Deepening the Demographic Mechanisms in a Data-Driven Social Simulation of Moral Values Evolution

Samer Hassan[1], Luis Antunes[2], and Millán Arroyo[1]

[1] Universidad Complutense de Madrid, Spain
{samer@fdi,millan@cps}.ucm.es
[2] GUESS/Universidade de Lisboa, Portugal
xarax@di.fc.ul.pt

Abstract. The "Keep It Simple, Stupid" principle is a recommended rule for modelling complex phenomena. However, there must be a compromise between simplification and expressiveness, determined by the results produced by the model. Here we propose to gradually increase the complexity of a model, so we can improve its behaviour. This incremental "deepening" process is an attempt to approach the real phenomena, so resulting in a better model, provided that an accurate analysis reveals the right steps. As application we propose an agent-based data-driven model of the evolution of moral values in the Spanish post-modern society. We focus on improving the demographic mechanisms so that the system output follows the evolution of Spanish population. In order to do that, we raise the amount of quantitative input information of the system, improve its statistical distributions, and change the time of evolution, together with other commented changes.

Keywords: agent-based modelling, agent-based social simulation, complexity, demography, values.

1 Introduction

Simulation for the study of complex social phenomena is often based on multi-agent systems, drawing on micro motives to explain and cause individual decisions and interactions of the participant agents [20]. The role of the experimenter is in this case to design the appropriate simulation setting and to observe the outcomes, both in terms of overall macro behaviours, as well as individual agent trajectories (histories over time) [2]. Many times, this aggregate behaviour is called emergent, as the collective and even individual behaviours could not be predicted or expected (or predictable) from the initial settings of the simulation [4].

The Mentat model [8] takes a relatively different approach from the one described above. In fact, it uses data from a survey conducted over thousands of people across Europe to set up the initial state of the simulation and to justify the behaviour that agents exhibit along the simulation. Instead of using individual motives for individual behaviour, in Mentat the agent heterogeneity

N. David and J.S. Sichmann (Eds.): MABS 2008, LNAI 5269, pp. 167–182, 2009.

is achieved by a thorough examination of the aggregate measures taken over those thousands of agents in the survey. Specific domain knowledge is used to isolate the relevant groups, and then aggregate statistical measures are used to characterise the agents in those groups and their behaviour.

Since the survey has been repeated in time, it was possible to test out the appropriateness of this approach by comparing the survey results of 2000 with the outcomes of 20 years of simulated time starting with real data for 1980. This was made by focusing on the evolution of moral and political values for the Spanish population, and results were quite accurate.

In this paper, we have identified and isolated several issues concerning demographic evolution that seemed to be unaddressed or overly simplistic in the current Mentat version. For each of these issues, we have proposed a solution and tried it out in Mentat in a progressive deepening manner. These mechanisms allowed us to avoid the overall decrease in population numbers, which were not found in the real survey.

The methodology used to improve the model was the incremental deepening of its mechanisms. It will improve the micro-behaviour for a better explanation of the macro-level, in the same line of [5]. However, as a result of its application we don't want to explain the whole process, but the importance of demographic dynamics in the evolution of values trends.

To tackle the initial absence of under-18s in the simulation we have used the real data to generate children, their mental setting, and introduced them early in the simulation. This fills the gap for about 20% of the population that was previously absent.

Marriages were also absent from the initial population, agents would only get married later during the course of the simulation. We have proposed a method for introducing an initial set of marriages, as a more realistic picture of the population.

Another issue concerned reproduction. Where previously only an overall population average was used, we have now considered some mechanisms building on age-based clustering of reproductive women. This allows us to observe the dynamics of reproduction among several groups of women, which is of course a crucial matter in the study of moral and behavioural values.

The paper is organised into 7 sections. Section 2 provides context of the research context for the study of moral values evolution in Spain in recent years. Section 3 details the variables to be studied. Section 4 describes Mentat model as an exploratory social simulation to address the issue at hand. Section 5 begins with the description of our methodology. Afterwards, the following subsections illustrates the demographic mechanisms previously used in Mentat, criticise their adequacy, and propose deeper ones, while we describe the details of the implementation used in the deepened mechanisms, and how they affected the development of the simulation. Section 6 presents and analyses the simulation results, and finally section 7 concludes and proposes some directions for future work.

2 Research Context: The Evolution of Moral Values

Social phenomena are extremely complicated and unpredictable, since they involve complex interaction and mutual interdependence networks. Quantitative sociological explanations deal with large complex models, involving many dynamic factors, not subject to laws, but to trends, which can affect individuals in a probabilistic way. According to [16], a social system is an interrelated and hierarchical set of components which interact to produce certain behaviours. So, we can conceive our target social system as a collection of individuals that interact between them, evolving autonomously and motivated by their own beliefs and personal goals, and the circumstances of their social environment.

The idea beneath Agent-Based Social Simulation (ABSS) is that we may be able to understand this huge complexity not by trying to model it at the global level but instead as emergent properties of local interaction between adaptive autonomous agents who influence one another in response to the influences they receive [11]. Because of that, the specification of characteristics and behaviour of each agent is critical, in what it can affect the dimensions of the studied problem.

As part of the middle-term objective of increasing the usability of ABSS tools for sociologists [18,19], who are usually not skilled in computer programming, we looked for a real sociological issue to analyse. Therefore, the case under study makes an analysis of the evolution of multiple factors in Spain between 1980 and 2000, focusing on moral values and mental attitudes. This Spanish period is very interesting for research, due to the big shift on values and attitudes that the society bore then. The almost 40 years of dictatorship finished on 1975, when the country was far from Europe in all the progress indicators, including the predominant values and modernisation level. However, the observed trends of the values and attitudes evolution since then are analogous to the ones found in its EU partners. Furthermore, the change in Spain has been developed with a special speed and intensity during the studied period. The problem we are facing is to study this complex problem: the shift in the people's mentalities in this society, in this period.

The Spanish society had in 1981 some more traditional and conservative values with respect to the European average, due to the dictator regimen influence. However, in 1999, the measurements exhibit a relatively high permissiveness, showing not only a strong convergence with European values, but also positioning itself as one of the most modern countries [10,15]. The strength and rapid velocity with which these attitude change occurred in the Spanish society imply, as a starting point in 1981, the existence of strong intergenerational differences, that is, between different age groups, which are more intense in Spain than in the other Western Europe countries, as shown by the European Value Survey. The age variable ability to discriminate values and mental attitudes is in general important in all the modern industrial societies such as those from Western Europe, but that ability is especially high in Spain, not only in 1981, but along the whole addressed time period. Hence the importance of demographic dynamics as a predictor of values change during the whole period, especially in this country, which was facing a quick change process.

The problem is faced using ABSS, but which cognitive model should be used for the design of the agents? The widely used Believes-Desires-Intentions architecture (BDI) [3] has been proposed as a possibility, but it has been discarded. The BDI model, as other models based on the rational choice theory, works very well for limited contexts, with clear objectives and roles, together with consistent defined rules. Good examples are the industrial task-driven agents or firms in a perfect market. But we found it useless in this general context, where the agents' global objectives cannot be defined. Besides, the values change process cannot adapt well to the $B + D \rightarrow I$ classical pattern: typically, the change comes through the world circumstances, that change the behaviour (Intentions), and only then, afterwards, the B and D are adapted in consequence. In this case the knowledge and mental state is more a result of the external forces than the opposite.

The values are influenced by a large number of factors. We need to cope with the most important ones: gender, age, education, economy, political ideology, religiosity, family, friend relationships, matchmaking and reproduction patterns, life cycles, tolerance regarding several subjects... factors that are usually interrelated. The statistics of these variables will evolve over time, together with the agent network.

We consider that the best way to deal with this overwhelming complexity is to reduce the number of independent variables to the minimum. But several studies in the field show the difficulty of reducing the number of parameters to only a few, as many other ABSS systems do, studying only particular aspects of the problem. Then, we have decided to solve the problem loading huge amounts of empirical data into the social simulation. This way, the variables loaded are not independent anymore: they are treated as known values. The dynamic processes that are not constant can be controlled and fixed too by empirical data: in this case, *Normal* distributions with known means (average of children per couple, average of female death age...).

Therefore, the initial data for agents of the simulation has been taken from the results of the European Values Survey (EVS). EVS provides a source of quantitative information and periodical results offering also data for validation of the simulation model. EVS is run periodically (each 10 years) in all European countries, regarding a huge number of characteristics. The data is aggregated by country, so we can easily extract the answers of the sample of Spanish individuals. Those 2303 individuals are a good representative sample of the Spanish society, and we can consider that their distributions of the studied variables are equivalent to the real ones.

In the case under study, we will import the data from the EVS of 1981. There is no available data from before that, during the authoritarian regime. The following EVS (1990 and 1999) will be used for validating the results of the system, as "snapshots" of the reality.

3 Variables Studied in the Simulation

The European and World surveys on values confirm an increasing change in the values of developed industrial societies. This has suddenly occurred with strength in the Spanish society during the time period we studied (1981–1999). Inglehart [9,10] has named the macro-trends that define these change in values post-materialism and more recently post-modernisation. Halman [7] has referred to the process of social individualisation. Both concepts are closely interrelated, as well as related to the secularisation process, that has been more intense in Europe than anywhere else in the World [14].

Not in every society we will see the demographic dynamics yielding the good prediction results seen in the Spanish society. In Spain we can see an especially high discrimination ability among the values and attitudes of individuals with several ages, sustained along at least three decades. In other countries this intergenerational differentiation is much smaller, and so is the impact of demographic dynamics. An especially appropriate example is Portugal, that went trough a similar democratisation and openness phase in the same years, but does not have this accentuated differentiation.

Our simulation aims to determine to which extent the demographic dynamics explains the magnitude of mentality change in Spain. The demographic factor does not cause nor determines the change in values, but it does exert an important influence on the velocity and intensity in which it manifests, so (and for this reason) it possesses an important predictive ability for its evolution, as we will show further on. This is mostly due to the fact that the changes in values have been chiefly (but not exclusively) generational changes, hence the generational replacements (say, the death of elders, carriers of the most traditional and conservative values, and the arrival of youngsters, bearers of emerging values) constitute a significant sociological inertia.

Therefore, the values of each individual remain constant, but not their aggregation in the whole society, as its demography is changing with time. This reflects the inter-generational changes, but not the intra-generational ones, also important enough. However, this isolation is the only way to analyze and appreciate the predictor effect of the demographic dynamics.

The variables chosen for mentality change indicators, attached to the macro-tendencies of individualisation, postmodernisation and secularisation have been, on one hand, a set of signals for individual moral tolerances: abortion, divorce, euthanasia and suicide, and on another, a signal for religiosity, which classifies the several degrees of religiosity of the population in the categories of: clericals, low-intensity religious, alternative-critical with church hierarchy, and finally 'not religious.'

Furthermore, we exerted control to ensure that the evolution of demographic variables provide an adequate response (we provide data relative to the most important: gender and age). We also considered other variables, such as the socioeconomic status, an indicator of education level, and another of ideological positioning. The latter variables are not the focus of study, nor they are controlled in the model, but were nevertheless considered to verify the degree to which the simulation was realistic in this other social and socioeconomic dimension. The indication of education level is defined by the age in which school was abandoned as a main activity. The indicator of social status is a standardised factor (average zero and standard deviation one), and ideally it should stay (close to) constant in time. The indicator for political ideology was obtained from a scale of ideological self-positioning in the spectrum of leftwing-rightwing positions: 1 means extreme leftwing and 10 means extreme rightwing.

4 MENTAT: Data Driven Social Simulation Model

In previous works, we can see a progressive increasing permissiveness and moral relativism, an important religious secularisation and a certain decrease in ideological dogmas, concomitantly with an increase in the 'postmodern' [10] and individualising [7] sensibilities. Reciprocally, these trends are interconnected among them. On the other hand, for the purpose of this study, it is especially relevant to show that, according to Ronald Inglehart [9,10], basic values are acquired in a relatively early stage in life, between the adolescence and the first youth, and these will change very little later in life. This is the theoretical principle that this prominent sociologist uses to explain the fact that the values changes that occurred in Spain in the second half of XX century and until now, are indeed intergenerational changes. The explanation of the changes arises from the environmental conditions in which the socialisation of the new generations occurs.

Having in mind the generational differences so accused in Spain, in what respects to values and attitudes, from the previous theory we infer that the demographic dynamics in Spain should reach a high predictive value of the prediction of such attitudes for several decades, since the generational gaps will tend to keep along time. A way of showing that this theory is right is to study the effect of demographic dynamics without taking into account the intergenerational changes. If the theory is right, we should observe that the simulation data from 1981 should adjust relatively well (even when time is taken into account) to those obtained empirically in the later curves of 1991 and 1999.

As it has been mentioned, the methodology of research chosen will be the ABSS, taking care in detail of the underlying sociological model, based on the works of studies of moral values of [12,13]. A first prototype of MENTAT was implemented, as explained in [17] and [8]. What is presented here will try to improve the results of these other works.

The Multi-Agent System (MAS) designed was developed using agents with several attributes: from the most simple ones such as sex or age, to others like economic class, ideology, or divorce acceptance. The attributes have been carefully selected as the most important for our purposes: the evolution of moral values. The population in the agents society (as in real societies) also experiments demographic changes: individuals are subject to life cycle patterns: they are born, can find a couple, reproduce and die, going through several stages where they follow some intentional and behavioural patterns. Besides, the agents can build and be part of relational groups with other agents. They communicate with their wide Moore neighbourhood, and depending on their rate of similarity, occasionally leading to friendship relationships. On the other hand, after finding a couple they can build family nuclei as children are born close to their parents.

The Mentat system may be configured to follow the parameters (such as average number of children per couple, or mean of male average age of death) from a specific country or import data from surveys that specify the attributes of the agents, reflecting the behaviour of the given population (and as mentioned, the EVS will be imported).

Besides, due to the relative simplicity of the agents, it can manage thousands of them, reaching the necessary amount for observing an emergent behaviour that results from the interactions of individuals, leading to the appearance of social patterns than can be studied [2]. Study is helped by the capacity of plotting several graphs, during and after the execution of the simulation, which reflect the evolution of the main attributes of the social system (some will be shown later on). These possibilities are possible thanks to the potential of Repast, the leading framework in java programming for social simulation. The system robustness has been tested enough to demonstrate the stability of the results, needed for the macroscopic comparative analysis.

Now we are focused in the study of the influence of demographic dynamics in inter-generational change. Therefore, not intra-generational evolution has been implemented yet: the agents' mental states don't evolve internally in time (they will always have the same level of education, ideology, or acceptance of divorce). The global evolution of the variables that can be observed is due to the changes on the demography (the old agents are dying, new ones are born after reproduction, but they have very different values).

Thus, in the micro level the agents' interactions were modelled, representing the socialisation processes. Note that these processes will not be completely represented until the model includes the intra-generational dynamics, with the influence between agents and so, values diffusion. Now they are limited to establishing friendship and couple links with the aim of reproduction, together with the family links. In the meantime, the evolution of the different patterns can be observed on the macro level.

Taking the EVS of 1981 of Spain, a sociological approach provided a spreadsheet for the characterisation of the group of individuals which statistically represents the Spanish population. In the original system, only 500 individuals were managed. However, nowadays Mentat is able to deal with all the individuals presented in the Spanish part of the EVS: 2303 individuals loaded as the same number of agents. This relation one-to-one allows a better behaviour.

Thus, these data were taken as input to generate the diverse and representative population of agents in the model, which was simulated for a period of nearly 20 years, till year 1999. As it has been mentioned, the evolution of the big number of variables studied can be compared with the data of the other two EVS, validating the results of the ABSS.

Note that as the system is non-deterministic, the graphical results have some variations at each execution. The outcome of the model should not be taken as a static output. In all our simulations the trends were always very similar, even though the exact data have some small comparison errors. Therefore, the system executions have structural similarity, as defined in [6].

It has to be remarked that, as long as we are analysing the evolution of moral values and socio-cultural phenomena (deeply interrelated) during a long period of time (like 20 years), and we pretend to achieve a simulated output in equivalent quantitative terms to their real evolution, it becomes a need to implement an analog demographic pattern of the Spanish one. And we will proceed with this task in the following section.

5 Demographic Mechanisms

5.1 Incremental Deepening as a Methodological Approach

When building up experimental designs, it is usual to defend and adopt the so-called KISS ("keep it simple, stupid!") principle [2]. In some sense, Sloman's "broad but shallow" design principle starts off from this principle [21,22]. Still, models must never be simpler than they should. The proposed solution for this tension is to take the shallow design and increasingly deepen it while gaining insight and understanding about the problem at hand. The idea is to explore the design of agents, interactions, institutions, societies and finally experiments (including simulations and analysis of their outcomes) by making the initially simple (and simplistic) particular notion used increasingly more complex, dynamic, and rooted in consubstantiated facts (see [1]). As Moss argued in his WCSS'06 plenary presentation, "Arbitrary assumptions must be relaxed in a way that reflects some evidence." This complex movement involves the experimenter him/herself, and according to Moss includes "qualitative micro validation and verification (V&V), numerical macro V&V, top-down verification, bottom-up validation," all of this whereas facing that "equation models are not possible, due to finite precision of computers." Therefore, this paper will follow the same line of his paper [5].

A possible sequence of deepening a concept, representing some agent feature, (say parameter c, standing for honesty, income, or whatever) could be to consider it initially a constant, then a variable, then assign it some random distribution, then some empirically validated random distribution, then include a dedicated mechanism for calculating c, then an adaptive mechanism for calculating c, then to substitute c altogether for a mechanism, and so on and so forth.

The deepening procedure is meant to be used in a broad methodology in which the space of possible designs for agents, societies and experiments is explored through a combination of techniques that allow to discover the most adequate features of models, having in mind the increase of insight that the models and their explorers possess. Some of the ultimate goals of social simulation include the explanation of phenomena, prediction of future outcomes, and even prescriptions to tune up the effects of policies. By traversing the design space, the best designs can be mixed and used to develop more accurate models to fulfill those purposes.

5.2 Introducing Missing Children

The initial demographic pattern that we have is a probabilistic one, with a representative distribution given by the EVS of 1981 (where the maximum random error limit is $\approx 2.08\%$, the trust level is 95.5% and $p = q = 50\%$). This distribution represents the Spanish population, with the adequate distribution in many different variables: education, economy, ideology, values, etc. But there is a methodological problem difficult to face: as it is based on a survey, we found no data from underaged people. Children don't make surveys, and the problem is not the limit of age, because we will always find a small range of age that is not able of giving us the appropriate answers. The problem could be ignored if we consider just the very short-term. But dealing with 20 years evolution, the problem is deep and important: in 1999, all the children in 1981 are adults (even the ones of 0 years, that in the end are 18) capable of reproducing and altering the global patterns. Furthermore, those "children" (some of them have 17 years, that would very likely reproduce in the 80's) represented the 23.33% of the total Spanish population at that time... a too big amount for being ignored, especially considering that lots of them will reproduce during those 20 years. As a side effect of these "missing children" (and as important as that: the missing children that should be born from them but are not) the population of the simulation drops more than a 20% during those 20 years (a whole old generation dies, but only a few children are born).

So the first task has been to introduce that big amount of individuals that we do not have data from. The characteristics of the underaged born after 1981 are based in the EVS-1981. Although the initial idea was to use the information of the EVS-1990 to generate them, it was discarded because of several reasons. A big part of them (the closest to 18 yeas) were much more similar to the 1981 group than to the 1990 generation. Besides, the 1990 data would give individuals already changed by the influence of other circumstances, and we want to study the influence of demography in these variables. This way we will still use the other two EVS just for validation, and the work will continue being general enough (considering that, in many cases when trying to make predictions, we don't have available "future data" compared with the initial time).

As we have available the population pyramids for these periods (in the official statistics), we can easily know how many individuals of each range of ages we need: in total, 716 new agents that are included. Their characteristics have been assigned from other existing agents of the EVS-1981, chosen randomly from the ones under 30 years. This way of proceeding is justified in that the children should be similar to the youngest ones available. We are aware that in the real society this new generation keeps values slightly more modern than the simulated ones, but with the options available, it turns to be an acceptable approximation for solving the problem without an exponential growth of complexity.

But solving this demographic problem, including the underaged with a group of characteristics equivalent to an adult, raises two new inconveniences. First, it disrupts reality as long as children are not mature enough for having stable values. Second, it makes it difficult to compare with the available empirical data (basically, surveys and studies made after 1981), always done only to people over 18 years. Both problems can be solved filtering the output of the simulation: the ones related to the variables measured (as if we were doing a survey in the actual population of the simulation). Anyway, some demographic statistics will remain unfiltered (total number of agents, percentage of children and adults, and others).

5.3 Introducing Initial Marriages

After the "jump" from the Mentat with (around) 2300 individuals to the Mentat 3000 (being precise, we should have called it Mentat 3019), we still found structural deficiencies concerning the demography management in the system. It can be observed that in the first years of simulation no agents are born. Obviously, this is completely unrealistic, and the reason of this misbehaviour is that the agents begin isolated from each other (close, but with no links between them). They invest their first years in the simulated world in finding friends, and maybe a couple. When the population as a whole has already established a robust linked network, they begin showing the expected macro output.

Then, for dealing with this small difficulty we have to introduce the needed initial marriages, as we did with the needed initial children. But, even though we have the percentage of people that should be married (and even the agents that are married or not, as it is given in the EVS), we can't know with whom (it depends on the network distribution, as an agent only have communication with his local environment).

As we couldn't load this missing information, and although we could have forced the links by inventing them, we decided to let the agents decide with their criteria. First, we loaded the information available: which agents are married/single in 1981. Second, we let the simulation begin... but freezing the years counter. In this special period (named "Phase B"), the agents neither get older, nor have children, nor die. But they do communicate with each others, building new friendship and couple links (taken into account if they should be married in 1981). After a certain period of "steps" (the minimum measure of simulated time) have passed, the "Phase B" finishes and the years counter begins. This way, the real simulation begins with the agents already linked, and since 1981 new agents are born, achieving a much more realistic global behaviour.

5.4 Deepening Population Dynamics

With the last changes, new dilemmas arise. In the previous Mentat 2300, before all these new changes, the agents only searched for a couple whenever they wanted to reproduce, so the number of total couples was the same as the number of couples with children. But now, taking into account the agents already married before 1981, this doesn't make sense. The most part of those old couples probably had already all the children they wanted to have, and shouldn't have more in the 80's and 90's. So we need a way to find out who wants to have more children.

This fact is determined by multiple factors (economic class, religiosity...) but the one that beats them all as the biggest weight in the election/possibility of having children is the age. With the data available, and with the statistics technique of nonlinear regression, we can build an equation that will allow us to determine the probability of having children of a given individual of certain age. For example, an agent of 35 years old has only a 23% of probability.

The same mechanism can be used for other needed choices. Instead of loading the information of being single from the EVS (as mentioned in the last subsection), we can calculate it through another regression equation, so it is generalised and usable for every agent (and not only for the loaded ones).

To continue to try to achieve the demographic convergence with the target, we decide to substitute the quite simple way of deciding the number of children (a Normal distribution with centre on the Spanish average in 1981) with the regression equation of birth rate considering the actual year. The change on the behaviour is immediate, as the birth rate has decreased rapidly in those years (from a 2.2 falling to 1.19). It has to be compensated with the introduction of a proper distribution for life expectancy (by the way, two equations, as they are completely different for men and women), replacing the previous Normal one. In this case, it has been increasing thanks to the improvements in health and quality of living. Lastly, we cannot ignore that the age of mothers when having the first child has been increasing significantly. Therefore, we exchanged the last normal distribution for another appropriate regression equation based on the data available.

Now the life cycle of an agent has changed, being more complex, but still easily understandable. Its much richer in information, as it uses more data from the EVS, together with some equations (generated from empirical data). However, this solution has its own limitations. The first two equations, based on age, try to solve the problem of "time," as an agent has different behaviour depending on his/her age. But the equations themselves evolve in time too. Maybe that agent of 35 years old has a 23% of probability in 1981, but a 28% in 1995. Anyway, the calculations would turn to be thorny and unnecessary complicated for the small real difference that it would entail. And the other equations, that take into account the year instead of the age, do not imply this difficulty.

6 Results and Discussion

In table 1 we can observe a comparison between the different evolutions of a sample of chosen variables, in the period studied. The measurements have been a) made by extracting the information from the three different EVS of those years, b) from statistical calculations (over several executions) in the previous version of the system, Mentat 2300, in the specified years, and c) equivalent calculations made in the new Mentat 3000. All variables are calculated considering only the individuals over 18 years. We extracted a wide range of different statistical measures to test the consequences of every change. Even so, we are showing here mainly means and percentages, as they can easily reflect the behaviour of the whole population. It is important to mention that the high stability of both versions of Mentat has simplified the analysis, and we can conformably assess that these values have a minimum error between executions.

We can begin the analysis of the results by the most simple measure: the gender. Its evolution is steady, but it has a small change in 1990. This change

Table 1. Validation: comparison between EVS, previous version and the more complex one

	EVS			Old Mentat 2300			New Mentat 3000**		
	1981	1990	1999	1981	1990	1999	1981	1990	1999
GENDER									
Men	49	47	49	49	48	48	49	48	49
Women	51	53	51	51	52	52	51	52	51
Age (mean)	45	43	46	45	51	57	45	47	49
% 65+ years	15	13	19	15	23	31	15	19	24
% 65+ years*	16*	18*	21*						
% SINGLE	28	29	29	100	82	79	100**	42**	35**
% SINGLE***							100***	34***	30***
% SINGLE****							100****	29****	28****
AGE END OF STUDIES	16	N/A	17	16	16	16	16	18	18
ECONOMIC STATUS	0	N/A	N/A	0.028	0.062	0.095	0	0.09	0.105
IDEOLOGY (mean)	4.85	4.65	4.75	4.85	4.82	4.73	4.85	4.74	4.59
IDEOLOGY (%)									
Left	29	33	31	29	30	32	29	33	36
Centre	18	19	23	18	17	18	18	18	17
N/A	30	25	24	30	30	29	30	29	27
Right	22	23	21	23	22	22	23	22	20
TOLERANCE MEASURES (scale 1 to 10)									
Abort	2.89	4.43	4.58	2.89	2.96	3.06	2.89	3.08	3.3
Divorce	4.79	5.65	6.24	4.79	4.92	5.09	4.79	5.13	5.4
Euthanasia	3.18	4.17	4.95	3.18	3.24	3.34	3.18	3.43	3.6
Suicide	2.26	2.25	2.95	2.26	2.29	2.35	2.26	2.36	2.5
RELIGIOUS TYPOLOGY	1981	1990	1999	1981	1990	1999	1981	1990	1999
Ecclesiastical	33	25	22	33	31	29	33	29	25
Low-Intensity	22	26	23	22	23	22	22	23	22
Alternatives	14	17	19	14	14	15	14	16	16
Non-religious	31	32	35	31	31	33	31	34	37
POPULATION GROWTH			+8%*			+1%			+7.2%**
POPULATION GROWTH***									+8.6%***
POPULATION GROWTH****									+10%****

*: The source is the Spanish population census of the years 1981, 1991 and 2001 (INE, Spain). EVS does not shows accurately these data.
,*,****: With a "Phase B" of 100 steps, 500 steps or 1000 steps, respectively

is not shown at all in the old system, but can be appreciated quite well in the new one. On the other hand, the age mean increases just a bit (+2%), but in both Mentat's we see an incorrect higher rising. Anyway, it's clear that the old version has a much worse output (+22%) than the new one (+8%).

To analyse the percentage of old people has more importance than the previous variables. Its evolution reflects an important part of the population pyramid structure, it is more sensible than the others, and it suffers bigger changes. The EVS data is not accurate here, because it takes into account other factors, but in the table it has been included the data from the census as empirical source. The increase, as before, is ridiculous in the old case, but moderately good in the new one.

The percentage of single individuals in the population is a factor related to the agents network. A single individual is the one that or a) does not have single adult opposite-sex friends to have as a couple, or b) it does not want to have a couple (for example, because it is a child). In the beginning, the agents find the problem (a), but with time the network should grow in complexity and cohesion, so the predominant problem is (b). We can see that the percentage of singles should remain quite stable near 30. However, in both Mentat's we see a surprising "100" in the beginning. This can seem weird, as in all the other measurements the 1981 value matches the EVS. The logic beneath this is that the measurement of 1981 has been done in the very beginning of the simulation, even before than "Phase B" (the introducing of friendships and marriages before the years counting). And so, in that moment, all the agents are isolated, so all find themselves with the problem (a). Obviously, the size of the "Phase B" is crucial for this variable: the more time you leave them to interact, the more couples you will have. We will leave apart the Mentat 2300, with useless output, and concentrate in the new one. We have tested with three different sizes: 100, 500 and 1,000 steps (one year, when they are been counted, is 50 steps), and the different outputs are in the table. After the "Phase B" the network acquires consistence and approaches a lot to the ideal value (that is just measuring the connectivity between nodes). We can see how a wider size gives more cohesion: with 500, we achieve the ideal after 1500 steps (500 of "Phase B" plus around a thousand of the nearly 20 years of simulation), and logically it can be achieved before with 1000, after the same 1500 steps (1000 of "Phase B" plus 500 of 10 years). The stability after that objective is achieved (in 1999 it is still nearly 30, instead of continuing the falling) matches accurately the observed reality, and can lead us to the situation where the problem (b) is the widely dominant. From the Social Network Analysis point of view, we could say that this is a property of small-world networks.

On the other hand, the size of "Phase B" has another logical effect: the more couples we have, the more children will be born. Anyway, as the most part of the young couples find someone always, it does not have a big influence in the population growth. As we can check in the end of the table, the real growth in that period is around +8%. The previous model fails again, but now, with different sizes, we achieve much better results. The side effect of the extra 1000 steps that gave great results with the percentage of singles yields a bigger error here. It has to be mentioned that there are no other important side effects, so we have not shown the other variables with other sizes different than the usual one of 100 steps.

The level of studies shows a slight rising that is approximately shown in the simulation. But the education is categorised (in five ranges), so it cannot be appreciated properly (only with the “jump” from 16 to 18). The economic status is not calculated properly in the EVS, but it is known that should remain stable or with a minimum increase, and this is exactly what happens in the simulation.

The political ideology follows a similar trend but with some more slope, in the means and in the different percentages. The tolerance measures are in the same case: slower increase than in reality, and in both measures the differences between the two ABSS are not particularly big. This is due to several facts. First, we have not modelled the intra-generational changes, so the agents main attributes remain static over time (and these variables are extremely sensible to those influences). Second, the new generation introduced (the 700 agents) have a very similar characteristics to the ones from 18 to 30 (as it has been explained), but they should be more modern than that. Finally, the simulation is not able to display the slight move to the right that occurred in the Spanish society during the governments of Joseé María Aznar (1996-2004). But that would be too much to ask from a simulation drawn from 1981 data, while Spain was still in the period of democratic transition.

One of the best indicators for the evolution of values that we have available here is the religious typology, strongly based on them. As we can see, the values are predicted with a very good accuracy, regardless the different curves that each type follows (rapid fall, hill, smooth rising and smooth growing, respectively).

7 Concluding Remarks

Overall, we can say that we accomplished our objective of improving the results of Mentat through the deepening process. The Mentat 2300 gives quite good trends in some important aspects of the simulation, like the evolution of the religious typology or the political ideology. But, as it has been commented, it has deep problems dealing with the demography of the agents population. With Mentat 3000 we have addressed these problems, keeping the good results that Mentat 2300 achieved, and improving a collection of other indicators. This methodology has been followed, isolating every part that should increase its complexity, re-implementing it, analysing the result of the step, and comparing it to the previous situation and the reality. The result is a model that is still quite simple, especially considering the problem studied. But it has been growing in complexity gradually from the previous version, so now it can deal better with several issues. We have successfully tried to combine simplicity and expressiveness.

As a sociological conclusion, we can extract that the determinism of demographic dynamics in the prediction of social trends is, as we can support by the good results, far more important than what it was expected. Further research should be addressed in this direction, better with other completely different environments that could support the same statement.

After the analysis of the current variables, we suggest that the use of Social Network Analysis (SNA) techniques in combination with the deepening process could be very helpful, as the monitoring of the consequences of every design option could be tracked. Ideally, with automatic testing programs that did not require lots of computational effort, each execution could be deeply analysed and compared with all the previous ones, to build a tree of changes with its SNA measurements associated.

Acknowledgments. This work has been performed in the context of the project "Metodos y herramientas para modelado de sistemas multiagente", supported by Spanish Council for Science and Technology, with grant TIN2005-08501-C03-01.

References

1. Antunes, L., Coelho, H., Balsa, J., Respício, A.: e*plore v.0: Principia for strategic exploration of social simulation experiments design space. In: Takahashi, S., Sallach, D., Rouchier, J. (eds.) Advancing Social Simulation: the First World Congress, pp. 295–306. Springer, Kyoto (2006)
2. Axelrod, R.: Advancing the art of simulation in the social sciences. In: Conte, R., Hegselmann, R., Terna, P. (eds.) Simulating Social Phenomena. LNEMS, vol. 456. Springer, Heidelberg (1997)
3. Bratman, M.E.: Intentions, Plans and Practical Reasoning. Harvard University Press, Cambridge (1987)
4. Conte, R., Castelfranchi, C.: Cognitive and Social Action. UCL Press, London (1995)
5. Edmonds, B., Moss, S.: From kiss to kids an 'anti-simplistic' modelling approach. In: Davidsson, P., Logan, B. (eds.) MABS 2004. LNCS, vol. 3415, pp. 130–144. Springer, Heidelberg (2005)
6. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist, 1st edn. Open University Press (April 1999)
7. Halman, L., Ester, P., de Moor, R.: The individualizing society. Tilburg University Press, Tilburg (1994)
8. Hassan, S., Pavón, J., Arroyo, M., Leon, C.: Agent based simulation framework for quantitative and qualitative social research: Statistics and natural language generation. In: Amblard, F. (ed.) Proceedings of the ESSA 2007: Fourth Conference of the European Social Simulation Association, Toulouse, France, pp. 697–707 (2007)
9. Inglehart, R.: Culture shift in advanced industrial societies. Princeton University Press, Princeton (1991)
10. Inglehart, R.: Modernization and postmodernization. Cultural, economic and political change in 43 societies. Princeton University Press, Princeton (2001)
11. Macy, M.W., Willer, R.: From factors to actors: Computational sociology and agent-based modeling. Annual Review of Sociology 28, 143–166 (2002)
12. Menéndez, M.A.: Cambio cultural y cambio religioso. Tendencias y formas de religiosidad en la España de fin de siglo. Ed Complutense. Servicio de Publicaciones, Madrid (2004)
13. Menéndez, M.A.: Hacia una espiritualidad sin iglesia. In: Tezanos, J.F. (ed.) Tendencias en identidades, valores y creencias, pp. 409–436. Fundación Sistema (2004)
14. Menéndez, M.A.: La fuerza de la religión y la secularización en Europa. Iglesia Viva 223 (2005)
15. Menéndez, M.A.: Individualización y religión en la Europa católica. Revista Española de Sociología (RES) 9 (2008)

16. Parsons, T.: The social System. Free Press of Glencoe IL, New York (1951)
17. Pavón, J., Arroyo, M., Hassan, S., Sansores, C.: Agent based modelling and simulation for the analysis of social patterns. Pattern Recognition Letters (June 2007)
18. Sansores, C., Pavon, J.: Agent-based simulation replication: A model driven architecture approach. In: Gelbukh, A., de Albornoz, Á., Terashima-Marín, H. (eds.) MICAI 2005. LNCS, vol. 3789, pp. 244–253. Springer, Heidelberg (2005)
19. Sansores, C., Pavon, J.: Visual modeling for complex agent-based simulation systems. In: Sichman, J.S., Antunes, L. (eds.) MABS 2005. LNCS, vol. 3891, pp. 174–189. Springer, Heidelberg (2006)
20. Schelling, T.C.: Micromotives and Macrobehavior. W. W. Norton & Company (October 1978)
21. Sloman, A.: Prospects for ai as the general science of intelligence. In: Proc. of AISB 1993. IOS Press, Amsterdam (1993)
22. Sloman, A.: Explorations in design space. In: Proc. of the 11th European Conference on Artificial Intelligence (1994)