감사합니ㅏㄷ

Weak influences in socio-cultural sphere: Method to measure consequences of contacts with art

Vladimir M. Petrov and **Lidia A. Mazhul**
State Institute for Art Studies
Apt. 11, dom 39, Zemlyanoy val, Moscow 105120, Russia
Tel / Fax: +7 495 9164453; E-mail: vladmpetrov@yandex.ru

KEYWORDS : Social impact, Art, Causality, Personal development, Cultural politics

When measuring various kinds of social interactions, we usually deal with two principal obstacles: low value of the measurand and the multiplicity of parameters which produce their impacts together with the influence to be estimated. For instance, impacts of contacts with poetry (into various potentials of personality) reveal strong links with impacts of contacts with music, as well as 'noise' caused by numerous socio-demographic parameters, etc. For such purposes, special method was derived, consisting in dividing the entire massif of respondents into pairs with close values of all parameters to be eliminated. The 'influencing' parameter **X** *(e.g., frequency of contacts with the given kind of art) and the 'output' parameter* **Y** *(e.g., the index of moral development of personality) are supposed to be linked for each* **i***-th respondent:* $\mathbf{Y}^{(i)} = \mathbf{Y}_0^{(i)} + \mathbf{a}_{XY}\,\mathbf{X}^{(i)}$, *where* $\mathbf{Y}_0^{(i)}$ *is the impact of all parameters to be eliminated. The 'zest' of the method is compiling pairs possessing minimal 'integral difference' between the eliminated parameters. Beside the value of 'isolated influence,' sometimes the causality can also be determined – by comparing values* $\mathbf{a}_{XY}$ *and* $\mathbf{a}_{YX}$*. One of typical results obtained concerns influence of theatre-going upon moral potential of personality: 13%-growth per each additional contact. The results obtained can be used in the practice of cultural politics.*

1. Introduction. General contours of the method derived, and fields of its application

There exist various fields which reveal the need to measure rather weak influences. Most of such fields belong to the sphere of human sciences, mainly sociology and psychology. (Because, for instance, in the sphere of engineering we usually deal with certain structures possessing quite definite destination, so here primarily strict, 'strong' influences should be taken into account.) Examples of such weak influences are the following:

– impact of decorative elements into the scales of consumption of some goods (including impacts of package, advertising, etc.);

– impact of theoretical knowledge into the activity in different fields (including social life, engineering, art, etc.);

– impact of cultural 'consumption' into some kinds of behavior and/or mentality (e.g., influence of contacts with music upon the socio-moral development of personality).

Two main peculiarities are inherent in most of such situations:

(*) small value of the impact to be measured; sometimes this value is about 1-3% or even less;

(**) multiplicity of parameters which produce their impacts together with the influence to be measured; for instance, when measuring impact of the personality's contacts with music (into various potentials of personality), we usually observe strong links between this impact and impact of contacts with poetry and painting, as well as 'noise' caused by socio-demographic parameters, etc.

These peculiarities are typical primarily for socio-cultural sphere. Nevertheless, the below method derived for measuring weak influences in this sphere, can be also applied in other spheres, including technical ones – when the above specific features are weighty for the phenomena to be estimated. However, all further description will appeal to the sphere of sociology; hence, appropriate terms will be used.

2. "Isolated" influences: Method of "binary comparisons"

As far as traditional regression analysis occurs not effective in such situations, special "method of binary comparisons" was derived, consisting in dividing the entire sample of respondents into pairs with close values of all parameters to be eliminated. Then for each pair of respondents (***i, j***), the difference between its members over their 'input parameter' ***X*** (e.g., frequency of regular contacts with music) is referred to their difference over 'output parameter' ***Y*** (e.g., the value of the index of socio-moral potential [1-3]). For each respondent, linear measurement model is supposed:

$$Y^{(i)} = Y_0^{(i)} + a_{XY}X^{(i)}, \quad (1)$$

$Y_0^{(i)}$ being the impact of all parameters to be eliminated, a being the 'strength' of the influence (of X upon Y) to be measured. (This 'strength' is supposed to be 'common' for all respondents.)

As a result, for each pair of respondents (i, j), the quantitative estimation of the influence studied can be obtained:

$$a_{XY}^{(i,j)} = [Y^{(i)} - Y^{(j)}] / [X^{(i)} - X^{(j)}]. \quad (2)$$

Such procedure being applied to all the pairs of respondents, permits to measure the average value of the influence in question: a_{XY}. Hence, the key problem is to form a set of due pairs of respondents, which would be close to each other on many parameters.

Because of specific requirements to such measurements [4], the following conditions are supposed to be valid for the variables involved:

– both 'input parameter' (i.e., the parameter which is supposed to be the 'cause') and 'output parameter' (i.e., the parameter supposed to be the 'effect') are measured in the ratio scale;

– all other parameters, i.e., those ones to be eliminated, are measured in nominal scales (e.g., the scale for the place where the respondent lives, can possess three gradations: big city, small town, and village).

Evidently, when the number of parameters to be eliminated, is great, then it would be almost impossible to form a set of pairs with 'exact coincidence' over all these parameters. Really, if it is needed to eliminate the influence of 6 parameters. possessing 2, 5, 4, 3, 8, and 7 gradations, the number of their combinations would be 2×5×4×3×8×7=6720. Taking into account certain correlations within the matrix of distribution of respondents over these gradations, we should expect the number of respondents needed to fill up most the cells of this matrix, to be about 70 000. (And even in this case, a large share of respondents will occur 'unpaired,' i.e., without due member of the pair, hence, the results obtained would be distorted.) So it seems unreal to form the pairs in question, requiring exact coincidence of the members over all parameters to be eliminated.

That is why we should resort to the help of so-called 'soft equalizing,' i.e., such a procedure which presupposes only approximate coincidence of the members' gradations: some discrepancies on separate eliminated parameters are admissible – in order to provide minimal integral distinction (over all the parameters eliminated). To do this, it seems reasonable to base such 'soft equalizing' on the values of distinction between each two gradations involved, concerning their link with average values of the 'influencing variable' ('input' one).

In other words, when considering each given n-th parameter to be eliminated, the more the difference – over the values of the 'influencing' parameter X – between a certain β-th gradation characterizing respondent i – and ε-th gradation characterizing respondent j, the more 'weighty' should be the difference between these two respondents. It means that we should first of all calculate these averaged (over all respondents) values X_β and X_ε , and then it would be possible to estimate the inter-respondent difference over this eliminated parameter:

$$\delta_{\beta,\varepsilon}^{(i,j)} = \left| X_\beta - X_\varepsilon \right|. \quad (3)$$

For instance, let our 'input variable' be the frequency of contacts with music, and we try to compare two respondents (i, j) eliminating such β-th variable as their age, and the scale for age consists of 10 gradations (0-5 years; 6-12 years; 13-19 years; 20-25 years; 26-30 years, …). Let i-th respondent belongs to the third gradation (13-19 years old), and j-th respondent to the fifth gradation (26-30 years). Let the average frequency of listening to music for all respondents belonging to the third gradation, be 20 contacts per year, and for all respondents belonging to the fifth gradation – 30 contacts per year. Then the distinction between the i-th respondent and the j-th one (over this eliminated parameter, i.e., age) equals 30–20 = 10 contacts. After that let's compare these two respondents (i, j) over the next parameter to be eliminated (e.g. profession); for this, we should know average values of listening to music for representatives of different professions, and so forth.

The resulting 'integral difference' (over all n eliminated parameters) between each two respondents (i, j) equals

$$\Delta^{i,j} = \sum_{n} \delta_{\beta,\varepsilon}^{i,j}. \quad (4)$$

Namely this 'integral difference' should be used when compiling due pairs of respondents.

So, the algorithm of calculating the 'isolated influence' of the parameter X upon the parameter Y looks as follows:

A. For each gradation of n parameters to be eliminated, the average value – over the entire sample of respondents – of the 'influencing' parameter X is calculated: X_β, X_γ, X_ε, and so on.

B. In the sample of respondents, a certain i-th one is chosen by chance. Then for this respondent, his/her 'integral difference' Δ from all other respondents, is calculated, using formula (4). Out of all respondents, such l-th respondent is singled out, which reveals minimal value of difference $\Delta^{i,l}$ from the i-th respondent. (If there occur several such 'minimalistic' respondents, one of them is chosen by chance.) This pair (i, l) is considered to be formed, and both members are withdrawn from the sample. Out of remaining respondents, again a certain j-th respondent is chosen by chance, and using the above procedure, the most close to him/her m-th respondent is found. Then this pair (j, m) is also withdrawn from the sample, then the procedure is repeated many times – till the moment, when all the sample occurs divided into pairs (if the number of respondent is even) or one respondent remains 'unpaired' (if the number of respondents is odd).

C. Out of all the pairs formed, a certain share is withdrawn (e.g., 25%) – those pairs which reveal rather great values of 'integral difference' Δ.

D. Out of a set of remaining pairs, those ones are withdrawn where the members do not differ on the values of the 'influencing' parameter X, i.e., the pairs with $X^{(i)} = X^{(j)}$.

E. For each of remaining pairs, the coefficient $a_{XY}^{(i,j)}$ is calculated using formula (2).

F. All the values $a_{XY}^{(i,j)}$ are averaged over all the pairs. The result of averaging is nothing else than the coefficient a_{XY} characterizing the 'isolated influence' of parameter X upon parameter Y.

G. The calculations following the above algorithm (***B*** – ***F***) are repeated several times (usually 15-20 times), each time proceeding from another choice of the first (i-th) respondent, in order to exclude the error caused by the first choice, as well as consequent choices.

H. The values a_{XY} obtained in these calculations, are averaged over all the variants (and appropriate standard deviation is also calculated), and we receive the final value of the influence studied.

This measurement procedure was used in several sociological investigations fulfilled in different regions of Russia and former USSR. It seems reasonable to present some of the results obtained, in order to illustrate the potentialities of the method derived, meaning first of all the possibility to eliminate the influence of rather large number of ‘collateral’ parameters, however dealing with rather small samples of respondents.

In one of such investigations [5, 6] a fragment of empirical data was realized, concerning industrial workers of two towns in Samara Region – in total 425 respondents. The ‘input variables’ (‘influencing’ parameters) were frequencies of various kinds of the personality’s contacts with art (per year); the ‘output variable’ was the labor activity of respondents, which was measured in the ratio scale – as percentage of fulfilling month plans. So, both main parameters were measured in the ratio scale. As for the variables to be eliminated, two versions of calculations were realized:

– elimination of 6 principal socio-demographic parameters: gender (2 gradations), age (7 gradations), education (6 gradations), presence of little children in the family (2 gradations), duration of living in the given town (5 gradations), and educational level of the respondent’s parents (6 gradations);

– elimination of the above 6 socio-demographic parameters, plus 14 kinds of cultural consumption (of course, excluding that kind of contacts with art, which was investigated, e.g., the frequency of theatre-going).

As a rule, both versions of calculations came to rather close results, though the first version usually responded to slightly higher values of the coefficient a_{XY} (15-20% higher than the values for second version), which seems to be natural – because of links between various kinds of cultural consumption, as well as their links with socio-demographic variables. However, this difference is commensurable with the error of measurements, so further we shall present the results averaged over both versions of calculations.

Strong link with labor activity was established for such kind of cultural consumption, as theatre-going: a_{XY} = .012. It means that each additional annual visiting theatre, is accompanied with 1.2%-growth of labor activity. Nevertheless, such results (as well as the below ones) cannot be treated as ‘absolute’: firstly, here not one occasional act of theatre-going is meant – but regular additional act (inherent in the style of life of the given respondent); secondly, the measurement model is based on the supposed linear relationships between the variables studied, but in reality this dependence may occur non-linear and even non-monotonic (moreover, usually such dependences show certain maximum, i.e., optimal frequency of the personality’s contacts with art). Hence, the results obtained describe only the behavior of most respondents, belonging to the most spread range of theatre-going (from zero to 3-4 visits per year).

Approximately the same was the strength of the influence of visiting artistic museums and exhibitions: each additional annual visit is accompanied with 1%-growth of the labor activity. As for visiting libraries, each additional book received, responds to .02%-growth of labor activity. Each additional annual variety show going is accompanied with 2.5%-growth of labor activity, whereas each additional participation in meetings of circles of amateur art responds to .15%-growth of labor activity.

Another investigation dealt with the influence of cultural consumption upon the socio-moral development of the population of Kostroma (sample of 176 respondents representing all adult population of the town [7]). In the questionnaire there were 20 indicators reflecting mental life and worldview of the respondent; these indicators were aggregated in the ‘index of socio-moral development’ of personality [1-4]. This ‘output variable’ was supposed to depend on such ‘input variables’ as frequencies of the personality’s contacts with different kinds of art. As for variables to be eliminated, they were four: gender (2 gradations), age (13 gradations), education (7 gradations), and educational level of the respondent’s parents (7 gradations).

Strong link was revealed for theatre-going: a_{XY} = .13, i.e., each additional regular (annual) theatre-going is accompanied with 13%-growth of the index studied. (Of course, all the above reservations are valid also in this situation.) Each additional variety-show going responds to 9%-growth of the index, and each home listening to recorded music – .2%-growth. Hence, socio-moral development of personality experiences much more strong influence of contacts with art than it takes place in the case of labor activity.

3. Causality: direction of the links observed

Another task is to establish causal orientation of links observed, meaning to clear up which parameter dominates in each link: X or Y? (In other words, it is necessary to determine if listening to music results in a certain growth of socio-moral development of personality, or perhaps, this development causes listening to music?) We shall speak of ‘soft causality,’ meaning the degree of adequacy of the hypothesis of determination of the value of Y by the value of X – or, on the contrary, quite opposite determination takes place: the value of X is determined by the value of Y? Or perhaps, none of these two variables is dominating in their link?

To realize rough proving the adequacy of any such hypothesis, it seems reasonable to divide all the respondents into two groups (desirably equal or almost equal, i.e., using median ‘splitting’):

– respondents with low values of parameter X, responding to its mean value X_{LOW} ;

– respondents with high values of parameter X, responding to its mean value X_{HIGH} .

Then due to immediate influence of X upon Y, the growth of Y when turning from the first group of respondents to the second one, would be

$$P^{Y}_{IMM.\,INFL.} = (X_{HIGH} - X_{LOW})\, a_{XY}\,. \qquad (5)$$

Meanwhile, we have the value of real growth of Y when turning from the first group to the second one:

$$P^{Y}_{REAL} = Y_{X\text{–}HIGH} - Y_{X\text{–}LOW}\,, \qquad (6)$$

$Y_{X\text{–}HIGH}$ and $Y_{X\text{–}LOW}$ being mean values of Y for respondents belonging to the above ‘high’ and ‘low’ groups, respectively.

The discrepancy between these two values of growth: real one and supposed one – can be treated as the degree of adequacy of the hypothesis to be proved:

$$\eta_{XY} = \left(\left|\, P^{Y}_{REAL} - P^{Y}_{IMM.\,INFL.} \right|\right) / \left|\, P^{Y}_{REAL} \right|. \qquad (7)$$

Using quite analogous procedure, it is possible to calculate the degree of adequacy for another, competing hypothesis (Y dominating

X): η_{YX} . As a result, three variants of their relationships are possible:

– for one of these hypotheses (e.g., ***X*** dominating ***Y***) its degree of adequacy occurs low, whereas for another hypothesis this adequacy is high; in such a case it is possible to interpret the second variable as dominating one, i.e., quite definite causality;

– for both hypotheses the degree of their adequacy is low; then it is impossible to speak about any dominance or causality;

– for both hypotheses their degrees of adequacy are high; hence, we deal with mutual influence of the variables studied.

Let's illustrate the first of these variants with the data of the above example concerning the link between the frequency of theatre-going and labor activity, this link being studied under the condition of eliminating 6 socio-demographic parameters. The degree of adequacy of the hypothesis that the theatre-going is dominating variable, occurred to be about .8, i.e., really observed difference on labor activity between 'high' and 'low' groups of respondents (meaning their frequency of theatre-going) almost coincided with the difference which would respond to the calculations based on formula (5); the difference was only 20%. Meanwhile, for the inverse hypothesis (i.e., treating labor activity as the cause of theatre-going) the degree of adequacy was about .1, i.e., the calculated difference on theatre-going between the two groups of respondents (with high labor activity and low one) occurred 8 times more than the difference really observed. Such a discrepancy evidences in favor of rather featured causal asymmetry of the link discussed: here theatre-going is the dominating variable.

So we may continue the above algorithm of calculations, focusing attention on the causality problem:

I. All the respondents are divided into two groups (desirably by a median border) over their values of the parameter ***X***: 'low' and 'high,' and their mean values X_{LOW} and X_{HIGH} are calculated.

J. Using formula (5), the value $P^{Y}_{IMM.\ INFL.}$ is calculated responding to immediate influence of the parameter ***X*** upon parameter ***Y*** when transition from 'low' group – to 'high' one, supposing ***X*** to be a cause of the change.

K. According to formula (6), the value of real change for the above transition is calculated.

L. Using formula (7), the degree of the adequacy of the hypothesis treating ***X*** to be the causal variable, is calculated: η_{XY} .

M. Quite analogous procedure (steps ***A – L***) is applied to the hypothesis treating variable ***Y*** as a causal one, and the degree of the adequacy of this hypothesis is calculated: η_{YX} .

N. Comparing values η_{XY} and η_{YX} , it is possible to conclude about the direction of causality in this link (taking into account the above considerations concerning three variants of interrelations of these values).

In the above investigation of links between cultural consumption and labor activity, exactly featured causal influence was observed in relation to theatre-going (1.2%-growth of labor activity per each annual regular contact), visiting artistic museums and exhibitions (1%-growth), and receiving books in libraries (.02%). All these kinds of consumption caused changes in labor activity of respondents. However, for such kinds of cultural consumption as variety show going and participation in meetings of circles of amateur art, it occurred impossible to establish the direction of their links with labor activity, though the above mentioned strength of these links was rather high – about 2.5% per each annual contact and .15%, respectively. Nevertheless, the values characterizing these links, are of interest both for researchers and specialists of the sphere of cultural politics [5-7].

As for another kind of links – connecting cultural consumption with the socio-moral development of personality, 'direct' influence was established for the above influence of theatre-going (13%-growth for each regular annual contact). As well such direct link was inherent in home listening to recorded music (.2%-growth for each annual contact). On the contrary, variety show going reveals no oriented link with socio-moral development; possibly this link is bilateral, as it was in the case of influence of this kind of cultural consumption upon labor activity.

The method derived for measurements of weak influences in the socio-cultural sphere (which are complicated by numerous collateral links) can be useful not only for practical application in the given sphere, but also for measurements in some other fields characterized by similar conditions.

REFERENCES

1. Petrov, V.M. "In search for new paradigm for humanitarian measurements: Information path between Scylla of subjectivism and Harybdis of operationalism," Key Engineering Materials, Vol. 437, pp. 509-513, 2010.
2. Mazhul, L.A., & Petrov, V.M., "Moral qualities of personality and the need for art: How to measure these delicate links?" V.P.Ryzhov (Ed.), Empirical Aesthetics: Informational Approach. Taganrog: Taganrog State University of Radio Engineering, pp. 133-140, 1997.
3. Golitsyn, G.A., & Petrov, V.M. "The principle of the information maximum, Zipf's law, and measurement of individual cultural development," L.Dorfman, C.Martindale, D.Leontiev, G.Cupchik, V.Petrov, & P.Machotka (Eds.), Emotion, Creativity, and Art. In two volumes. Perm: Perm State Institute of Arts and Culture, Vol. 1, pp. 179-221, 1997.
4. Petrov, V.M., "Studies of Direct and Indirect Influence of Art: Methodological and Methodical Problems," Moscow: Russky Mir Publishers, 1997 (in Russian).
5. Focht-Babushkin, Yu.U. (Ed.), "Man in the World of Artistic Culture (Mastering Art and Problems of Controlling)," Moscow: Nauka, 1982 (in Russian).
6. Petrov, V.M., "Measurement of the socio-economic effectiveness of cultural consumption," Ekonomika i Matematicheskiye Metody (Economics and Mathematical Methods), Vol. XXIV, No. 5, pp. 832-840 (in Russian).
7. Focht-Babushkin, Yu.U. (Ed.), "Artistic Culture and the Development of Personality (Problems of Long-Range Planning)," Moscow: Nauka, 1987 (in Russian).