Psicológica (2010), 31, 477-489.

Initial conditions in the averaging cognitive model

Noventa S.[1 *], Massidda D.[2] and Vidotto G.[2]

[1]Department of Developmental ad Social Psychology, University of Padova, Italy; [2]Department of General Psychology, University of Padova, Italy

The initial state parameters s_0 and w_0 are intricate issues of the averaging cognitive models in Information Integration Theory. Usually they are defined as a measure of prior information (Anderson, 1981; 1982) but there are no general rules to deal with them. In fact, there is no agreement as to their treatment except in specific situations such as linear models where they can be merged with the arbitrary zero inter-response scale C_0. We present some considerations on their meaning and usefulness in the Functional Measurement approach, starting from different points of view. Furthermore, we suggest a method to deal with their complexity both within each single trial of a factorial design, and between the overall trials of an experiment.

Despite their algebraic simplicity initial state parameters of averaging models are very complex issues in the domain of Functional Measurement. There are no general rules to deal with their effective meaning and interpretation, particularly in comparison with others parameters. However, this kind of difficulty in measuring or sometimes even understanding the meaning of a parameter is typical of most multi-attribute modeling traditions like those of Numerical Conjoint analysis or Information Integration Theory (Lynch, 1985; Oral & Kettani, 1989).

Without any claim of being exhaustive, in this paper we attempt to analyze initial parameters from several points of view that each deserves consideration. With some simple observations we will try to suggest a possible solution to deal with their complexity.

Furthermore, after a brief survey on their meaning in Functional Measurement (Anderson, 1965), we will review the problems of their analytical and statistical uniqueness and identifiability, and sketch out a

---

[*] Correspondence to: Stefano Noventa; *e-mail address*: stefano.noventa@unipd.it.
Acknowledgements: We would like to thanks the reviewers for the fundamental issues brought up on the topic. We would also like to thank Cheng L. Qian for the precious help he gave us.

relation between their meaning and the Dynamical System Theory from which they were borrowed (Anderson, 1959).

At the end we will try to suggest a method to treat them considering all the problems raised in the previous sections; this method, however, implies a different interpretation of their meaning.

ORIGINAL DEFINITION

The first definition of an initial state parameter was proposed in the context of an averaging-type model for serial presentations (Anderson, 1959); the variable, named X_0, was a measure of the opinion of a subject before the presentation of a communication.

In the framework of Functional Measurement (Anderson 1965, 1967) that previous term was then transformed into an I_0 constant defined like an initial or neutral impression that was combined, in a weighted average, with the scale values of the stimuli of a set of attributes.

In a later work Anderson (1981) renamed that constant s_0 and weighed it with an w_0 term for similarity with the averaging models formalism. The w_0s_0 term was defined as an internal state variable representing subject's initial belief or an attitude concerning the experimental situation (Zalinski, 1984); more generally a prior memorial information (Zalinski and Anderson, 1991). Also, in order to protract the similarities with the other subjective measure parameters s in the averaging model, the s_0 parameter was thought to be the internal representation of a stimulus S_0 referring to a subject's initial attitude or prior belief. Indeed, Anderson (1981, pp. 63-64) claims that: "*S_0 need not be considered as a unitary entity. Rather, it may be some complex field of cognitive elements. The w_0 and s_0 parameters may therefore represent the resultant of some integration operation over some internal stimulus field.*"

These parameters represent the "initial state" of the cognitive integration process that leads to manifest responses (impressions, judgments and so on), a sort of *bias* that modulates the integration process. Whether or not a *bias* should be considered either an undesirable and unfortunate side-effect or a natural functioning of past attitudes and beliefs in the integration process is an important issue recently raised by Anderson (2009)[1].

[1] In this paper we will hence use the term *bias* in a broad way, as a sort of shift on the manifest response that influences the attribute evaluations of stimuli, generated by an initial state, regardless of its real nature.

Starting by these considerations, the averaging integration rule stated that:

$$R = \sum_{i=0}^{N} \frac{w_i}{\Omega} s_i \qquad \textit{with} \qquad \Omega = \sum\nolimits_{i=0}^{N} w_i \tag{1}$$

whereas $i = 1, \dots N$ is a multi-index taking into account both the overall number of levels and attributes included in each set of stimuli. The condition $i = 0$ represents the initial state parameters.

However, averaging model possesses an intrinsic difficulty: the number of parameters that have to be estimated generally differs from the number of available equations. Hence, the analytical equation systems are generally underestimated or overestimated: for instance, including the initial parameters, a 3 × 3 design has fourteen parameters to be estimated whereas it gives only nine equations, a 2 × 2 design has four equations for ten parameters, a 5 × 5 design has twenty-two parameters against twenty-five equation. Obviously, problems of identifiability and uniqueness of the solution do not exist whereas an analytical solution can be found. Nevertheless, an analytical solution is impossible in most of the cases and the problem is solved by searching for least square solutions.

ANALYTICAL AND STATISTICAL SOLUTIONS

When is a factorial design analytically solvable?

If we take a $N \times N \times \dots \times N$ factorial design, with n factors, in absence of initial conditions and without using the method of sub-designs (Norman, 1976; Anderson, 1982), simple algebra shows that all the analytically solvable designs are solutions of the equation:

$$\begin{cases} y = N^n \\ y = 2nN \end{cases} \longrightarrow N(N^{n-1} - 2n) = 0 \tag{2}$$

thus the only natural number solutions that corresponds to concrete experimental designs, for an *s*-*w* representations, are the trivial case N = 0 and the two sixteen-equations case of a 4 × 4 design ($N = 4$, $n = 2$) and a 2 × 2 × 2 × 2 design ($N = 2$, $n = 4$).

Generalizing equation (2) to factors that possess different number of levels, namely a full factorial design of the type $N_1 \times N_2 \times \ldots \times N_n$, leads to:

$$\begin{cases} y = \prod_k N_k^{m_k} \\ y = 2\sum_k N_k \end{cases} \longrightarrow \prod_k N_k^{m_k} - 2\sum_k N_k = 0 \qquad (3)$$

where k indexes the set of levels of all the factors. This equation has more solutions that correspond to analytically solvable experimental designs than the previous case (in addition to the trivial solution $N_j = 0$ for each j between 1 and n), and their number depends on the number of factors. For instance a $N_1 \times N_2$ design has three solutions, a $N_1 \times N_2 \times N_3$ has fifteen solutions, a $N_1 \times N_2 \times N_3 \times N_4$ has forty-nine solutions.

As an example, the $N_1 \times N_2$ design, in absence of initial condition and without sub-designs, shows that analytically solvable designs are those that fulfill the following equation:

$$N_1 N_2 = 2N_1 + 2N_2 \qquad (4)$$

namely the 3 × 6, the 4 × 4 and the 6 × 3 design.

What happens if we add to the problem the two initial state parameters required by Anderson?

Equation (4) becomes:

$$N_1 N_2 = 2N_1 + 2N_2 + 2 \qquad (5)$$

and, if the relation (4) had three solutions, equation (5) is satisfied by four experimental designs (that are the 3 × 8, the 4 × 5, the 5 × 4 and the 8 × 3).

On the same line of thought if we allow a variable number of parameters p we obtain:

$$N_1 N_2 = 2N_1 + 2N_2 + p \qquad (6)$$

and now each (overestimated) $N_1 \times N_2$ design can be solved by adding a number p of parameters, as reported in Table 1.

Table 1. Number p of additional parameters needed to analytically solve equation (6).

N	2	3	4	5	6	7	8	9	10
2	–	–	–	–	–	–	–	–	–
3	–	–	–	–	–	1	2	3	4
4	–	–	–	2	4	6	8	10	12
5	–	–	2	5	8	11	14	17	20
6	–	–	4	8	12	16	20	24	28
7	–	1	6	11	16	21	26	31	36
8	–	2	8	14	20	26	32	38	44
9	–	3	10	17	24	31	38	45	52
10	–	4	12	20	28	36	44	52	60

Extending the procedure to more than two factors, as per equation (3), results in several analytically solvable systems (namely the equations provided by different full factorial designs) once added an adequate number of parameters p. Nevertheless, there will be some factorial designs that will remain analytically unsolvable.

Therefore, if we apply the method of sub-designs to a $N \times N \times \ldots \times N$ full factorial design, providing the system with other equations, the analytically solvable factorial designs will be all the solutions of the following equations system, by means of an adequate number p of parameters:

$$\begin{cases} y = \sum_{k=1}^{n} \binom{n}{k} N^k \\ y = 2nN + p \end{cases} \longrightarrow p = \sum_{k=1}^{n} \binom{n}{k} N^k - 2nN \qquad (7)$$

Hence, a classical 3 × 3 factorial design could be analytically solved by adding three parameters (for instance C_0 and the initial state parameters s_0, w_0 as in Anderson, 1981), and yet a 3 × 3 × 3 designs (with sub-designs) needs forty-five additional parameters to be analytically solvable. This huge disparity is a weakness of factorial designs. Indeed, Anderson (1982) and Zalinski (1987) suggest that, in spite of a unique analytical solution, a statistical one should be searched for, that is, uniqueness and identifiability of solutions can be obtained by means of statistical methods. The method of sub-designs provides an additional number of equations that underlie statistical analysis. In other words, a least square solution is searched for the overall system of analytical equations.

In the previous section we have briefly reviewed the problem of parameters identifiability and uniqueness for averaging models, highlighting the fact that there are several systems (although not as many as would be desirable) that could be analytically solved. The number of these uniquely solvable systems could be raised by simply allowing a number of initial state parameters different from the two suggested in the Information Integration Theory framework (Anderson, 1981), but there is the need for a logical and meaningful way in which this may be achieved.

Furthermore, for such systems an analytical-statistical hybrid methodology could be useful: indeed, using different subsets of the equations provided by the method of sub-designs, and setting an adequate p number of additional parameters (which should not necessarily be interpreted as initial state parameters), different analytical solutions to the problem may be found. Statistical methods could then be layered over the analytical results of all the possible equations systems obtainable from the experimental data. Otherwise, parameters obtained by solving one (or more) of these systems could be used as starting values for an iterative algorithm like AVERAGE (Zalinski, 1987; Zalinski and Anderson, 1991) or R-Average (Vidotto and Vicentini, 2007).

To build such a p-variable system we need to investigate more in detail the initial conditions.

DYNAMIC AND STATIC APPROACH

Whereas the parameter s can be easily identified with the subjective measure of some attribute and w can be related with its importance, a similar interpretation for s_0 and w_0 is far from simple: for instance, should they be interpreted as a dynamic or a static *bias*?

This point deserves careful considerations: initial state parameters are generally important features of Dynamical System Theory and, if we conceive the acts of perception and integration as dynamical phenomena, then we should deal with initial state parameters as if they were starting points in the evolution of the system. Indeed, they have been inspired by a parallelism with an iterative discrete serial presentation (Anderson, 1959) in which X_0 was the first term of the series.

However, although in the averaging model a dynamical structure is introduced by showing to the subjects a sequence of designs and sub-designs, there is not a truly iterative formulation and initial state parameters are more of a common feature of all the sub-designs, rather than an effective starting point. In other words, they can be considered a sort of starting point only for each single presentation and when the assumption of independence of the responses holds.

In such a perspective, what effectively constitutes the prior information? Moreover, how does the system behave and evolve over time if w_0 and s_0 are the same in every experimental condition? Does the analogy to the initial state constant of Dynamical Systems still hold?

Considering the system from a dynamical perspective is a fascinating approach, but the structure of an averaging model and of the presentation sequence of a factorial design is not the structure of a real iterative system; it is instead a static one. Namely, there's a static evaluation of attributes that lead to a decomposition of the judgment into subjective parameters inside each cell of the design.

Hence, we question whether the initial state parameters could be considered local constants that regard each single cell of the design, instead of global constants relative to the entire experiment. Certainly, we are aware that there exist features which equally affect all the sub-designs and the full factorial design: for instance, in a classical integration task in which responders evaluate the sexual attractiveness of fictional persons varying in beauty and personality, sexual orientation affects both the full-factorial and all the sub-designs in a rather similar way, as it seems hardly conceivable that one's sexual orientation is dependent on the stimuli of a certain sub-design.

Nevertheless, probably these two perspectives are not mutually exclusive: testing the equality of different parameters could be a further way to assess the effective existence of a unique *bias*. In the same way, since the hypothesis of allowing initial state parameters to record *bias* toward a specific factor or collected information of previous trials violates the experimental assumption that responses are independent, we could directly use the experimental results to test if the independence assumption itself holds.

On the other hand, if we allow each sub-design to have its own initial state parameters, we incur a complex model with too many parameters. We need a sort of “principle of parsimony” to set up the appropriate number of initial conditions. At the same time, we wonder whether there really exists an ideal number of initial state parameters for every experiment?

ONE WAY SUB-DESIGNS

Do we really need initial conditions for one way sub-designs? Theory states that the *bias* should be averaged even in one-way design (Anderson, 1981). Nevertheless, should a *bias* be averaged within a one-way design? We could consider a *bias*, over a single dimension, as a linear or additive effect. So, we could model one-way designs as:

$$R = s + b + \varepsilon \tag{8}$$

where s is the subjective evaluation of the attribute, b is the subject's *bias* and ε is the normal error term.

If we are analyzing group data the *bias* should be normally distributed, like the normal error, hence the mean of R should be the same as the mean of s, similarly as it is for an additive model (Anderson, 1981). On the other hand, if we are analyzing a single subject's data, we probably do not need to identify the *bias.* This is because, if it does in fact exist, the *bias* is intrinsic in the subject's behavior, and it will affect both the successive sub-designs and the full factorial design. Hence the mean of R could be interpreted again as equal to the mean of s.

Thus, in both cases we could neglect the initial state parameters as a *bias* measure; since, in group data, they should theoretically vanish, whereas in single subject data they are necessary to the analysis: if we want the final parameters s and w to be a measure of the subjective values of the attributes, we need to consider the *bias* as a part of the parameters themselves. Otherwise, interpretation of n-th way design would be influenced by the absence of *bias*.

Finally, if one way sub-designs could be considered linear models they can be used to reduce the total number of necessary parameters: s-parameters can be identified by the one way sub-designs.

Eventually, in the case of a non-linear assumed model for the one way sub-designs, s-parameters could be still identified by neglecting initial parameters for one way sub-design, and the result could be used as a baseline for iterative algorithmic analysis such as those carried out by AVERAGE or R-Average.

N-TH WAY DESIGNS

An opposite conclusion can be reached for n-th way designs: here the initial state parameters seems to be necessary. In single subject data analysis, since the s-parameters are supposed to be *biased* and measured by the one-way designs, the s_0 and w_0 parameters in the *n*-th way designs (see equation 9 for an example of a two-way design) become a measure of the *bias* of that single design of the experiment. Then, if they are the same in all N-th way designs, we have the case of a prior information term that equally affects all designs. Otherwise, our experimental design likely violates the fundamental assumption of independence.

$$R_{ij} = \frac{w_0 s_0 + w_{Ai} s_{Ai} + w_{Bj} s_{Bj}}{w_0 + w_{Ai} + w_{Bj}} \tag{9}$$

The same happens for group data: since *s*-parameters now are supposed to be *unbiased* and measured by the one-way designs, the s_0 and w_0 parameters could became a measure of an objective *bias* of that single design of the experiment.

A POSSIBLE APPROACH

What we mentioned before implies an additional initial state parameters for every cell of every design (except for the one-way sub-designs). This is not a parsimonious model. However, we can further reduce the number of these additional parameters by means of at least two different criteria:

1- Experimental hypotheses
2- Statistical analyses

EXPERIMENTAL HYPOTHESES

Let us try to shape the relation between experimental hypotheses and additional parameters assuming for instance a 3 x 3 x 3 full factorial design. If we are interested in verifying the presence of a possible *bias* due to a certain combination of factors, independently from the levels of these factors, we could for instance build a model of this kind:

1- One-way designs without bias
2- Each two-way sub-design with its own bias
3- The full factorial design with its own bias

then we will have to deal with 18 *s* and *w* parameters having a total of 27 + 27 + 9 = 63 equations; the number of chosen *biases* is 2 + 2 × 3 = 8. If we add a C_0 parameters we have exactly 27 parameters that have to be estimated. We can then choose between 1) an analytical solution by means of the equations provided by only the full factorial design, 2) a statistical solution provided by the method of sub-designs, or 3) a hybrid methodology in which the results of analytical equations system can be used as a baseline for iterative algorithmic procedures.

Another example could be the following one: let us consider a 3 x 3 experimental design. If we are interested in verifying the relation between possible *biases* induced by the levels of factor A, in the presence of factor B but independently from its levels, we could build this model:

1- One way designs without bias
2- In the full factorial design bias should be introduced like in table 2.

Table 2. Structure of the bias parameters.

	B_1	B_2	B_3
A_1	Same *bias* $w_{0,1}$ and $s_{0,1}$		
A_2	Same *bias* $w_{0,2}$ and $s_{0,2}$		
A_3	Same *bias* $w_{0,3}$ and $s_{0,3}$		

On the contrary, if we were interested in verifying the relation between the *biases* induced by the levels of the factor B, in presence of factor A, we could build the opposite model. Comparison of these models, by means of fit indexes like BIC (Schwarz, 1978; Raftery, 1995) or AIC (Akaike, 1974;1976), could be useful to understand eventual difference in *bias* induced by the two factors.

Hence, the number of parameters in the previous examples was completely defined by the experimental hypotheses. Analytical, statistical or hybrid analysis could be used, depending on the nature of the model, to estimate the parameters. Nevertheless, statistical analyses could also be used to reinforce the decision of removing some s_0 and w_0 parameters from the model.

STATISTICAL ANALYSIS

If we assume the hypothesis where the mean of one-way R is equal to the mean of s, as previously stated, we could use regression analysis to estimate the relation between w-parameters and the existence of an s_0 and w_0

in a *n*-th way sub-design. Taking for instance a two-way sub-design we can hypothesize the following relation (that holds in each cell of the design):

$$R_{ij} = \beta_0 + \beta_1 R_{Ai} + \beta_2 R_{Bj} \quad (10)$$

whereas the hypothesized linear dependence between *R* and *s* leads to:

$$\begin{cases} \beta_0 = \dfrac{w_0 s_0}{w_0 + w_{Ai} + w_{Bj}} \\ \beta_1 = \dfrac{w_{Ai}}{w_0 + w_{Ai} + w_{Bj}} \\ \beta_2 = \dfrac{w_{Bj}}{w_0 + w_{Ai} + w_{Bj}} \end{cases} \quad (11)$$

Linear regression could then be used to infer the importance of the w and $w_0 s_0$ parameters of the model. If a beta regression coefficient is statistically null (or lower than a fixed threshold value) the correspondent parameter would be negligible in the final analysis. Furthermore, ratios between the regression coefficients are a measure of the ratios between w-parameters. Hence, once a scale unity is decided upon, all the w and $w_0 s_0$ parameters could be calculated (using for instance a multivariate linear regression or, better, a path analysis). If all the resulting betas are statistically coherent in their proportion from cell to cell there will be no need for analytical or iterative solutions. Otherwise, they could be used as baseline for iterative algorithms as we mentioned before.

A final notice regards the use of equation (10) with a non linear dependence hypothesized for *R* and *s:* in this case the procedure still holds, but beta regression coefficients have a different meaning depending on the functional form that relate *R* to *s*. Nevertheless, they could still be useful for determining whether some $w_0 s_0$ parameters can be ignored, reducing in such a way the number of parameters that have to be estimated.

GENERAL DISCUSSION

All the previous considerations suggest a possible way to handle the initial state parameters s_0 and w_0 in a sort of "dynamical" framework; although not in the proper sense of the Dynamical System Theory, but from a flexible perspective in which there is no predetermined number of such parameters.

The global number of s_0 and w_0 parameters could depend upon the experimental hypotheses we would like to test and the importance of each parameter in the global model. This latter result could be obtained through statistical analyses on the collected data: linear regression (or path analysis), in fact, under suitable hypotheses, could be used to evaluate the importance of w_0 parameters and to calculate the ratios between *w*-parameters (and their values) both for single subject and group data analysis. However, in such a model, the meaning of initial state parameters is slightly different from the original interpretation described by Anderson. In the perspective of the present work, these parameters may be a measure of the *bias*, if it exists, in each cell of a design and not a sort of starting point for the system evolution.

Moreover, in such a framework, a model may often be analytically solvable by means of an adequate number of parameters. If they could be subdivided among the sub-design in a well-advised way that fits the experimental hypotheses, analytical solutions can be found. Otherwise previous results can always be used as a baseline for iterative methods.

Finally, it is worthy of notice that this interpretation of the initial state parameters needs much more theoretical considerations and empirical data: indeed, this interpretation could not hold for all integration tasks. This interpretation, also, does not substitute the "prior information" approach. It is instead meant to be largely complementary as a means to analyze experimental data from different perspectives.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* 19 (6), 716–723.

Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion. In R. K. Mehra & D. G. Lainotis (Eds.), *System identification: Advances and case studies* (pp. 52–107). New York: Academic Press.

Anderson, N. H. (1959). Test of a model for opinion change. *Journal of Abnormal and Social Psychology*, 59, 371-381.

Anderson, N. H. (1965). Averaging versus adding as a stimulus combination rule in impression formation. *Journal of Experimental Psychology*, 70, 394-400.

Anderson, N.H. (1967). Averaging model analyses of set-size effect in impression formation. *Journal of Experimental Psychology*, 75 (2), 158-165.

Anderson, N. H. (1981). *Foundations of Information Integration Theory*. New York: Academic Press.

Anderson, N. H. (1982). *Methods of Information Integration Theory*. New York: Academic Press.

Anderson (2009). *Unified social cognition*. Psychology Press.

Lynch, J. G., Jr. (1985). Uniqueness issues in the decompositional modeling of multiattribute overall evaluations: an information integration perspective. Journal of Marketing Research, vol. 22 (1), 1-19.

Norman, K. L. (1976). A solution for weights and scale values in functional measurement. Psychological Review, 83 (1), 80-84.

Oral, M. & Kettani, O. (1989). Modelling the process of multiattribute choice. The Journal of the Operational Research Society, vol. 40 (3), 281-291.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111-163.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals if Statistics*, 6 (2), 461-464.

Vidotto, G. & Vicentini, M. (2007). A general method for parameter estimation of averaging models. *Teorie e modelli*, 12 (1-2), 211-221.

Zalinski, J. (1984). Parameter estimation: An estimated variance-covariance matrix for averaging model parameters in information integration theory. *Behavior Research Methods, Instruments, & Computer*, 16 (6), 557-558.

Zalinski, J. (1987). *Parameter estimation for the averaging model of information integration theory*. Unpublished doctoral dissertation, University of California, San Diego, La Jolla, CA.

Zalinski, J., & Anderson, N. H. (1986). *Average: a user-friendly fortran-77 program for parameter estimation for the averaging model of information integration theory* [computer software]. San Diego, CA.

Zalinski, J., & Anderson, N. H. (1991). Parameter estimation for averaging theory. In N. H. Anderson (Ed.), *Contributions to Information Integration Theory* (Vol. 1: Cognition, pp. 353-394). Hillsdale, NJ: Lawrence Erlbaum Associates.

(Manuscript received: 22 July 2009; accepted: 5 November 2009)