. The period of evidence presentation (which lasted from four to 12 minutes across the 15 cases used, with an average of seven minutes) was split into segments of 30 seconds each. Each such segment was defined as a "piece of evidence". After each segment, the subjects were given five seconds to rate it on two scales (described below). To prevent a biased selection of the 15 cases from the total pool of episodes, the sequence of the five cases presented to the subjects was determined by the sequence in which the programs had been aired. The 15 cases were divided into three blocks of five cases and balanced for episode length, and the gender and credibility of the plaintiffs.

Video and sound condition was considered as one containing both legally relevant and emotionally relevant information. In the Video Only and Sound Only conditions, information was presented on the respective channels. The Video Only condition was assumed to present exclusively emotional (non-factual) information in terms of behavioral and facial cues. In contrast, the Audio Only condition presumably involved the presentation of paralinguistic emotionally relevant cues, as well as evidentiary information. The Transcript Only condition involved emotionally relevant information only to the extent that the content itself was emotionally colored.

The four conditions thus represented a gradual series of emotionally/substantively relevant information communicated to the subjects. Some of the issues arising from the use of these different communication channels have been previously investigated. For example, facial cues have been shown to convey emotional information to a large degree (Fujita et al., 1980; Rosenthal & DePaulo, 1979; Harris & Rosenthal, 1985). Different studies have shown that visual cues alone can transmit experimenter expectations (Blanck & Rosenthal, 1984). Kraus, Apple, Moency, Wenzel, and Winton (1981) showed that verbal, paralinguistic and visual factors play a role in the judgements of another's affect (see also Ekman, 1980).

Design, procedure and dependent measures A 2 (sex of "judge") × 4 (modes of presentation: Video and Sound, Video Only, Sound Only, Transcript Only) design was used. Each subject was randomly assigned to a block of cases and thus viewed and rated five cases, all in one mode. The subjects were told that the experiment was designed to answer the question of "how accurately a lay person can predict the outcome of civil liability cases when the evidence is presented on different channels". The subject's task was to indicate on a 120 mm rating scale, anchored by 0.00$ and the full amount that the plaintiff was suing for in the case in question, the extent of a defendant's liability for the damages allegedly suffered by the plaintiff. The subjects were instructed to rate quickly and to use all available information.

A measure of the subject's accuracy in predicting the real judge's verdicts was obtained by subtracting the "People's Court" judge's verdict in a particular case from

the subjects' "verdicts". Because of the different demands (in $) in every case the values were transformed into units of the 120 mm graphic rating-scale. A measure of the subject's accuracy in predicting the real judge's verdicts was obtained. An analysis of disparity within gender groups was also carried out. For both of these measures the *last* of the subjects' sequential dollar verdicts in the series (consisting of 30 second segments) was used (in units of the graphic rating scale).

Another dependent variable was a measure of the influence of emotionally-laden (non-legal) information on the two sexes' verdicts using multiple time series analyses in which different channels of information could be conceptualized as independent and dependent variables, respectively.

Results

"Judges'" verdict accuracy Accuracy of an *individual* verdict is defined as the absolute difference (in mm of the graphic rating scale) between the verdict of a particular subject and the actual verdict of the "People's Court" judge. The *accuracy* of a *group* of six male or female subjects is defined as the *disparity* among their verdicts (within a particular case and a single mode of presentation). The individual accuracy measure indicates how much an individual of either sex differs from the real world outcome, the disparity reflects the homogeneity of the verdicts within the two gender groups. It is notable that the correlations between the two measures are higher for females and that they are the lowest in the two conditions that can be thought of as the most difficult conceptually (Video Only and Sound Only).

An analysis of variance of the individual accuracy scores (see Table 1) revealed a main effect of the mode of presentation ($F(3,719) = 2.78, p = 0.04$)[2] which was due to the fact that the Video Only condition resulted in the lowest accuracy: the average absolute deviation - for the two sexes combined - from the "People's Court" judge's verdicts was 43.8 (in mm of the 120 mm graphic rating scale) or 37%. The other three conditions did not differ from each other (the range, for the two sexes combined, was from 40 mm for the Video and Sound condition to 42 mm for the Sound Only condition).

Table 1: Means and standard deviations for accuracy as a deviation from the actual verdict as a function of the sex of "judge" and experiment condition

	Video & Sound	**Video Only**	**Sound Only**	**Transcript**
Female "Judge"				
M	42.7	49.7	45.7	44.1
SD	36.6	41.0	40.5	41.6
Male "Judge"				
M	37.6	52.7	38.9	38.4
SD	36.7	44.9	43.5	37.8

Note: The deviation between the individual verdict and the actual verdict is the absolute difference of the two ratings (in mm of the 120 mm graphic rating-scale). High numbers indicate a larger difference from the actual verdict, that is low accuracy.

Both the main effect of the gender of "judges" ($F(1,719) = 1.48$, p - 0.23) and the interaction ($F(3,719) = 0.55$, p - 0.65) were not significant. The latter was true even though in three of the four conditions, those including verbal information, men's verdicts approximated "People's Court" judge's verdicts more closely than did women's. The only reversal occurred in the Video Only condition, with women more accurate than men.

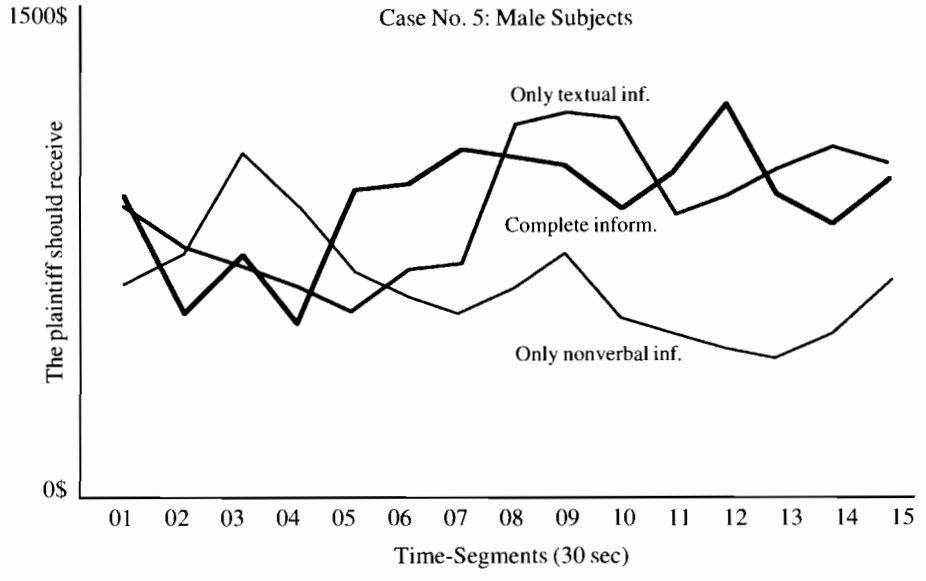
If a significant gender difference were to be obtained in the analysis of *group disparity*, that would imply a gender difference in factual/emotional processing, or differential within-gender variability in response to the type and difficulty of tasks, decision-making strategies, and possibly other factors (our experiment was not designed to distinguish between such differences), or both. Disparity was defined as the standard deviation of the verdicts by male and female subjects, respectively, calculated for each of the 15 cases (120 standard deviations: 2 (gender in groups of six) $\times$ 4 (modes of presentation) $\times$ 15 (cases)). There was, in fact, no main effect of gender, $F(1,119) = 0.10$, $p > 0.05$; the interaction was not significant ($F(3,119) = 1.25$, $p >$ 0.10). However, a marginally significant interaction between these two factors emerged in a theoretically important analysis: 2 (gender) $\times$ 2 (contrast of the three communication channels involving factual information, Video and Sound + Sound Only + Transcript vs. only non-verbal information (Video Only): $F(1,119) = 3.49$, p - 0.06. As can be seen from Table 2, the interaction is due to the considerably greater variability in verdicts reach by men, as opposed to women, in the Video Only condition, in the context of a mode-of-presentation main effect that was due mainly to the Video and Sound condition resulting in the least variability for both men and women, especially compared to the Video Only condition.

Table 2: Means and standard deviations for disparity of verdicts as a function of the sex of "judge" and experiment condition

	Video & Sound	**Video Only**	**Sound Only**	**Transcript**
Female "Judge"				
M	31.6	36.5	35.3	37.4
SD	13.2	8.9	16.1	14.2
Male "Judge"				
M	28.9	46.0	35.6	33.8
SD	13.1	8.0	19.8	11.7

Note: Disparity is the standard deviation of the individual verdicts (in mm of the 120 mm graphic rating-scale) within one cell and one case; that is, the means in the table are calculated from the 15 standard deviations according to the 15 cases.

Weighing of emotional information The central issue of the degree of weighing of emotionally-laden information, especially in the light of the assumption that such information may be differentially used by the two sexes, was approached by time-series analyses. We employed the multiple time-series method (cf. Möbus & Nagl, 1983), such that two "independent" (Video Only and Transcript Only) and one



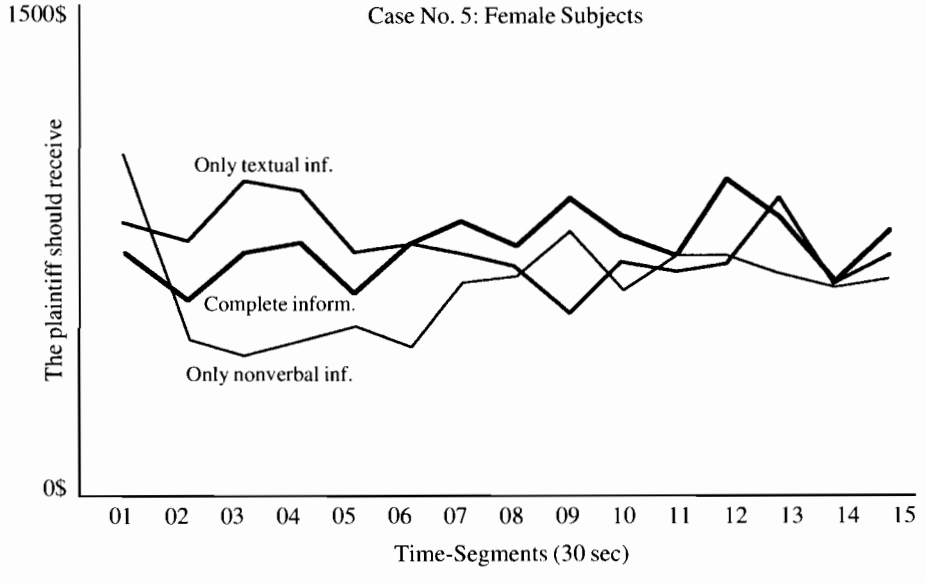


Figure 1: Sample time series for male and female subjects in one case and three presentation modes; independent series: only nonverbal and only textual information; dependent: condition with complete information (video & sound).

"dependent" variable (Video and Sound) were conceptualized.[3] The Video Only condition was construed as providing exclusively emotionally-relevant information, whereas the Transcript Only condition was assumed to provide predominantly factual, legally-relevant information. The dependent-variable channel was the Video and Sound condition which combines both types of information. The data for the time-series analyses were obtained by averaging, for males and females separately, the ratings at each 30-second point, in each of the 15 cases. Another way of putting this is that individual subjects' time-series data in the three channels mentioned above were averaged, separately for each sex.[4]

The statistical procedures used are illustrated in Figure 1 for one of the 15 cases. In the top panel of this figure are presented the averaged data for male subjects' verdicts in the Video Only, Transcript, and Sound and Video conditions across the 15 30-second segments. The bottom panel presents the data for female subjects. In the analysis, the "predictive" power (σ-coefficients) of the two "independent" variables was compared in the data for the two sexes. Whereas in female subjects the non-verbal (Video Only) and textual (Transcript) information are equally good "predictors" of the "dependent" variable, the male subjects' verdicts in the complete-information channel (Video and Sound) were "predicted" only by factual/legally-relevant information (Transcript), as evidenced by the divergence of the Audio Only time-series line. In 11 out of 15 cases, women weigh the non-verbal/emotional information more than did men. The σ-coefficients and the appropriate t-test documenting these statements are presented in Figure 2.

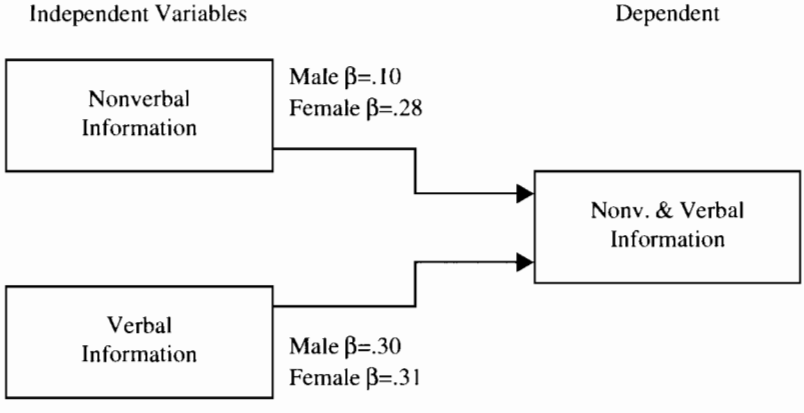


Figure 2: Sample time series for male and female subjects in one case and three presentation modes; independent series: only nonverbal and only textual information; dependent: condition with complete information (video & sound).

Time-series auto-correlations An analysis of the autocorrelations of the 720 (144 subjects rating five cases each) individual time series was also carried out. The Function of Auto-Correlations (ACF) or Function of Partial-Auto-Correlations (PACF, see Mobus & Nagel, 1983) can be interpreted as "memory" of events in the

time series (at time-lags t minus 1, t minus 2, ... t minus n), that is, the memory of previously seen "pieces of evidence" (or previously *given ratings* of those "pieces of evidence", i.e. "consistency": these two interpretations cannot be separated easily). Such data can perhaps clarify any observed gender X mode-of-presentation interactions, in that a different pattern of auto-correlations would be observed in the relevant cells.

High auto-correlations could, of course, also be caused by the nature and order of the introduction of evidence. (For example, if the presented evidence favored one side throughout, then each new piece would cause a move to that side, and there would be a pseudo-trend in the time series.[5]) To separate "memory" effects from the irrelevant, contaminating type-of-case effects, such trends have to be statistically eliminated (by differentiation). If auto-correlations remain statistically significant, this suggests that the previous verdict decisions are "dragged along". One implication of this is that a superior memory for certain types of previously presented information (for example, in the Video Only channel) would substantially influence the subsequent verdict-related decision-making process.

The 720 auto-correlations were used as data in a 2 (sex of subject) × 4 (mode-of-presentation) design: no significant effects emerged. An attempt was then made to focus on the two communication channels which repeatedly differentiated between men's and women's verdicts in other analyses. In a 2 (sex of subject) × 2 (Video Only vs. Transcript Only) design, 360 auto-correlations served as data points. This analysis yielded no main effects, but a statistically significant interaction, $F(1,359) = 3.82, p = .05$. There were lower auto-correlations for men than for women in the Video Only condition, and a reversal in the Transcript condition (see Table 3). This pattern exactly parallels the results obtained with other dependent measures.

Table 3: Means of autocorrelations of the individual time-series of $-amount-ratings as a function of the sex of "judge" and experiment condition

	Video & Sound	**Transcript**
Female "Judge"		
M	.39	.28
SD	.29	.29
Male "Judge"		
M	.33	.34
SD	.31	.33

Note: The autocorrelation (time lag minus 1) indicates the "memory" of the individual (differenced) time series. N=90 (18 Ss per cell, rating five cases each).

Discussion

For men, in contrast to women, facial configurations, gestures and body language of the litigants did not "predict" the decision-making process and its outcomes (verdicts over time; Figures 1 and 2). In addition, women's relative reliance on non-verbal

behavior was not associated with greater decoding accuracy: if anything the data (Table 1) indicate that men were somewhat (though not significantly so) more accurate than women in three of the four mode-of-presentation conditions - all those involving verbal/factual information. To the extent that the litigants were likely to engage in all sorts of non-verbal simulation and dissimulation, women's apparent (conscious and unconscious) greater sensitivity to, and attempts to capitalize on, the non-verbal "leakage" resulted in poorer accuracy.

Furthermore, women were mor homogeneous than men in responding to the information in the Video Only channel (Table 2), yet their greater homogeneity (smaller disparity) was associated with greater accuracy only in the conceptually most difficult (or unreliable) circumstances (Video Only, Sound Only)

There was thus no evidence in our data to suggest that women's relatively more homogeneous, more non-verbally-sensitive judgements led to objectively "better" decisions - in terms of fairness or equity. In fact, one could conclude that their reliance on emotionally-colored information (presumably for the most part conveyed non-verbally) was "unfair" to the litigants. However, such a statement is based on the assumption that the "People's Court" judge's verdicts are "fair" by some objective criterion - a tentative supposition indeed. Besides, the lone "People's Court" judge is male and it is entirely possible that female reliance on non-verbal information would bring them closer to the verdicts of female judges.

Nevertheless, the observed sex differences should be explored in further research to examine the possibility of differential utilization of emotion- vs. content-based information in legal contexts. Such research, directed specifically at the relative weights given to emotional expression as opposed to verbal content (whether true or false) might explain the differences between female and male jurors (and perhaps judges) noted in the introduction.

The fact that measures of accuracy in terms of a quasi-real world legal outcome were obtained in the present study is significant because of the existing literature that suggests that, in general, people rely more on non-verbal cues than verbal content in certain situations, including legal contexts, especially when, for example, the tone of voice and verbal contents are in conflict (e.g. Domangue, 1978; Efran, 1974; Ekman & Friesen, 1974; Kulka & Kessler, 1978; Littlepage & Pineault, 1981; Mehrabian & Wiener, 1967). To the extent that one is willing to grant the "People's Court" judge the status of an objective and fair (or at least highly trained) arbiter, the juror's reliance on non-verbal or emotional input (provided by the accused, the victim, the witnesses or civil-law litigants) is perhaps detrimental to fairness and justice (cf. Maier and Thurber, 1968). If so, the question is whether this state of affairs can be improved (for example, by instructions to the jury).

With regard to the essential Video Only vs. Audio Only vs. Transcript contrasts, one might note that the symbol of justice stands with a sword in her right hand, scales in her left, and a blindfold: she can only hear.

References

Blanck, P. D. (1981). Sex differences in eavesdropping on nonverbal cues: Developmental changes. *Journal of Personality and Social Psychology, 41*(2), 391-396.

Blanck, P. D., & Rosenthal, R. (1984). The mediation of interpersonal expectancy effects: The counselor's tone of voice. *Journal of Educational Psychology, 76*, 418-423.

Christensen, D., & Rosenthal, R. (1982). Gender and nonverbal decoding skill as determinants of interpersonal expectancy effects. *Journal of Personality & Social Psychology, 42*(1), 75-87.

Domangue, R. (1978). Decoding effects of cognitive complexity, tolerance of ambiguity, and verbal-nonverbal inconsistency. *Journal of Personality, 46*, 519-522.

Efran, M. G. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *Journal of Research in Personality, 8*, 45-54.

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin, 94*(1), 100-131.

Ekman, P. (1980). Relative importance of face, body and speech in judgements of personality and affect. *Journal of Personality and Social Psychology, 38*, 270-277.

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology, 29*, 288-298.

Fujita, B. N., Harper, R. G., & Wiens, A. N. (1980). Encoding-decoding of nonverbal emotional messages: Sex differences in spontaneous and enacted expressions. *Journal of Nonverbal Behavior, 4*(3), 131-145.

Gruhl, J., Spohn, C., & Welch, S. (1981). Women as policymakers: The case of trial judges. *American Journal of Political Sciences, 25*, 308-322.

Guilford, J. P. (1954). Psychometric Methods (2nd ed.). New York: McGraw.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin, 85*(4), 845-857.

Harris, A., & Rosenthal, R. (1985). The mediation of interpersonal expectancy effects. *Psychological Bulletin, 97*, 363-382.

Kette, G. (1994). *Die Signalerkennengstheorie als Modell der laienrichterlichen Urteilsbildung.* Wien: Springer.

Kette, G., & Brandstätter, H. (1990). Times series analysis of a murder trial. *Journal of Social Behavior, 5*, 15-31.

Kombos, N. A., & Fournet, G. P. (1985). Effects of dominance-submissiveness and gender on recognition of nonverbal emotional cues. *Educational and Psychological Research, 5*(1), 19-28.

Konečni, V. J., & Ebbesen, E. B. (1982). An analysis of the sentencing system. In V. J. Konečni & E. B. Ebbesen (Eds.) *The criminal justice system: A social-psychological analysis.* San Francisco: W. H. Freeman.

Kraus, R. M., Apple, W., Moency, N., Wenzel, C., & Winton, W. (1981). Verbal, vocal and visible factors in judgements of another's affect. *Journal of Personality and Social Psychology, 40*, 312-320.

Kritzer, H. M., & Uhlmann, T. M. (1977). Sisterhood in the courtroom: Sex of judge and defendant in criminal case disposition. *Social Science Journal, 14*, 366-379.

Kulka, A., & Kessler. (1978). Is justice really blind? The influence of litigant physical attractiveness on juridical judgement. *Journal of Applied Social Psychology, 8*, 366-379.

Littlepage, F., & Pineault. (1981). Detection of truthful and deceptive interpersonal communications across information transmission modes. *Journal of Social Psychology, 114*, 57-65.

Maier, E., & Thurber, G. (1968). Accuracy of judgements of deception when an interview is watched, heard and read. *Personnel Psychology, 21*, 23-40.

Mehrabian, A., & Wiener, M. (1967). Decoding of inconsistent communications. *Journal of Personality and Social Psychology, 6*, 108-123.

Möbus, C., & Nagl, W. (1983). Messung, analyse und prognose von veranderungen. In J. Bredenkamp & H. Feger (Eds.), Hypothesenprüfung. *Enzyklopädie der Psychologie* B15 (pp. 239-470). Göttingen: Hogrefe.

Myers, M. A., & Talarico, S. M. (1987). *The Social Context of Criminal Sentencing*. New York: Springer.

Noller, P. (1980). Misunderstanding in marital communication: A study of couples' nonverbal communication. *Journal of Personality and Social Psychology, 39*(1-6), 1135-1148.

Noller, P. (1986). Sex differences in nonverbal communication: Advantage lost or supremacy regained? *Australian Journal of Psychology, 38*(1), 23-32.

Rosenthal, R., & DePaulo, B. M. (1979). Sex differences in eavesdropping on nonverbal cues. *Journal of Personality and Social Psychology, 37*(2), 273-285.

Schneider, J., & Schneider-Duker, M. (1984). Sex roles and nonverbal sensitivity. *Journal of Social Psychology, 122*(2), 281-282.

1. This study was sponsored by the Österreichischer Forschungsfond Project Nr.: J70 to Gerhard Kette and by various research grants at UCSD to Vladimir J. Konečni. We thank Mag. Gernot Fillip for assistance in the statistical analysis and also Professor Hermann Brandstätter for a helpful discussion of the methodological and theoretical questions of the study. Correspondence should be addressed to Gerhard Kette at the Department of Psychology, University of Linz, A-4040 Linz, Austria or Vladimir J. Konečni at the Department of Psychology, University of California, San Diego, La Jolla, California 92093, USA.
2. Mixed-design ANOVA: 18 Ss rating five episodes each in each cell of the 2 (gender) × 4 (mode of presentation) design.
3. As an indication of the reliability of ratings, the average intercorrelation was used, which is known as the intraclass correlation (Guilford, 1954, p. 395). The reliability of the ratings in all four modes is acceptable. The reliability of judgements of one rater ranged from, r 1/1=.30 to .54, while the reliability of the averaged judgements of the nineteen rater per cell ranged from r 18/18=.61 to .90. Therefore it was possible to integrate the individual time series into one group time series.
4. Since trends in this type of data may cause artificially high correlations between the "independent" and "dependent" times-series, reflecting nothing more interesting than certain case and evidence-presentation features (e.g. cases may differ in how gradually the pro-plaintiff evidence is introduced), such trends were statistically removed - a procedure common in multiple time-series methodology.
5. We attempted to minimize an unduly large effect of each new piece of information by printing the following instruction above the rating scales: "Considering all available information thus far, the plaintiff will receive:".