RUDN Journal of Sociology. ISSN 2313-2272 (print), ISSN 2408-8897 (online) **2024 Vol. 24 No. 1 241–258**
Вестник РУДН. Серия: СОЦИОЛОГИЯ http://journals.rudn.ru/sociology

DOI: 10.22363/2313-2272-2024-24-1-241-258
EDN: ZTTKWR

Do indirect measures of attitudes improve our predictions of behavior? Evaluating and explaining the predictive validity of GATA*

O.L. Chernozub

Institute of Sociology of FCTAS RAS,
Krzhizhanovskogo St., 24/35–5, Moscow, 117218

(e-mail: 9166908616@mail.ru)

Abstract. The generalization of the results accumulated to date has shown that the implicit measures of attitudes (some even suggest defining them with a less pretentious term "indirect") show a disappointingly weak predictive potential in relation to real behavior. Thus, the predictive validity of the Graphical Association Test of Attitude (GATA), which also claims to be an indirect method, has been questioned. To check this assumption, we analyzed the results obtained with GATA in 64 predictions provided that the predicted outcome could be verified by real action. Such forecasts cover the domains of electoral, consumer and communicative behavior. In some cases, the prediction based on the data from a representative sample was checked referring to the actual behavior of the group represented by the sample, e.g., the electorate, or the consumers of a certain category of goods, etc. In other cases, the accuracy of the forecast was checked for each respondent. This allows to avoid the effect of "mutual compensation" of erroneous forecasts with opposite valence. The test method consisted of a comparison of the prediction accuracy of pairs of "control" and "experimental" prediction models: the only difference identified was that the latter used the data from indirect measurements of GATA as an additional factor of action. In the article, all models are presented in their simplest and most transparent versions. The results of the conducted meta-analysis do not fully correspond to the general trend: the use of the GATA data significantly and continuously improves the accuracy of predicting behavior. In addition, the incremental effect on the accuracy of individual forecasts (for each respondent) turned out to be higher than that of the sample-based group forecasts.

Key words: indirect measurement; criterion validity; predictive validity; factors of behavior; dual system theories; structural theory of attitude; implicit attitudes; GATA

In theory, indirect measures of social attitudes are an important element for explaining and predicting social phenomena. If the available methods really measure attitudes, they should explain human behavior. If, being "indirect", they really mitigate the problems of respondents' deliberate misreporting and lack of introspection, they

The article was submitted on 15.12.2023. The article was accepted on 15.02.2024.

should explain behavior better than “direct” measures of attitudes. In practice, this is not always the case. Recent debates about the validity and the predictive power of indirect measures of social attitudes question the theoretical validity and practical usefulness of such measures [26; 27; 29]. It is argued that while the predictive validity of indirect measures is low at the individual level, it is quite high at the level of group behavior [20; 31]. The low predictive validity may be explained by the low temporal stability of indirect measurements, which is easily eliminated by averaging the results of several consecutive measurements [21]. This phenomenon can also be explained by an imprecise correspondence between the object of the measured attitude and the object of the actual action [16; 23]. In addition, it is argued that implementation of the forecast depends on additional situational factors [2; 3; 33]; therefore, the prediction of behavior based on the results of indirect measures of attitude should be considered rather probabilistic than causal. Finally, there are suggestions that “indirect” measures indicate such components of attitudes that do not directly influence behavior but are a substantive part of a more complex mechanism for identifying the relative preference of each action/inaction [30].

This paper presents an evaluation of the predictive validity of the Graphical Associative Test of Attitude (GATA). By “predictive” (a form of “prospective criterion”) validity we mean the ability of the examined indicator to act as a theoretically assumed predictor of independently measured parameter (criterion). In this study, the tested parameter is the output of GATA, and the independent parameter is the fact of social action/inaction. In GATA, attitudes are understood as the tendency to consciously or unconsciously [40] perceive the object of the attitude as attractive or repulsive [12; 34]. The sociological significance of attitudes is determined by their influence on person’s social actions, encouraging him to act in accordance with the valence of the attitude towards the object [1]. If GATA measures components of attitudes, as is theoretically assumed, the results of such measurements should explain and predict social behavior.

Instrumentally, GATA attempts to avoid conscious activities of respondents, replacing them by an associative test [4]. As respondents do not have to evaluate and report their attitudes towards the tested objects, GATA should be classified as an indirect measurement instrument [2; 9; 10; 16]. GATA was introduced in 2015 as a supporting tool for poll-based election forecasting but still suffers from the lack of validation, and we attempt to fill this gap by testing the effects of GATA on social action.

The first assumption: if the interpretation of the GATA output as an indicator of attitudes is correct, then corresponding measurement results should affect behavior. This casual effect is explained by the structural theory of attitudes [36] and any dual-process theory of action [5; 6; 13–15; 17–19; 24; 28; 32; 33; 35; 37; 38]. The second assumption: if GATA can detect something more than the results of “direct” measurements reveal, then a prediction of social action based on GATA will be more accurate. This paper presents a critical observation of the accuracy of predicting social behavior for both types of prediction algorithms: GATA-free "direct-only determinants models" vs GATA-contributed "direct and indirect determinants models". In the available for our meta-analysis data of predictions for electoral, consumer and communicative behaviors, the first type is represented by conventional models based on the explicitly expressed attitudes or intentions ("control" models). "Experimental" models use the results of the GATA measurements as an additional factor presumably affecting behavior. The comparison of the accuracy of the control and experimental predictions should explain whether or not incorporating the GATA output into a predictive model leads to a better explanation and, thus, to a better prediction of social behavior.

Thus, this paper aims at clarifying some methodological issues related to the interpretation of GATA as an indicator of specific fractions of attitude, both influencing behavior and not reducible to the fractions detected by "direct" measurement. Methodologically, we would interpret the results of our analysis in the context of theories supporting the validity of "indirect" measures.

It is generally accepted that respondents may not be able or willing to fully express the true drivers of their behavior, some of which remain unrecognized by both the researcher and the respondent; the knowledge of such "hidden" or "implicit" factors of behavior should improve our ability to explain and predict social behaviors. Theoretically, it is possible to suppress these confounding effects by avoiding the respondents' self-assessment of their attitudes and self-reports of the results of these assessments. GATA was introduced to solve this task with two sequential associative procedures. First, the respondent is shown a primary stimulus representing an object of interest, followed by a set of target stimuli represented by a set of abstract graphical shapes (Fig. 1) .

Figure 1. An example of the GATA set of graphical shapes

The respondent is asked to select the graphic shape (s) that is "most appropriate" for the object under study. This task can take the form of picking one or more shapes or ranking them. The result of the first step is the graphical shape (s) that the respondent associates with the object under study. Then we take the "distracting pause" of exposure to stimuli that are not correlated with the GATA procedure: typically, these are common self-report questions from the non-GATA sections of the questionnaire. Second, the phrase with verbal markers of the approach — avoidance tendency — is presented as the primary stimulus. As a rule, the phrase includes such words as "would like to look at", "would be nice to have around", "would like to touch" and so on. The presentation of the stimulus phrase is followed by the same set of graphic shapes. At both stages, the respondent is to select from the target stimuli the graphical shapes that are the most relevant to the primary stimulus.

Technically, the procedure is structured as follows:

a. The respondent considers the studied object presented as a verbal concept on the screen of the CAPI device.
b. The set of graphic shapes is presented to the respondent on the screen of the CAPI device to choose graphic shapes for the studied object.
c. The respondent is asked other questions, preferably not related to the studied object.
d. The respondent reacts to the approach — avoidance phrase, ranking graphic shapes from the most to the least preferable for longer contact.
e. An "individual scale" of preferences for graphic shapes is based on this ranking.
f. The implicit preference score according to the "individual scale" is presented for the studied object based on the association from phase "a".

Thus, each tested object receives a score on an ordinal scale, regardless of which particular shape each respondent may prefer or dislike due to psychological, cultural, mental, physical or other factors. The predictive validity of measurements is the practical confirmation of the theoretically predicted influence of a measurand on the phenomena it is presumed to determine. Technically it implies statistically significant associations of the testing parameter with independently measured parameters or "criteria" representing presumed pairs of explanans and explanandums: the former are the results of GATA, which, if they indicate the status of attitudes, should influence social actions that are corresponding explanandums. If GATA's results improve our predictions of social action, we can argue that the results of empirical testing do not contradict the theoretically presumed properties of the method. The control criterion is the outcome of action/inaction, as identified by direct observation (not self-report). Given the available empirical data, we consider two forms: group actions (voting or consumption) and individual actions (keeping or refusing a discount coupon, filling in or skipping a feedback form, etc.).

The test algorithm used is a combination of a generally accepted "direct measurement only" prediction models with the models enriched with the indirect measurement data supplied by GATA: the former are "control", the latter are "experimental" models; both are assumed capable of predicting action/inaction.

Control models include three categories based on verbal questions as stimuli to directly test attitudes towards the activity in the prediction:

- EA — attitude towards object of anticipated action. "Is this object preferred or rejected?" (for instance, "Which candidate do you prefer?");
- AI — act intentions. "Which way do you intend to act?" ("For which candidate will you vote?");
- LAAI — likelihood to act, intentions. "Do you intend to act somehow? What do you intend to do?" ("Do you intend to vote? For which candidate?").

Thus, control models consider as potential actors all respondents who explicitly express a positive attitude towards the object of action (EA) or action (AI) or type of action aimed at a particular object (LAAI). In some cases, these models are additionally supported by the control question "Are you sure, or your attitude/intention can alter?". These models are marked with "/c" — "confirmed": EA/c, AI/c, LAAI/c; and only respondents who additionally confirmed their attitude/intentions are considered potential actors.

In this way, for each sample, we made a set of control predictions that depend on the models we can construct with the available directly measured variables. Then we applied to each control model an additional filter of the indirectly measured component (GATA). This filter excluded as potential actors all respondents whose indirectly measured attitude was negative (four "lower" or "negative" points of the GATA scale). The result was an alternative ("experimental") prediction. Next, we counted the modules of "fact minus prediction" errors for each model. Then we expressed these errors as a share of the actual outcome of actions. For example, if the election forecast is 22 % and the actual result is 20 %, the error is 2 % and 2 % / 20 % = 10 % is the "normalized error".

Thus, the validity criterion is defined as the ratio/difference of deviation between the actual and predicted outcomes for the control model and between the actual and predicted outcomes for the experimental model, normalized to the actual outcome of the event:

$$Vc = \frac{\sqrt{\left(F - Pc\right)2} - \sqrt{\left(F - Pe\right)2}}{F} \qquad (1)$$

Vc — validity criterion (degree of improvement in forecast accuracy);

Pc — predictive value of the control model;

Pe — predictive value of the experimental model;

F — actual value.

The study's main hypothesis is $H_0 1$: "There is no statistically significant differences between control and experimental models' predictions of real actions". The supporting hypothesis is $H_0 2$: "There is no statistically significant differences for GATA data's incremental effect between subsamples of group and individual actions". The empirical basis of the study consists of 64 pairs of control and experimental prediction models from 14 empirical projects that used GATA as an indirect measure of attitude. Today, large empirical material allows to make conclusions about the comparative accuracy of the predictions based on GATA measurements. For the analysis we used only empirical data sets that contain both (a) direct only and (b) GATA based indirect measurements of behavior predictors together with the data on (c) actual behaviors. The mentioned 14 surveys are as follows:

1. forecast of the results of the election of deputies to the State Duma in 2016 (A, B);
2. forecast of the results of the May 2017 elections of the heads of executive power in the regions of the Russian Federation (C1–C4);
3. forecast of the results of the Presidential election from March 2017 and December 2017 (D, E);
4. forecast of the dynamics of the Russian residential real estate market in 2021 (F1–F4);
5. forecast of the next wave of voters' answers in the panel survey in 2016 (B);
6. forecast of the behavioral choice (submitting a request to get feedback) from the 2020 methodological experiment (G);
7. forecast of the behavioral choice (submitting a request to get feedback) from the 2022 methodological experiment (H);
8. forecast of the behavioral choice (requesting or rejecting discount coupon) from the 2016 brand associations and consumer behavior research (I).

This general sample splits into two methodologically contrasting subsamples:

- Group actions prediction (1–3) includes 38 pairs of models, prediction is made for the sample, but outcome is registered for the society, which creates the risk of additional errors due to sample biases. In this subsample, reciprocal forecast errors can cancel each other. For instance, when predicting group behavior, if action is predicted for 50 % of respondents and inaction for other 50 %, it may turn out that both parts acted contrary to the prediction. In such a case, a predictive model for individual behavior will detect a prediction accuracy of zero. On the contrary, a predictive model for group behavior will not even see its own fiasco and will announce a prediction accuracy of 100 %.
- Individual actions prediction (4–8) includes 26 pairs of models that predict actions not for the group but for every respondent. For this subsample, we can assume the absence of both the risks of sample bias and the effect of reciprocal error compensation.

The main characteristics of the used empirical data are presented in Table 1.

Table 1

General characteristics of the dataset: pairs of control/experimental predictions

Prediction model	Group behavior	Individual behavior	Sum
EA	0	2 (G, H)	2
AI	18 (A, B, D, E)	20 (B, G, H, F, I)	38
LAAI	10 (A, B, C)	0	10
EA/c	0	1 (H)	1
AI/c	10 (A, B, E)	3 (H, I)	13
LAAI/c	0	0	0
Total	38	26	64

According to the Table 1, we have a large sample and subsample for predicting group behavior, but the subsample for individual predictions looks less reliable, i.e., we should carefully compare our subsamples, while the general sample is sufficient to identify the main tendencies. The general characteristics of the dataset obtained are presented in Table 2.

Table 2

General description of the dataset

№	Object	Year	Domain*	Prediction**	Model	Gap***	Accuracy improvement
1	2	3	4	5	6	7	8
1	UR (A)	2016	E	G	AI	64	–7.9 %
2	UR (A)	2016	E	G	AI/c	64	–10.3 %
3	UR (A)	2016	E	G	LAAI	64	–9.8 %
4	CPRF (A)	2016	E	G	AI	64	–7.5 %
5	CPRF (A)	2016	E	G	AI/c	64	25.6 %
6	CPRF (A)	2016	E	G	LAAI	64	9.8 %
7	LDPR (A)	2016	E	G	AI	64	14.5 %
8	LDPR (A)	2016	E	G	AI/c	64	0
9	LDPR (A)	2016	E	G	LAAI	64	16.8 %
10	FR (A)	2016	E	G	AI	64	19.4 %
11	FR (A)	2016	E	G	AI/c	64	12.9 %
12	FR (A)	2016	E	G	LAAI	64	19.4 %
13	Incumbent–1 (C1)	2017	E	G	AI	87	2.2 %
14	Incumbent–2 (C2)	2017	E	G	AI	59	7.9 %
15	Incumbent–3 (C3)	2017	E	G	AI	60	7.9 %
16	Incumbent–4 (C4)	2017	E	G	AI	52	7.3 %
17	Pretender–1 (C1)	2017	E	G	AI	87	10.8 %
18	Pretender–2 (C2)	2017	E	G	AI	59	35.1 %
19	Pretender–3 (C3)	2017	E	G	AI	60	10.1 %
20	Pretender–4 (C4)	2017	E	G	AI	52	16.2 %
21	Putin-March (D)	2017	E	G	AI	347	–2.2 %
22	Putin-March (D)	2017	E	G	AI/c	347	12 %
23	Putin-March (D)	2017	E	G	LAAI	347	–3.8 %
24	Zyuganov-March (D)	2017	E	G	AI	347	–2.7 %
25	Zyuganov-March (D)	2017	E	G	AI/c	347	14.4 %
26	Zyuganov-March (D)	2017	E	G	LAAI	347	1.8 %
27	Zhirinovsky-March (D)	2017	E	G	AI	347	26.3 %
28	Zhirinovsky-March (D)	2017	E	G	AI/c	347	17.5 %
29	Zhirinovsky-March (D)	2017	E	G	LAAI	347	–1.8 %

End of the Table 1

1	2	3	4	5	6	7	8
30	Putin-December (E)	2018	E	G	AI	104	–3.4 %
31	Putin-December (E)	2018	E	G	AI/c	104	–4.3 %
32	Putin-December (E)	2018	E	G	LAAI	104	2.3 %
33	Zyuganov-December (E)	2018	E	G	AI	104	–2.7 %
34	Zyuganov-December (E)	2018	E	G	AI/c	104	2.7 %
35	Zyuganov-December (E)	2018	E	G	LAAI	104	1.8 %
36	Zhirinovsky-December (E)	2018	E	G	AI	104	21.1 %
37	Zhirinovsky-December (E)	2018	E	G	AI/c	104	0
38	Zhirinovsky-December (E)	2018	E	G	LAAI	104	1.8 %
39	Brand–1 (I)	2016	C	I	AI	0	5.1 %
40	Brand–1 (I)	2016	C	I	AI/c	0	1.1 %
41	Brand–2 (I)	2016	C	I	AI	0	34.4 %
42	Brand–2 (I)	2016	C	I	AI/c	0	9.4 %
43	Housing-Q 1 (F)	2021	C	I	AI –1	365	22.8 %
44	Housing-Q1 (F)	2021	C	I	AI –2	365	32.6 %
45	Housing-Q1 (F)	2021	C	I	AI –3	365	29.2 %
46	Housing-Q2 (F)	2021	C	I	AI –1	365	12.5 %
47	Housing-Q2 (F)	2021	C	I	AI –2	365	11.5 %
48	Housing-Q2 (F)	2021	C	I	AI –3	365	–4.1 %
49	Housing-Q3 (F)	2021	C	I	AI –1	365	27.2 %
50	Housing-Q3 (F)	2021	C	I	AI –2	365	24.7 %
51	Housing-Q3 (F)	2021	C	I	AI –3	365	9.3 %
52	Housing-Q4 (F)	2021	C	I	AI –1	365	27.1 %
53	Housing-Q4 (F)	2021	C	I	AI –2	365	39.7 %
54	Housing-Q4 (F)	2021	C	I	AI –3	365	52.4 %
55	UR-Panel (B)	2016	O	I	AI	32	0.5 %
56	CPRF-Panel (B)	2016	O	I	AI	32	–2.5 %
57	LDPR-Panel (B)	2016	O	I	AI	32	1.8 %
58	FR-Panel (B)	2016	O	I	AI	32	3.6 %
59	Test (G)	2020	O	I	EA	0	54.3 %
60	Test (G)	2020	O	I	AI	0	60.4 %
61	Test (G)	2020	O	I	AI/c	0	2.9 %
62	Volunteers (H)	2021	O	I	EA	0	12.7 %
63	Volunteers (H)	2021	O	I	EA/c	0	34.5 %
64	Volunteers (H)	2021	O	I	AI	0	25 %

* Domains of social actions: E — electoral, C — consumer, O — online communications
** Prediction mode: G — group, I — individual
*** Gap presented in days

Table 3 presents the main descriptive statistics for the general sample of GATA incremental effects on the accuracy of prediction, and Figure 2 — its graphical form.

Table 3

Descriptive statistics for the general sample

Statistic	Value
N	64
Mean	12.31
Median	9.5
SD	15.7
Min	–10
Max	60
Range	70
Excess	0.87
Asymmetry	1.02

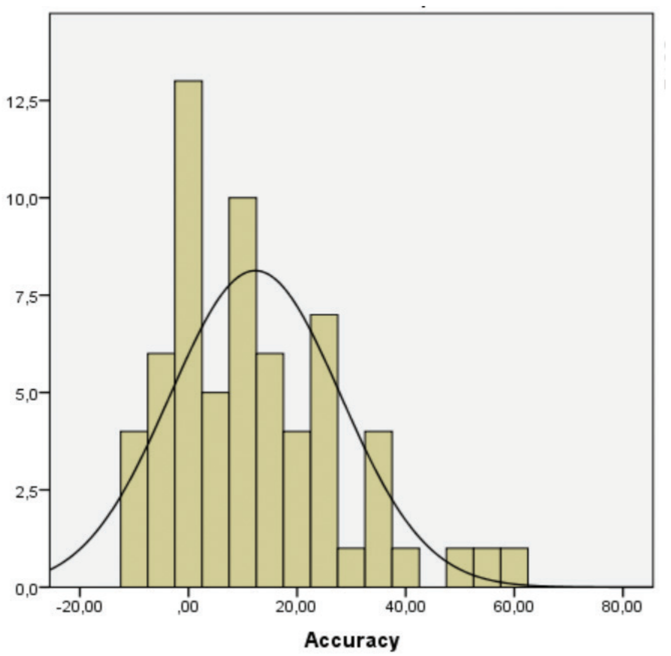

Figure 2. Distribution of the GATA prediction accuracy for incremental effects

According to Table 2, the distribution has a range of 70 % — from -10 % to 60 %. The excess and asymmetry statistics suggest a slightly “wide” distribution, with a long tail towards higher values. The mean and median are in a confidently positive position, suggesting the GATA average incremental influence of 10 %–12 %, which means that the GATA effects can be ambivalent and potentially produce negative effects; however, the magnitude of positive effects is greater.

Negative results were obtained mainly for the election forecasts for the ruling party United Russia and V. Putin (7 out of 13 negative cases — 1–3, 21, 23, 30, 31). More recent studies showed that at least the partial explanation of this error is the specific voting culture of some Russia’s national regions, which was not taken into account in the sample design. If we consider these cases “atypical” and recalculate the descriptive statistics without them, we get a median of 11.5 % and a mean of 14.6 % for the corrected sample of 57 cases, i.e., errors of experimental models decrease but do not disappear completely.

On the other hand, several observations with the extremely high values of forecast improvement stand out (54, 59, 60), albeit in different studies and revealed by different methods. The only thing they have in common is predicting personal behavior. If we exclude these values from the dataset, we get a median of 9.3 % and a mean of 10.2 % for 61 cases, which still keeps these statistics within a confidently positive interval.

Therefore, for further analysis, we decided to use the initial results of forecasts as the most cautious and balanced approach. Thus, the simultaneous ability of GATA to show both moderately negative and strongly positive results was registered as a reliable phenomenon. We considered as its explanation the contradictory nature of situations in which GATA is used: in some circumstances it tends to improve the accuracy of the forecast, in others — to worsen. Our data allows to test this hypothesis in relation to two possible determinants of this phenomenon: the first potential determinant may be differences in the prediction of group vs individual behavior (the focus of ongoing discussionas about the relatively poor predictive validity of indirect measures [27]); the second one may be the absence or presence of the “intention inflation” effect (general problems of predictive validity of attitude measures [39]).

The data on the mode of prediction is presented in Table 4 and indicates a clear difference in the distributions of the GATA effects for the compared groups. The tendency of the mean for individual predictions is much better than for group predictions (20.27 vs 6.87). For individual behavior predictions, the min, max and range values significantly shift towards the positive pole of the scale; the asymmetry also shows the longer positive tail for distribution. Both excess measures are “tighter” compared to the general sample. All these peculiarities support the assumption of different processes represented by distributions. Thereby, the effect of GATA on the results of group and individual predictions is different, and the prediction accuracy is more improved for individual behavior.

Table 4

Descriptive statistics and ANOVA test for the prediction mode

Statistic	All	Group	Individual
N	64	38	26
Mean	12.31	6.87	20.27
Median	9.5	5	18
SD	15.7	10.99	18.21
Min	–10	–10	–4
Max	60	35	60
Range	70	45	64
Excess	0.87	–0.31	–0.48
Asymmetry	1.02	0.49	0.62
F	NA	13.47	
	NA	0.001	

To check the effect of the "intention inflation", we divided each subsample of group and individual predictions approximately in half. The first half corresponds to a relatively small value of the gap between measurement and action; the second half — to a relatively large gap. For the group prediction subsample, a relatively small dataset was made up of cases with a gap of 64 days or less (18 cases out of 38); for the individual prediction subsample — 32 days or less (14 cases out of 26). For general reasons, we consider small gap cases to be less vulnerable to the effects of "intention inflation". The analysis of variance did not support the assumption of a significant difference in the distribution of GATA effects in these groups of potentially less and more inflated intentions (Tables 5–8).

Table 5

Descriptive statistics by the behavioral domains

Domain	N	Mean	SD	Min	Max	Range	Excess	Assymetry
Electoral	42	6.33	10.6	–10	35	45	–0.06	0.63
Consumer	16	20.81	15.36	–4	52	56	–0.51	0.22
Communicative	6	31.5	22.46	3	60	57	–1.59	0.13
Total	64	12.31	15.7	–10	60	70	0.87	1.02
F								13.914
P								0

Table 6

Descriptive statistics by the modes of prediction

Mode	N	Mean	SD	Min	Max	Range	Excess	Assymetry
Group	38	687	10.99	-10	35	45	-0.31	0.49
Individual	26	20.27	18.21	-4	60	64	-0.48	0.62
Total	64	12.31	15.7	-10	60	70	0.87	1.02
F								13.469
P								0.001

Table 7

Descriptive statistics by the models of prediction

Model	N	Mean	SD	Min	Max	Range	Excess	Assymetry
VI/Int	38	14.84	16.45	–8	60	68	0.33	0.79
VIC/IntC	13	6.54	9.87	–10	26	36	–0.17	0.33
LVVI	10	3.8	9.08	–10	19	29	–0.28	0.51
EA	3	33.67	20.5	13	54	41	0	–0.07
Total	64	12.31	15.7	–10	60	70	0.87	1.02
F								4.339
P								0.008

Table 8

Descriptive statistics by the “intention inflation” gap

Gap_Ordinal	N	Mean	SD	Min	Max	Range	Excess	Assymetry
Minutes	6	31.5	22.46	3	60	57	–1.59	0.13
Days	8	6.75	11.49	–2	34	36	6.13	2.4
Months	29	6.9	11.19	–10	35	45	–0.07	0.48
Year	21	16.43	15.68	–4	52	56	–0.36	0.4
Total	64	12.31	15.7	–10	60	70	0.87	1.02
F								6.171
P								0.001

Thus, the use of GATA measures in predicting social behavior is steadily improving the accuracy of such predictions. Although in some cases the accuracy of the forecast deteriorates, the frequency of such cases and their negative impact are relatively low; on the contrary, positive effects occur in most cases and are relatively strong. This allows to reasonably reject H01: “There is no statistically significant differences between control and experimental models’ predictions of real actions" — our data shows the opposite. Similar conclusions have already been made on the basis of a more detailed analysis of some pre-election surveys [5; 6]. Thereby, we extend our conclusion to some other behavioral domains, at least consumer and communication behavior.

Our conclusion contradicts the results of the above-mentioned general meta-analysis of the effects of indirect measures for predicting behavior, which can be explained by two interrelated sets of factors. First, the presumed way in which the results of indirect measures of attitudes are related to behavior: direct links between indirectly measured attitudes and actual behavior tend to be weak. For instance, the prediction of election results on the basis of the GATA data alone in some cases led to a normalized forecast error (from 80.5 % to 566 % [5. P. 84]); therefore, we analyzed incremental effects obtained in complex models that combine the results of both direct and indirect measurements. In the theoretical perspective, our approach is based on the assumption that action is not determined by the attitude but by the result of the interaction of its various components: according to the "structural theory of attitudes", these components are unequal in nature and potentially conflicting. Each measurement (direct and indirect) presumably records the state of only a fraction of the attitude.

The second set of factors we interpret as follows: the efficiency of methods combined to describe the state of attitudes can be verified by the orthogonality test of their measurement results; some popular indirect measurement methods are not in order with regard to the orthogonality of their and direct measurements results [7; 8], which can explain why the results of such measurements add little to the results of conventional direct measurements; on the contrary, GATA demonstrate reliable orthogonality to direct measurement results [4–6]. If the measurement methods are effective in the sense of complementarity, their combination should lead to an increase in the completeness of the explanation and in the accuracy of the action prediction. Thus, the methodological features of GATA, which provide information inaccessible to direct measurements, constitute a second set of factors that improve the accuracy of behavioral predictions.

The observed effect of increasing the accuracy of predicting social behavior is stable: we did not find any significant influence on its manifestations of any potential factors considered, with the exception of the prediction mode (group/individual behavior). Thus, the supporting hypothesis H02: "There is no statistically significant differences for the GATA data's incremental effect between subsamples of group and individual predictions" should be rejected.

Again, our results are not in line with the general trend: the conventionally accepted norm is that the predictive power of indirect measures is higher in relation to group rather than individual behavior. The theory that explains this phenomenon is based on the assumption that the fractions of attitudes measured by indirect methods are relatively unstable and constantly fluctuate under the influence of random causes. Being averaged for a group, such measurements are less sensitive to fluctuations in the moods of each respondent, i.e., group measurement reflects the actual state of group attitudes, and when attitudes change, it filters out stochastic noise and reveals actual changes driven by systematic factors, which explains the relatively higher predictive potential of group indirect measures [31]. How can we explain our results? Perhaps, GATA does not measure exactly the same fractions of attitudes as most other indirect methods, which is supported by some previous studies (for instance, the comparison of IAT and GATA showed that they share a common latent variable but interact with other variables in significantly different ways [7; 8]). Thus, we can claim the scientific validity of the GATA measurements as a factor of social action and a tool for predicting such actions, although the sources of the relatively high predictive power of GATA (which is not typical for indirect measures) remain unclear.

Notes: Brief methodological descriptions of the empirical sources

(1) Pre-election poll 2016. The GATA methodology was used for the first time to test the assumption about the influence of implicit factors on the attitudes and electoral intentions. The sampling was multistage, representing all social-economic macro-regions of Russia and the structure of population by type of settlement, gender, age. Survey method was interview at home, CAPI. N = 1611. The maximum standard error is 2.24 %. The validity criteria are actual results of voting for four most popular candidates; the intention inflation gap — 4 months; the object of prediction — group behavior.

(2) Pre-election panel survey 2016. The GATA methodology was used as an additional tool for forecasting the voting results within one of four waves of the panel study. The panel was representative for voters residing in Russia; by gender, age, macro-region and type of settlement. N = 3721. Survey method — online interview. The maximum standard error is 2.24 %. The validity criteria are actual results of voting for four most popular candidates; the intention inflation gap — 1 month; the object of prediction is group behavior.

(3) (B’) Study uses the same data as above but differ by subject. The validity criteria reproduce the “explicit” choice to vote for the favorite candidate in the next wave of the panel survey; the intention inflation gap is 2 weeks; the object of prediction — individual behavior.

(4) (C1-C4) Pre-election polls in the subjects of the federation 2017. The GATA methodology was used to improve the accuracy of forecasts. Four independent surveys were conducted. The samples were multistage, representing local sub-regions and population structure by type of settlement, gender, age. The combined sample size was 4,000 (N = 1,000 in each region). Survey method was interview at home, CAPI. The maximum standard error for each region is 2.32 %. The validity criteria are actual results of voting for four incumbents and four most popular candidates (one per region); the intention inflation gap — 2 months; the object of prediction — group behavior.

(5) Study of the prospective presidential candidates’ ranking in the 2018 elections from March 2017. The GATA methodology was first used to verify the fact and the nature of the influence of implicit factors on electoral attitudes and intentions. The sampling was multistage, representing all social-economic macro-regions of Russia and the structure of population by type of settlement, gender, age. Survey method was interview at home, CAPI. N = 1607. The maximum standard error is 2.24 %. The validity criteria are actual results of voting for three most popular candidates; the intention inflation gap — 11 months; the object of prediction — group behavior.

(6) February 2018 pre-election poll. The GATA methodology was used to improve the accuracy of predicting the results of the 2018 Presidential Election. The sample is multistage, representing all social-economic macro-regions of Russia and population structure by type

of settlement; gender, age. Survey method was interview at home, CAPI. N = 1614. The maximum standard error is 2.24 %. The validity criteria are actual results of voting for three most popular candidates; intention inflation gap — 1 month; the object of prediction — group behavior.

(7) (F1-F4) Monitoring of the demand dynamics in the Russian housing market in 2021–2022. The GATA method was used to improve the accuracy of forecasting the demand. Four quarterly measurements were made in 2021 to forecast the market dynamics in 2022. The sample consisted of respondents visiting websites — integrators of housing market offers. N = 600. The maximum standard error is 3.44 %. The validity criteria are actual "next 12 months" volumes of the market as per moving average shifted by quarter; the intention inflation gap — 12 months; the object of prediction — group behavior.

(8) (G) Methodological experiment in 2020. To test its ecological validity, the GATA method was presented to respondents as a "psychological test". The sample consisted of users of the Runet. Quota sampling control: gender, age, type of settlement. N = 1204. The maximum standard error is 2.26 %. Online survey. The validity criterion is a request for the results of the "test"; the intention inflation gap — a few seconds; the object of prediction — individual behavior.

(9) (H) Methodological experiment in 2022 for the complex theoretical validation of GATA in a form of the all-Russian mass survey (of Russian citizens-users of the Russian-language segment of the Internet). The sample was controlled by gender, age, type of locality. N = 2100 respondents. The maximum standard error is 2.14 %. The validity criterion is a feedback to contact the favorite candidate's local representative; the intention inflation gap — a few seconds; the object of prediction — individual behavior.

(10) Brand association and consumer preferences research in 2016. The sample consisted of the Russian-speaking consumers of the brand category. N = 1200. The maximum standard error is 3.4 %. The validity criterion is requesting a discount coupon for the target brand; the intention inflation gap — a few seconds; the object of prediction — individual behavior.

References

1. Allport G. Attitudes. *A Handbook of Social Psychology*. Clark University Press; 1935.
2. Brownstein M., Madva A., Gawronski B. What do implicit measures measure? *Wiley Interdisciplinary Reviews: Cognitive Science.* 2019; 10 (5).
3. Brownstein M., Madva A., Gawronski B. Understanding implicit bias: Putting the criticism into perspective. *Pacific Philosophical Quarterly.* 2020, 101 (2).
4. Chernozub O. Affective components of electoral behavior: Design and validity of visual association test of attitude. *Monitoring of Public Opinion: Economic and Social Changes.* 2018; 3. (In Russ.).
5. Chernozub O. Implicit factors and inconsistency of electoral behavior: From attitude to behavior. *Monitoring of Public Opinion: Economic and Social Changes*. 2020; 5.
6. Chernozub O. Graphic associative test of attitudes as a convenient implicit measurement tool for mass polls. *RUDN Journal of Sociology.* 2023; 23 (1).
7. Chernozub O., Belonozhko M. Comparative analysis of implicit GATA and IAT measures: Unity in diversity. *Monitoring of Public Opinion: Economic and Social Changes.* 2023; 5. (In Russ.)
8. Chernozub O., Shuraeva L. Orthogonality of IAT and GATA results: The worse the better? *Monitoring of Public Opinion: Economic and Social Changes.* 2023; 6. (In Russ.).
9. Corneille O., Hütter M. Implicit? What do you mean? A comprehensive review of the delusive implicitness construct in attitude research. *Personality and Social Psychology Review.* 2020; 24 (3).
10. De Houwer J., Moors A. How to define and examine the implicitness of implicit measures. B. Wittenbrink, N. Schwartz (Eds.). *Implicit Measures of Attitudes: Procedures and Controversies*. Guilford; 2007.
11. Eagly A., Chaiken S. The psychology of attitudes. *Journal of Marketing.* 1993; 34 (2).

12. Eagly A., Chaiken S. The advantages of an inclusive definition of attitude. *Social Cognition.* 2007; 25 (5).
13. Evans J. Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology.* 2008; 59.
14. Fazio R. Attitudes as object-evaluation associations of varying strength. *Social Cognition.* 2007; 25 (5).
15. Fazio R. The role of attitudes in memory-based decision making. *Journal of Personality and Social Psychology.* 1990; 59 (4).
16. Gawronski B., de Houwer J., Sherman J. Twenty-five years of research using implicit measures. *Social Cognition.* 2020; 38.
17. Gilovich T., Griffin D. Introduction — Heuristics and biases: Then and now. T. Gilovich, D. Griffin, D. Kahneman (Eds.). *Heuristic and Biases: The Psychology of Intuitive Judgment.* Cambridge University Press; 2002.
18. Greenwald A., Poehlman T., Uhlmann E., Banaji M. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology.* 2009; 97 (1).
19. Greenwald A., Smith C., Sriram N., Bar-Anan Y., Nosek B. Implicit race attitudes predicted vote in the 2008 U.S. Presidential Election. *Analyses of Social Issues and Public Policy.* 2009; 9.
20. Greenwald A., Banaji M., Nosek B. Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology.* 2015; 108 (4).
21. Greenwald A., Brendl M., Cai H., Cvencek D., Dovidio J., Friese M., Wiers R. Best research practices for using the Implicit Association Test. *Behavior Research Methods.* 2021; 20.
22. Hassan L., Shiu S., Shaw D. Who says there is an intention–behavior gap? Assessing the empirical evidence of an intention–behavior gap in ethical consumption. *Journal of Business Ethics*. 2016; 136 (2).
23. Irving L., Smith C. Measure what you are trying to predict: Applying the correspondence principle to the Implicit Association Test. *Journal of Experimental Social Psychology.* 2020; 86.
24. Kahneman D., Frederick S. Representativeness revisited: Attribute substitution in intuitive judgment. T. Gilovich, D. Griffin, D. Kahneman (Eds.). *Heuristic and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press; 2002.
25. Kahneman D. *Thinking, Fast and Slow*. Farrar, Straus and Giroux; 2011.
26. Machery E. Anomalies in implicit attitudes research. *Wiley Interdisciplinary Reviews: Cognitive Science.* 2021. https://doi.org/10.1002/wcs.1569.
27. Machery E. Anomalies in implicit attitudes research: Not so easily dismissed. *Wiley Interdisciplinary Reviews: Cognitive Science.* 2022. https://doi.org/10.1002/wcs.1591.
28. Metcalfe J., Mischel W. A hot/cool-system analysis of delay of gratification: Dynamics of willpower. *Psychological Review.* 1999; 106.
29. Mitchell G., Tetlock P. Popularity as a poor proxy for utility: The case of implicit prejudice. S. Lilienfeld, I. Waldman (Eds.). *Psychological Science under Scrutiny: Recent Challenges and Proposed Solutions*. Wiley; 2017.
30. Moors A., Koster M. Behavior prediction requires implicit measures of stimulus-goal discrepancies and expected utilities of behavior options rather than of attitudes toward objects. *Wiley Interdisciplinary Reviews. Cognitive Science.* 2022. https://doi.org/10.1002/wcs.1611.
31. Payne B., Vuletich H., Lundberg K. The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry.* 2017; 28 (4).
32. Perugini M. Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology.* 2005; 44 (1).
33. Perugini M., Richetin J., Zogmaister C. Prediction of behavior. *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*. Guilford Press; 2010.

34. Petty R., Duane T., Wegener D., Fabrigar L. Attitudes and attitude change. *Annual Review of Psychology*. 1997; 48 (1).
35. Roccato M., Zogmaister C. Predicting the vote through implicit and explicit attitudes: A field research. *Political Psychology.* 2010; 31.
36. Rosenberg M., Hovland C., McGuire W., Abelson R., Brehm J. *Attitude Organization and Change: An Analysis of Consistency among Attitude Components*. Yale University Press; 1960.
37. Strack F., Deutsch R. Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review.* 2004; 8.
38. Strack F., Neumann R. Furrowing the brow may undermine perceived fame: The role of facial feedback in judgments of celebrity. *Personality and Social Psychology Bulletin.* 2000; 26.
39. Sussman R., Gifford R. Causality in the theory of planned behavior. *Personality and Social Psychology Bulletin.* 2019; 45 (6).
40. Thurstone L. Attitudes can be measured. *American Journal of Sociology.* 1928; 33.

DOI: 10.22363/2313-2272-2024-24-1-241-258
EDN: ZTTKWR

Улучшают ли косвенные измерения социальной установки прогноз поведения: прогностическая валидность GATA*

О.Л. Чернозуб

Институт социологии ФНИСЦ РАН,
ул. Кржижановского, 24/35, к. 5, Москва, 117218, Россия

(e-mail: 9166908616@mail.ru)

Аннотация. Обобщение большого количества накопленных к настоящему моменту данных показало, что имплицитные измерения социальной установки (предлагается даже заменить их название на менее претенциозное — «косвенные») показывают разочаровывающе слабый прогностический потенциал по отношению к реальному поведению. На этом фоне прогностическая валидность «Графического ассоциативного теста отношения» (ГАТО), который также претендует на роль косвенного метода измерения, также оказалась под вопросом. Мы проанализировали 64 прогноза поведения, которые использовали данные ГАТО в области избирательного, потребительского и коммуникативного поведения, где предсказанный результат был подтвержден или опровергнут реальными действиями. В одних случаях прогноз по данным репрезентативной выборки проверялся по отношению к фактическому поведению рассматриваемой группы (например, корпуса избирателей или потребителей определенной категории товаров). В других случаях точность прогноза проверялась для каждого респондента, что позволяет избежать эффекта «взаимной компенсации» ошибочных прогнозов с противоположными знаками. Использованный метод тестирования состоял в сравнении точности прогноза для пар «контрольных» и «экспериментальных» прогнозных моделей. Вторые отличались от первых только тем, что в качестве дополнительного фактора

Статья поступила 15.12.2023 г. Статья принята к публикации 15.02.2024 г.

использовали косвенные измерения ГАТО. Все модели были использованы в своих наиболее простых и очевидных форматах. Оказалось, что результаты нашего метаанализа не вполне соответствуют общей тенденции: данные ГАТО значительно и устойчиво повышают точность прогнозирования поведения; его влияние на точность индивидуальных прогнозов (для каждого респондента) оказалось выше, чем на точность групповых прогнозов.

Ключевые слова: косвенные измерения; критериальная валидность; предиктивная валидность; факторы поведения; теории дуальной системы; структурная теория установки; имплицитная установка; ГАТО