Ecological validity test of laboratory studies of information integration

*Eric Fruchart** (Marne-la-Vallée), *Patricia Rulence-Pâques*** (Boulogne), & *Etienne Mullet**** (Toulouse)

Hundreds of laboratory studies have to date been carried out examining how human beings use and combine informers to make a judgment. Many of these studies were conducted using Information Integration Theory (Anderson, 1996). Within the framework of this theory, various judgment schemata have been evidenced. These judgment schemata have been found by different teams doing research on different populations. These schemata are robust. One can cite the blame schemata as a particularly robust example (Anderson, 1991a, 1991b; Hommers & Anderson, 1991).

Blame schemata suggest that when an observer judges the degree of blame to be attributed to a person who has committed a negative act, she takes into account the degree of intent and the severity of the consequences of this act. In most cases, intent of the act and severity of the consequences are integrated by using an additive rule: Blame = Intent + Consequences (Anderson, 1991a, 1991b). This schemata is very general and has been observed even in 5-year-old children (Leon, 1980, 1984).

Ecological validity of laboratory studies

Calling into question the ecological validity of the judgment schemata drawn from observations carried out in laboratory conditions amounts to asking whether the same schemata would have been observed if the real reactions of the persons directly involved had been studied and compared. In the case of blame schemata, the question would be to know whether, after having observed a negative act that had really been committed by a person, an observer (a) really takes into account the degree of intent of the act as

* Department of Physical Education, University of Marne-la-Vallée.
** Department of Physical Education, Cote d'Opale University.
*** Ethics and Work Laboratory, Ecole Pratique des Hautes Etudes.

perceived through the author's declarations or other indications, (b) really takes into account the severity of the consequences of this act evaluated from the observed real damage or from the victim's declarations, and (c) combines these two elements in a strictly additive way to judge the degree of blame.

The way in which people integrate information on intent and on the severity of consequences to make a judgment on the degree of blame, is relatively easy to study in laboratory conditions. The minimum requirements are the creation of a set of cards with both elements of information and a rating scale. If two levels of intent (intent to damage, no intent to damage) and three levels of consequences severity (low, rather high, very high) are considered, then six cards (six scenarios) are sufficient to provide the whole set of combinations.

The way in which people integrate information on intent and on consequences severity to make a judgment on the degree of blame is much more difficult to study outside the laboratory. In real life conditions (family life, work environment) it is necessary to make a large number of observations about situations where a negative act was committed and for which reliable information is available about (a) the author's intent, (b) the severity of the consequences of the act for the victim, and (c) the level of blame attributed by an observer who is really involved in the situation. Furthermore, these situations must be homogeneous enough for their categorization to be feasible and meaningful.

In addition, the process of comparing results observed in laboratory conditions with results observed in everyday life presents its own difficulties. To make direct comparisons, it is necessary to categorize the multiple everyday life observations in terms of both level of intent and level of consequences severity. Ideally, in the example, it would be necessary to categorize the multiple observations in the six sets corresponding to the six scenarios used in the laboratory.

The categorization process can only be approximate considering (a) the multiplicity and the complexity of the observed real life situations and (b) the variety of actors (perpetrators and victims). In particular, we can always have doubts in such a case about the real orthogonality of the two variables, even if all necessary precautions were taken to realize an adequate categorization. Furthermore, it is necessary to ensure that the measurement of the attributed level of blame possesses interval scale properties. In most of the cases, this last constraint is practically insurmountable.

Early tests of the ecological validity of laboratory studies of information integration

The numerous difficulties mentioned above explain why no study has so far been carried out to relate together and compare patterns of judgment observed in the laboratory and patterns of behavior observed in real life situation. Authors examining the ecological validity of the laboratory observations have adopted more or less indirect approaches (Levin, Louviere, Schepanski, & Norman, 1983).

For example, Lerman and Louviere (1978) compared (a) attractiveness judgments for different living places based on the size of the town (and the facilities provided) and the commuting time from home to place of work, and (b) the percentage of persons recently arrived in the town and working in a given place. In the first phase of their study, students and teachers were asked to evaluate the degree of attractiveness for 36 living places characterized by the size of the town and its facilities (6 levels) and the time required to go to work (6 levels) using a continuous scale of attractiveness. The pattern of judgments was shown to be a non-additive pattern (interactive pattern). The effect of the size of the town appeared stronger when transport time was short than when transport time was long.

In the second phase of their study the judgment model, which has just been described, was used to evaluate the level of attractiveness of several towns situated just outside an industrial complex recently built in a rural area. Finally, in the third phase of their study the attractiveness levels that were calculated were linked to official statistics on the location of homes of workers newly recruited in the industrial complex of these towns. The correlation between the two series of values was 0.92. Furthermore, none of the econometric models usually used in this kind of study was able to predict the effective place of residence of workers with as much precision as the model derived from the analysis of judgments made in the laboratory.

Several studies of this type were published in the 80's – for a review, see Levin et al. (1983) and Louviere (1988). All these studies conclude that analysis of people's judgments under laboratory conditions produces valid information because this information enables one to predict better than corresponding econometric models, certain types of behavior that have actually been observed (e.g., consumption, habitat, choice of mode of transport).

The Present Study

In comparison with the studies just mentioned, the present study had the ambition to directly test the ecological validity of the results observed

in the laboratory. The field of human activity that had been chosen was collective sport. This domain had been chosen because it presented a certain number of particularities which made it possible to make a direct connection between judgments observed in the laboratory and behavior in everyday life (Rulence-Pâques, Fruchart, & Mullet, 2005).

On the one hand, it is usual to record entire matches. These recordings are made for both important and less important matches. Their function is to enable players to observe their past behavior and possibly to correct it in the future. They are the basis for discussion between trainers and players. In handball and in football there therefore exist important video databanks, which are available to the experts in these sports.

On the other hand, an important advantage in situations found in collective sport is that these situations are standardized. The action always takes place on the same type of ground. The players are young adults in good physical shape. Rules of the game are clear. It is possible to know without ambiguity the parameters for various situations in a match. For example, it is possible to know precisely whether a specific action was carried out at the beginning of the match or a few minutes before the end of the match, or whether the opposing team had less or the same amount of players on the pitch, or whether the team of the player in question had a superior or lower score than the opposing team's score. The actions that are carried out are of a type that can easily be categorized (e.g., shot in goal, quick restart of play). And so, as a result of the natural standardization of situations in collective sport and as a result of the availability of a large amount of video archives, many of the difficulties mentioned previously (interpretations of the indicators, scale linearity) are greatly lessened.

In the present study, 20 professional handball players evaluated the frequency of a certain game strategy used at the end of a match – quick restart of play – using three variables: numerical status of the team (inferiority, equality, superiority), the score at the moment in the match (the team was winning, drawing, or losing), and the time left to be played (little time, very little time). An example taken from the 18 scenarios used is as follows: "During a championship match, a goal has just been scored by the opposing team. You must restart the game. Your team has less players than the opposing team's. The opposing team is leading by one goal. There are thirty seconds left to play. In this situation, is it usual to proceed with a quick restart of play?". The reply to the question is given by using a continuous scale ranging from "Not at all usual" to "Completely usual".

At the same time that the laboratory study was carried out, 200 video recordings randomly selected were analyzed, either showing or not showing a quick restart of play at the end of a match. Each situation where a goal
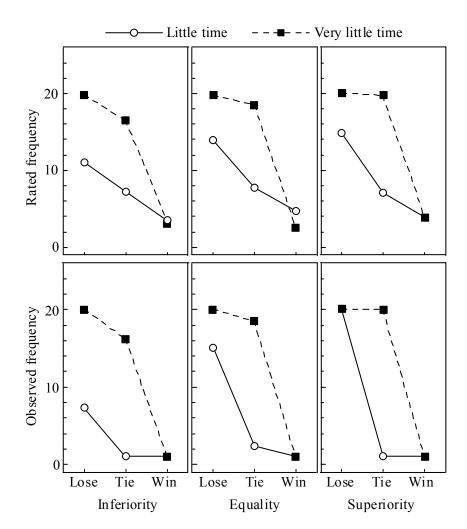
Figure 1. Patterns of judgments about the frequency of quick restarts of play as a function of score, composition of the team, and time left to play (top). Patterns of observed quick restarts of play as a function of the same circumstances (bottom).

was scored at the end of the match was used as the starting point of the analysis and each subsequent action (quick restart of play or not) was noted as well as the circumstances of this action (time left to play, numerical status, score). The percentage of the quick restart of play was calculated for

each of the 18 possible circumstances (and divided by five for making the comparison easier).

Figure 1 compares the patterns of results obtained in each of the two conditions. First of all, in the set of six panels, the curves go down rightward: in other words, an effect of the score factor is present. A quick restart of play is not considered as frequent or is not done at all unless the team is losing. Next, in each of the six panels, the dashed curve (Very Little Time) is higher than the solid curve (Little Time). In other words the factor Time left to play also plays an important role. A quick restart of play is above all considered usual or is above all carried out when there is very little time left to play. It must also be noted that the curves in the six panels are always at the same level in relation to the vertical axis; that is, the effect of the factor Numerical Status of the team is weak or even non existent whatever is the condition. Finally, in each of the six panels, the two curves are always very separated from one another at the level that corresponds to a draw. In other words it is when the two teams are drawing that the factor Time left to play appears to be absolutely essential. In the two other cases (winning or losing) its importance is less, indeed non existent. An analysis of variance showed that the main effects of score and time, and the score $\times$ time interaction were significant.

One can say that, globally, the results observed in the judgment condition are compatible with the results observed in the ecological condition. A correlation coefficient calculated on the two series of 18 values was 0.97.

Conclusion

The present study has illustrated the fact that (a) when necessary precautions are taken to ensure the scale linearity of responses in the judgment conditions (use of a paradigm known to produce linear responses) and in the ecological condition (calculation of percentages of behavior of a specific type), (b) when necessary precautions are taken to ensure the total intelligibility of the circumstances of observation (situations governed by clear rules that are, if possible, standardized), and (c) when the laboratory situation and the observation of real-life events are based on the same realities (judgment of the frequency of a given behavior and observation of this frequency), then it is possible to directly assess the compatibility of laboratory data and real life data.

In the present study, a reasonable degree of compatibility between a judgment schemata derived from a study in the laboratory and a judgment schema inferred from observations in real-life situations has been shown.

Future studies are needed to assess to degree to which such compatibility is or is not the rule.

References

Anderson, N. H. (1991a). Moral-social development. In N. H. Anderson (Ed.), *Contributions to information integration theory. Vol. III* (pp. 137-187). Hillsdale, NJ: Erlbaum.

Anderson, N. H. (1991b). Psychodynamics of everyday life: Blaming and avoiding blame. In N. H. Anderson (Ed.), *Contributions to information integration theory. Vol II* (pp. 243-275). Hillsdale, NJ: Erlbaum.

Anderson, N. H. (1996). *A functional theory of cognition*. Hillsdale, NJ: Erlbaum.

Hommers, W., & Anderson, N. H. (1991). Moral algebra of harm and recompense. In N. H. Anderson (Ed.), *Contributions to information integration theory: Vol. II* (pp. 101-141). Hillsdale, NJ: Erlbaum.

Leon, M. (1980). Integration of intent and consequences information in children's moral judgments. In F. Wilkening, J. Becker, & T. Trabasso (Eds.), *The integration of information by children*. Hillsdale, NJ: Erlbaum.

Leon, M. (1984). Rules mothers and sons use to integrate intent and damage information in their moral judgments. *Child Development, 55,* 2106-2113.

Lerman, S. R., & Louviere, J. J. (1978). On the use of direct utility assessment to identify the functional form of utility and destination choice models. *Transportation Research Record, 673,* 78-86.

Levin, I. P., Louviere, J. J., Schepanski, A. A., & Norman, K. L. (1983). External validity of laboratory studies of information integration. *Organizational Behavior and Human Performance, 31,* 173-193.

Louviere, J. J. (1988). *Analyzing decision making: metric conjoint analysis.* Newbury Park, CA: Sage.

Rulence-Pâques, P., Fruchart, E., Dru, V., & Mullet, E. (2005). Algebraic rules in sport decision-making. *Theory and Decision, 58,* 387-406.

Abstract

The ecological validity of the judgment schemata observed in the laboratory using the functional measurement methodology was directly tested. The field of human activity that has been chosen is collective sport. It was shown that (a) when necessary precautions are taken to ensure the scale linearity of responses in the judgment conditions and in the ecological condition, (b) when necessary precautions are taken to ensure the total intelligibility of the circumstances of observation (situations governed by clear rules, as in collective sports), and (c) when the laboratory situation and the observation of real-life events are based on the same realities

(judgment of the frequency of a given behavior and observation of this frequency), then it is possible to directly assess the compatibility of laboratory data (judgment schemata) and real life data (corresponding patterns of behaviors).

Riassunto

La validità ecologica degli schemata di giudizio osservati in laboratorio tramite la metodologia della misurazione funzionale è stata controllata direttamente. Il campo di attività umana che è stato scelto è lo sport collettivo. È stato mostrato che (a) quando le necessarie precauzioni vengono prese per assicurare la linearità della scala delle risposte nelle condizioni di giudizio e nella condizione ecologica, (b) quando le necessarie precauzioni vengono prese per assicurare la totale intelligibilità delle circostanze di osservazione (situazioni governate da regole chiare, come negli sport collettivi) e (c) quando la situazione di laboratorio e l'osservazione di eventi di vita reale sono basati sulle stesse realtà (giudizio della frequenza di un dato comportamento e osservazione di questa frequenza), allora è possibile determinate direttamente la compatibilità dei dati di laboratorio (schemata di giudizio) con quelli della vita reale (corrispondenti configurazioni di comportamenti).

Addresses. Eric Fruchart, Department of Physical Education, University of Marne-la-Vallée, F-77454 Marne-la-Vallée, France (fruchdk@wanadoo.fr.). Patricia Rulence-Pâques, Department of Physical Education, Cote d'Opale University, F-62327 Boulogne, France (patricia.rulence@wanadoo.fr). Etienne Mullet, Ethics and Work Laboratory, and Laboratory CLLE-LTC (UTM, CNRS, EPHE), F-31058 Toulouse, France (etienne.mullet@wanadoo.fr).