DEVELOPING AND CALIBRATING AN ABM OF THE PROPERTY LISTING TASK

Enrique Canessa
Facultad Ingeniería y Ciencias, CINCO
Universidad Adolfo Ibáñez
Av. P. Hurtado 750, Viña del Mar, Chile
E-mail: ecanessa@uai.cl

Sergio E. Chaigneau
Escuela Psicología, CINCO
Universidad Adolfo Ibáñez
Av. D. Las Torres 2640, Santiago, Chile
E-mail: sergio.chaigneau@uai.cl

Carlos Barra
Independent Researcher
Las Pimpinelas 880, Concon, Chile
E-mail: c.barra@vtr.net

KEYWORDS
Property listing task, Conceptual Agreement Theory, Agent-based modeling, intersubjective variability.

ABSTRACT

We present an Agent Based Model (ABM), which shows how subjects may perform the Property Listing Task (PLT), which is widely used across psychology. In the PLT, subjects list properties that describe concepts in their minds. This ABM views the PLT as a communicative process, by which agents list properties for a concept, so that they can achieve a given level of agreement among them, i.e. produce properties that are more or less the same among agents. Using Conceptual Agreement Theory (CAT), we model that agreement in the ABM and are able to derive functional forms that the ABM's outputs should follow. The results show that agents produce agreement curves that indeed match the stated functional forms and also reasonably follow the corresponding curves obtained from empirical data. Thus, the ABM can now be used to better understand the PLT and also be further developed to model how subjects may stop listing properties for a concept according to some criteria based on the achieved level of agreement.

INTRODUCTION

The Property Listing Task (PLT) is widely used across psychology to learn about concepts that occur in natural language (e.g., CANARY, CAR, ALLIGATOR). In the PLT, subjects are asked to list properties that are true of a given concept (e.g., to the request of listing properties true of the concept CANARY, *has feathers, is a pet* and *sings* might be produced). When responses are coded (e.g., *has feathers* and *is feathered* are counted as tokens of the same property) and lists are accumulated across individuals, frequency distributions of conceptual properties are obtained. It is widely assumed that these distributions reflect concepts' underlying semantic structure (Cree and McRae 2003), which in turn is thought to reflect the statistical structure of the categories under study (e.g., McNorgan, Kotack, Meehan and McRae 2007).

Although the PLT is widely used (e.g., in cognitive psychology, social psychology, cognitive neuroscience, neuropsychology, consumer psychology), little is known about what people are doing when they perform the task (Santos, Chaigneau, Simmons and Barsalou 2011; Wu and Barsalou 2009). Perhaps, this is because the task is deceptively simple; just a report of those properties that are associated with a given concept in an individual's mind. A fact that can be easily overlooked in the PLT (and one that speaks against its purported simplicity) is that conceptual content varies across individuals. Note that unless one wants to argue that people have identical concepts in their minds, variability is a necessary feature of concepts. As will become clear throughout our current work, and in contrast to other approaches, we see this variability as central to understanding the PLT. In other approaches, the solution to the problem of variability in PLT has been to treat inter-subject variability as measurement error which needs to be discarded. Consequently, a common practice in PLT is to delete those properties produced by less than a given proportion of the sample (e.g., in McRae et al. 2005, properties produced by less than 5 out of 30 participants were dropped from the analyses).

Notably, in contrast to all previous research on this topic, in the current work we deliberately focus on conceptual variability. As will become clear in what follows, in our analysis the PLT is not a neutral report of conceptual content. Rather, it is a communication task in which people try to produce a list of properties with which others would agree. To achieve this agreement, and given that people probably know that others may view a given concept differently (e.g., due to learning, context, or point of view), they need to carefully tailor their property lists. Thus, our main goal in the present study is to model the PLT as a communication task using Agent-Based Modeling (ABM) and see whether we can relate variability in conceptual content to the PLT. Additionally, we want to assess whether the ABM's outputs are similar to those obtained from the corresponding empirical data. Having such an ABM, will then allow us to better understand the PLT.

CONCEPTUAL AGREEMENT THEORY, CONCEPTUAL VARIABILITY AND THE PLT

To achieve the stated goals, we use our Conceptual Agreement Theory (CAT; Canessa and Chaigneau 2016; Chaigneau, Canessa and Gaete 2012), and use the probability of true agreement (*p(a1)*) and the probability of illusory agreement (*p(a2)*). These probabilities represent how likely it is to find that others agree with

Proceedings 32nd European Conference on Modelling and Simulation ©ECMS Lars Nolle, Alexandra Burger, Christoph Tholen, Jens Werner, Jens Wellhausen (Editors)
ISBN: 978-0-9932440-6-3/ ISBN: 978-0-9932440-7-0 (CD)

one's list in the PLT. To understand them, consider the following. In the PLT, after raw data have been collected and properties have been coded, all that researchers have are different lists of properties produced by individual subjects. Each of these lists can be thought of as a sample of conceptual content taken from the total population of properties produced by participants in the study. These samples vary in length and in content (respectively, the number of properties and the specific properties they contain). True agreement probability (*p(a1)*) is defined as the probability that any given property contained in one randomly chosen list produced by a subject in a PLT study in response to a given focal concept, will be also contained in another randomly chosen list produced in response to the same focal concept (i.e., the probability that agreement will be found regarding specific properties that are associated to the concept). This is a measure of the agreement that an individual can expect to find for his list, and it will achieve a value of 0 (zero) only when all subjects in a PLT study produce lists with completely different contents in response to a given concept. In contraposition, it will achieve a value of 1.0 only when all subjects in a PLT study produce the same list. Similarly, illusory agreement probability (*p(a2)*), is defined as the probability that a property contained in one random list produced by a subject in a PLT study in response to a given alternative concept, will be also contained in a randomly chosen list produced in response to the focal concept (i.e., the probability that agreement will be found regarding general properties that are associated to the concept). As will become clear further ahead, this measure will achieve a value of 0 (zero) only when there are no properties common to both concepts. Note that, when applied to concepts that are ordinates within the same superordinate (e.g., CANARY and PARROT are ordinates relative to BIRDS), *p(a2)* should achieve a value greater than zero because there should exist common properties for both concepts (i.e., otherwise, there would be no way for them to share the superordinate).

As shown in previous work, the mathematical formulation of CAT easily allows computing an estimate of *p(a1)* and *p(a2)* by using Equations (1) and (2) (Canessa and Chaigneau 2016):

$$p(a1) = \frac{s_1}{k_1} \tag{1}$$

$$p(a2) = \frac{s_1 u}{k_1 k_2} = p(a1)\,\frac{u}{k_2} \tag{2}$$

In Equations (1) and (2), s_1 is the average number of properties listed for a concept by subjects, k_1 is the total number of properties that describe a focal concept *C1* (i.e., the total number of properties after coding), k_2 is the same for an alternative concept *C2*, and u is the number of properties that simultaneously describe both concepts (i.e., the number of properties that belong to both concepts *C1* and *C2*). In the PLT, as people list properties, s_1 (or s_2 for the alternative concept) increases and so does k_1 (k_2). Importantly, the longer the individual lists of properties produced by subjects are, the higher the probability that those lists will contain low frequency properties that are not repeats in other lists. Take for example the list *has feathers*, *is a pet*, *tweets*, produced in response to CANARY. The longer this list becomes, the higher the probability that it will contain a property like *I have one at home*, which is probably seldom found in other lists. This implies a nonlinear relation between s_1 and k_1. To understand why, consider the following. When s_1 is small, a given increase in the mean number of properties produced yields a relatively small number of low frequency properties, and therefore, a modest increase in k_1. In contrast, when s_1 is larger, the same given increase in the mean number of produced properties yields a relatively larger number of low frequency properties, and thus a larger increase in k_1. Thus, we can state that the increase in k_1 due to an increase in s_1 depends on the value of s_1, or that the rate of increase of k_1 relative to the increase in s_1 is related to the value of s_1. This can be mathematically expressed by the following differential equation:

$$\frac{dk_1}{ds_1} = c\, s_1 \tag{3}$$

This simple equation can be solved by separating variables and integrating both sides:

$$\int dk_1 = \int c\, s_1 ds_1 \tag{4}$$

$$k_1 = c\,\frac{s_1^2}{2} + a \tag{5}$$

, where *a* is an arbitrary constant, and also without losing generality, we can write:

$$k_1 = a + b\, s_1^2 \tag{6}$$

Given that Equation (6) reflects only part of the relation between s_1 and k_1, i.e., other processes may also intervene in that relation, we can use Equation (6) as the guiding functional form in a regression analysis to see whether it empirically holds. Indeed, regression equations like Equation (6) explain on average almost 50% of the variance in k_1, and parameters *a* and *b* are statistically significant across many concepts (Chaigneau, Canessa, Barra and Lagos 2017). Thus, if our ABM models the PLT, the corresponding ABM´s outputs should also show that relationship. Moreover, replacing Equation (6) in Equations (1) and (2) yields expressions for *p(a1)* and *p(a2)*, which should also be obtained in PLT's ABM:

$$p(a1) = \frac{s_1}{k_1} = \frac{s_1}{a + b\, s_1^2} \tag{7}$$

$$p(a2) = \frac{s_1 u}{k_1 k_2} = \frac{s_1 u}{(a + b\, s_1^2)^2} \tag{8}$$

Recall that earlier we stated that the PLT is a communicative task in which individuals attempt to tailor property lists with which others will agree. The reader possesses now the necessary information to understand why. In our view of the PLT, depending on their communicational goal (i.e., being specific, being general, pursuing a combination of both, or being precise), people would probably select different properties to tailor their lists. Using *p(a1)* and *p(a2)*, we can express these goals as: (1) trying to be specific implies maximizing *p(a1)*; (2) trying to be general implies maximizing *p(a2)*; (3) trying to combine specificity and generality implies simultaneously maximizing *p(a1)* and *p(a2)*; (4) trying to be precise implies maximizing *p(a1)* and simultaneously minimizing *p(a2)*.

Regardless of which of the above-stated goals better represents what people attempt to do when performing the PLT, our ABM should reasonably model those possibilities, i.e., model *p(a1)* vs. s_1 per Equation (7) (corresponding to the first alternative: maximize *p(a1)*); *p(a2)* vs. s_1 per Equation (8) (corresponding to the second alternative: maximize *p(a2)*); *p(a1)* + *p(a2)* vs. s_1 (corresponding to the third alternative: simultaneously maximize *p(a1)* and *p(a2)*); and *p(a1)* – *p(a2)* vs. s_1 (corresponding to the fourth alternative: simultaneously maximize *p(a1)* and minimize *p(a2)*).

DESCRIPTION OF PLT'S ABM

We designed a simple ABM that implements the PLT, abiding by the *KISS* (keep it simple stupid) principle (Axelrod 1997), with the goal of modeling people's observed behavior when performing it. Note that we necessarily make some assumptions regarding how concepts are represented in their minds, which we will specify shortly. Note that because of space constraints, we don't strictly follow the ODD (Overview, Design concepts, Details) protocol (Grimm et al. 2006) to present the ABM, but we comply with including most of the material suggested in it.

First, the ABM creates *N* agents and assigns to each of them two frequency distributions of properties that describe the focal (*C1*) and alternative concept (*C2*). These frequency distributions were obtained from an actual PLT study (Devereux et al. 2014) and the ABM code has a routine that can input and process those two frequency distributions. These distributions correspond to our first representational assumption: property distributions obtained through the PLT, reflect people's average mental representation of the concepts at hand. To represent the properties, numbers are used, which identify each property. Each of those numbers has a given frequency. Note that *C1* and *C2* may have properties with the same number, which represent the common properties, and the total number of those common properties is *u*.

Second, per our previous discussion on conceptual variability, the ABM assigns to each agent's *C1* and *C2* concepts a given number of different properties (*NDP*) (properties that are not shared by any other agent). That number is sampled for each agent from U(0, *MAXNDP*), where *MAXNDP* can be set and is the maximum number of different properties that an agent can have. These *NDP*s properties correspond to our second representational assumption: as discussed above, conceptual content varies across individuals. The ABM code automatically creates those different properties, so that their respective identification numbers are different from the ones assigned to properties shared by all agents and different from the exclusive properties assigned among agents. To illustrate this whole process, imagine two agents (A1 and A2), and that for A1 *NDP* = 3 and for A2 *NDP* = 1. Also assume that *C1* = {5100,2300,4500,1400} and *C2* = {3000,5100,1400} with *FC1* = {10,8,9,10} and *FC2* = {15,16,14}, where *FC1* and *FC2* are the frequency of each property in *C1* and *C2*. Note that in this case *u* = 2 and that those frequency distributions are an input to the ABM. For this situation, A1 and A2 will have a set of properties equal to *C1* and *C2* with respective frequencies *FC1* and *FC2*. Also, A1 will have in its set of properties that represent *C1* and *C2*, three exclusive properties, for example {1, 2, and 3} for *C1* and {10, 20, 30} for *C2*. Similarly, A2 will have one exclusive property in its set of *C1*, for example {100}, and one exclusive property in set *C2*, for example {130}. These two sets are the potential properties for *C1* and *C2* in each agents' mind, which each agent might list.

After the initialization stage, the ABM performs the following actions, which comprise a simulation step:

a) From all the *N* agents, randomly select without replacement one agent.
b) This agent samples with replacement and using roulette-wheel selection, a property for *C1* (from its set of properties that represent *C1*) and another for *C2* (from its set of properties that represent *C2*), and adds them to separate lists that represent the properties listed for *C1* and *C2*.
c) Repeat actions a) and b) until all agents had done them.

It is worth noting that in step b), roulette-wheel selection, means that a property's probability of being selected and listed is equal to the ratio of its frequency to the sum of all the frequencies of the properties contained in the given set (*C1* or *C2*) for a specific agent, which include the corresponding exclusive properties. For example, for A1 and concept *C1* and property 5100, the probability of A1 selecting it is: 10 / (10 + 8 + 9 + 10 + 1 + 1 + 1) = 10 / 40 = 0.4

In each simulation step, after all agents have performed actions a) and b), the ABM calculates k_1, k_2 and *u*. That is done in exactly the same way as it is completed in the PLT coding stage. The ABM collects the lists of listed properties for *C1* and *C2* from all agents, merges them and counts the non-repeating properties for *C1* (*C2*), obtaining k_1 (k_2). For calculating *u*, the ABM counts the total number of non-repeating properties that are contained in both merged *C1* and *C2* lists. Additionally, to calculate s_1 (s_2, although we do not use this output for now), each agent reports to the ABM the number of non-repeating properties that are contained in its list of listed properties for *C1*(*C2*), and the ABM calculates the mean of those numbers. Note that per Equations (1) and (2), having k_1, k_2, s_1, and u is all we need to calculate $p(a1)$ and $p(a2)$. An example of the calculation will make the whole process clear.

For simplicity, consider the same two agents A1 and A2 as before. Suppose that A1 has listed for *C1* = {5100, 2, 5100, 2300} and for *C2* = {1400, 1400, 10, 5100}. Similarly, A2 listed for *C1* = {2300, 5100, 5100, 5100} and for *C2* = {3000, 3000, 1400, 130}. Then, for the given simulation step, the merged list of properties for *C1* = {5100,2,5100,2300,2300,5100,5100,5100} and for *C2* = {1400,1400,10,5100,3000,3000,1400,130}. Hence, k_1 = 3 (3 non-repeating properties = {5100,2,2300}), k_2 = 5 (5 non-repeating properties = {1400, 10, 5100, 3000, 130}), u = 1 (1 non-repeating property belongs to the merged list of properties for *C1* and *C2* = {5100}). Also, A1 has listed 3 non-repeating properties for *C1* and 3 for *C2*, whereas A2 has listed 2 for *C1* and 3 for *C2*. Thus, s_1 = (3 + 2) / 2 = 2.5 and s_2 = (3 + 3) / 2 = 3.0. Applying Equations (1) and (2) for *C1*: $p(a1)$ = 2.5 / 3 = 0.833 and $p(a2)$ = 0.833 x 1 / 5 = 0.167. All those values are the outputs of the ABM, which will be used in analyzing the results of experiments. Finally, for the interested reader, the ABM is coded using Netlogo version 5.3.1 (Wilensky 1999) and is available upon request from the authors, along with data that allow to replicate the experiments.

EXPERIMENTS AND RESULTS

Our general question is whether the ABM might produce curves of k_1 vs. s_1 that follow the relation stated by Equation (6) and also whether the curves of $p(a1)$ and $p(a2)$ vs. s_1 (calculated by the ABM as already shown in the previous section), mimic the ones implied by Equations (7) and (8), and also the corresponding values of $p(a1) + p(a2)$ and $p(a1) - p(a2)$ (per the discussion in the section Conceptual Agreement Theory, conceptual variability and the PLT). Additionally, those curves should reasonably match the corresponding ones obtained from empirical data.

To that purpose, we used data obtained from the Centre for Speech, Language and the Brain concept property norms (from here and on, the CSLB norms; Devereux et al. 2014). These norms contain empirical conceptual property frequency distributions obtained from a sample of subjects performing the PLT for a large number of concepts. These concepts can be organized in several superordinate sets. From the CSLB norms, we selected 5 superordinates: Musical Instruments (M.I.), Fruits (Fr.), Clothing Items (C.I.), Birds (Br.) and Weapons (W.). From each superordinate, we selected a focal (*C1*) and alternative concept (*C2*). Table 1 shows those superordinates along with the corresponding concepts and other information that will be used. Those superordinates and corresponding concepts were selected so that we could get different values for u (number of common properties between *C1* and *C2*), and for the a and b coefficients obtained from regression analysis using Equation (6) as the relational form. Note that in the regression equations, R^2 ranged from 0.23 to 0.59 (mean 0.43), so that we can generalize our findings to concepts that reasonably or more than reasonably follow Equation (6). We must note that to obtain the points necessary to calculate the coefficients of the parabolas for the CSLB data, we used the following procedure. The a and b parameters are not computed for a single concept, but for all concepts in the same superordinate (e.g., Musical Instruments) by regressing their corresponding k_1 values over their s_1 values. In contrast, our analysis of the ABM's outputs was done on each pair of concepts (thus, on only two k_1 and s_1 values; not enough to compute a single parabola). Hence, to obtain more data points at the concepts' level, we obtained the a and b coefficients for the corresponding superordinate and then used the ABM to obtain multiple points for the two concepts, which are part of the superordinate. To do so, we set up the ABM with the frequency distribution of each concept and adjusted *MAXNDP* (the only free input parameter to the ABM), so that we could approximately obtain the a and b coefficients of the superordinate. Then, we run the ABM to obtain (s_1, k_1) pairs for the concepts, i.e. we run the ABM and looked at the $p(a1) + p(a2)$ vs. s_1 curve and visually determined the corresponding values of s_1 and k_1 where that curve reached its maximum. We did that 20 times and obtained 20 different (s_1, k_1) pairs. Finally, using those pairs, we calculated a regression equation for obtaining the a and b coefficients for the pair of concepts, shown in Table 1.

Table 1: Data for superordinates and other information obtained/calculated using the CSLB norms

Superordinate	Concepts	Values calculated for a, b and u
Musical Instruments	*C1* = accordion *C2* = bagpipes	a = 26.7 b = 0.110 R^2 = 0.44 u = 10
Fruits	*C1* = apple *C2* = apricot	a = 26.1 b = 0.070 R^2 = 0.59 u = 11
Clothing Items	*C1* = blouse *C2* = cloak	a = 23.8 b = 0.096 R^2 = 0.37 u = 7
Birds	*C1* = canary *C2* = parakeet	a = 29.2 b = 0.114 R^2 = 0.54 u = 16
Weapons	*C1* = axe *C2* = machete	a = 24.7 b = 0.210 R^2 = 0.23 u = 10

Note: a and b see Equation (6) were obtained from regressing k_1 on s_1, all coeffs. are statistically significant at least at the 0.05 level

The experiments done with the ABM consisted in keeping fixed the number of agents N = 30 (because that was the number of subjects used in the CSLB norms), and entering the frequency distribution of properties for *C1* and *C2* obtained from the CSLB norms for each superordinate, one at a time. Then, for each superordinate, we adjusted by trial-and-error, *MAXNDP* (the maximum number of different properties that an agent can have), so that we could obtain a reasonable match between the outputs of the ABM and the corresponding values calculated from the CSLB data. Note that in our experiments, the only free input parameter to the ABM is *MAXNDP*. Recall that this value represents conceptual variability across individuals, and arguably cannot be measured, which justifies it being a free parameter. Given that the ABM has random processes, the final analyses were done on the mean outputs of 20 replications, for each superordinate. The stopping condition for each run was set at 400 time-steps. We selected that condition because visual analyses indicated that the values of the time series for *p(a1)*, *p(a2)*, *p(a1)* + *p(a2)* and *p(a1)* – *p(a2)* were all past the values of s_1 that optimize those expressions, per the discussion in the section Conceptual agreement theory, conceptual variability and the PLT.

To present the results, we first show the graphs of k_1 vs. s_1 (for concept *C1*) for each superordinate in Figure 1 for data obtained from the ABM.

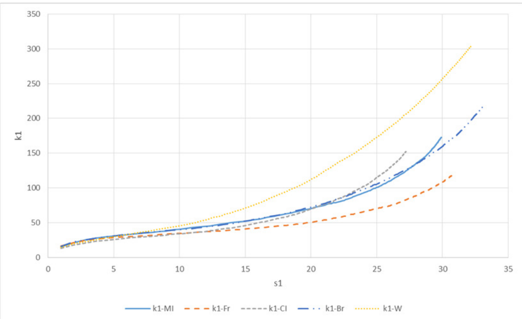


Figure 1: Avg. k_1 vs. s_1 for superordinates (data obtained from the ABM)

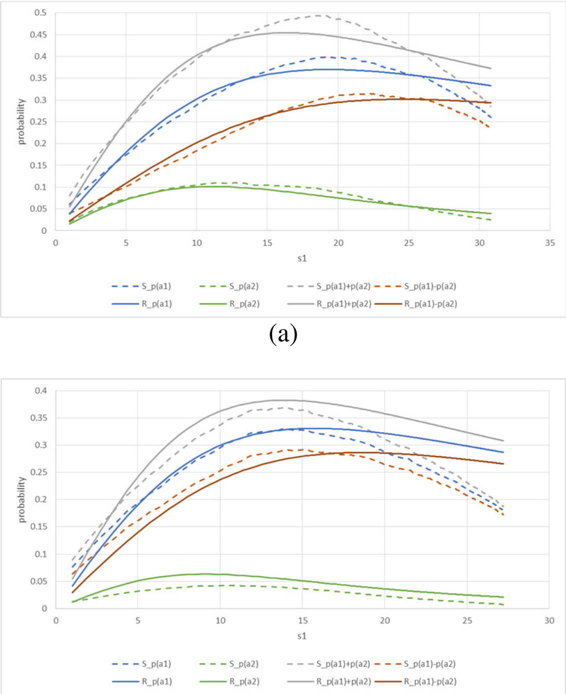
A visual analysis of Figure 1 indicates that indeed the curves of k_1 vs. s_1 for all superordinates exhibit a shape consistent with Equation (6) for positive *a* and *b* coefficients (a parabola whose axis is parallel to the y-axis and opens in the positive y-axis direction). To quantitatively corroborate that, we regressed k_1 on s_1 for a relevant range of s_1 (a range corresponding to the mean s_1 of *C1* for each superordinate ± 3 std. deviations). Table 2 shows the results of the regressions. As can be seen, the *a* and *b* coefficients obtained from the ABM data are very close to the ones of the CSLB data. Additionally, note that the R^2 values for the regression equations using the ABM data are very high, which means that the parabolas fit very well that data, i.e., the ABM data exhibit a shape very similar to the predicted parabolas. We got these suspiciously large R^2 values, because we are modelling data obtained from the same ABM, as explained before. In summary, we can say that the curves of k_1 vs. s_1 obtained from ABM's corresponding outputs fit the ones of the CSLB data for all five superordinates. Note also that these fits are achieved with only one free input parameter (i.e., the *MAXNDP*).

Table 2: Regression results of k_1 on s_1 for superordinates

Superordinate and *MAXNDP* for best fit	Reg. coeffs. for ABM data	Reg. coeffs. for CSLB data	Range of s_1 and R^2 for reg. eqs. for ABM data
Musical Instruments *MAXNDP* = 10	a = 26.5 b = 0.150	a = 26.7 b = 0.110	s_1 min = 2 s_1 max = 11 R^2 = 0.96
Fruits *MAXNDP* = 7	a = 26.0 b = 0.072	a = 26.1 b = 0.070	s_1 min = 1 s_1 max = 17 R^2 = 0.89
Clothing Items *MAXNDP* = 10	a = 23.82 b = 0.100	a = 23.8 b = 0.096	s_1 min = 3 s_1 max = 16 R^2 = 0.99
Birds *MAXNDP* = 15	a = 29.31 b = 0.103	a = 29.2 b = 0.114	s_1 min = 3 s_1 max = 18 R^2 = 0.99
Weapons *MAXNDP* = 23	a = 24.6 b = 0.211	a = 24.7 b = 0.210	s_1 min = 4 s_1 max = 18 R^2 = 0.99

Note: *a* and *b* coeffs. are statistically significant at least at the 0.05 level

The next analyses are aimed at corroborating whether the curves corresponding to Equations (7) and (8) for the real CSLB data are similar to the ones obtained from the ABM. To do that, we calculated and graphed the *p(a1)*, *p(a2)*, *p(a1)* + *p(a2)* and *p(a1)* – *p(a2)* vs. s_1 curves for the CSLB data using Equations (7) and (8) and the corresponding *a* and *b* coefficients and *u* shown in Table 1, for concept *C1*, and thus varying s_1. Then, we calculated the same curves, but using the *p(a1)* and *p(a2)* obtained from the ABM's outputs. Recall that the ABM computes *p(a1)* and *p(a2)* per Equations (1) and (2) and using the k_1, k_2, s_1, and *u* calculated by the procedure explained at the end of section Description of PLT's ABM. Thus, we are comparing different curves obtained from different data and using different procedures. Given that we do not have enough space to show all the five graphs for each of the superordinates, Figure 2 presents those curves for only two superordinates: Fruits and Clothing Items. Fruits exhibits the best match based on the Mean Absolute Percentage Error (MAPE) for the *p(a1)*, *p(a2)*, *p(a1)* + *p(a2)* and *p(a1)* – *p(a2)* vs. s_1 curves, and Clothing Items the worst one (see Table 3, second column). Figure 2 indicates that the better fit of Fruits over Clothing Items stems from the right tail of the curves. For that region, the difference between the simulated curves (dashed lines) and the real ones (solid lines) is smaller for Fruits than for Clothing Items. However, the match around the values of s_1 which maximize the curves is better for Clothing Items than for Fruits. Remember that per the discussion in the section Conceptual agreement theory, we want to model the maximization of those curves.



(a)

(b)

Figure 2: $p(a1)$, $p(a2)$, $p(a1) + p(a2)$ and $p(a1) - p(a2)$ vs. s_1 curves for real CSLB and ABM data: (a) Fruits, (b) Clothing Items (simulated = dashed lines, real = solid lines)

Thus, using only the MAPE for the entire range of s_1 might be misleading for our purposes. Hence, to better asses the fit between the simulated and real curves we calculated several goodness of fit figures. Along with MAPE, we also computed the Root Mean Squared Error (RMSE). Additionally, those two indices were calculated for the entire range of s_1 and also for a selected range of s_1. We established that range for each superordinate by taking into account the minimum s_1 and the value of s_1 that is located past the maximum value of $p(a1)$, $p(a2)$, $p(a1) + p(a2)$ and $p(a1) - p(a2)$ in those respective ABM's simulated curves (dashed lines). For example, looking at Figure 2 (a), the range is approximately from $s_1 = 1$ to $s_1 = 20$ for Fruits and Figure 2 (b) shows that for Clothing Items that range is about $s_1 = 1$ to $s_1 = 16$. We acknowledge that we used an approximate visual procedure to set those ranges, but we think it is appropriate for our purposes (see a further discussion of this point later on). Table 3 presents the indices along with the corresponding ranges (actually only the upper s_1, since the lower s_1 is always 1). From Table 3 we can see that the match of the real curves to the ones obtained from the ABM are rather good for Musical Instruments, Fruits and Weapons, all of them with a MAPE for all points below 20%. However, recall that one of our goals for building the ABM is to model the behavior of subjects when listing properties in the PLT. As already discussed, we hypothesize that people performing the PLT may try to list a number of properties (s_1) that may maximize $p(a1)$, $p(a2)$, $p(a1) + p(a2)$, or $p(a1) - p(a2)$.

Table 3: Goodness of fit indices for assessing the match between real and simulated $p(a1)$, $p(a2)$, $p(a1) + p(a2)$ and $p(a1) - p(a2)$ vs. s_1 curves

Super.	MAPE all points	MAPE range s_1	RMSE all points	RMSE range s_1
M.I.		s_1 *max* = 25		s_1 *max* = 25
	16.1%	3.59%	0.0442	0.0086
	23.0%	16.9%	0.0059	0.0072
	16.5%	4.28%	0.0490	0.0128
	16.3%	5.42%	0.0396	0.0094
	$\bar{x} = 18.0$	$\bar{x} = 7.55$	$\bar{x} = 0.0347$	$\bar{x} = 0.0095$
Fr.		s_1 *max* = 20		s_1 *max* = 20
	10.8%	5.97%	0.0427	0.0173
	20.5%	11.8%	0.0105	0.0106
	11.9%	6.44%	0.0528	0.0264
	9.59%	6.46%	0.0329	0.0113
	$\bar{x} = 13.2$	$\bar{x} = 7.67$	$\bar{x} = 0.0347$	$\bar{x} = 0.0164$
C.I.		s_1 *max* = 16		s_1 *max* = 16
	23.8%	4.67%	0.0768	0.0089
	50.4%	30.7%	0.0138	0.0174
	26.0%	7.38%	0.0890	0.0205
	22.2%	11.9%	0.0652	0.0184
	$\bar{x} = 30.6$	$\bar{x} = 13.7$	$\bar{x} = 0.0612$	$\bar{x} = 0.0162$
Br.		s_1 *max* = 23		s_1 *max* = 23
	16.4%	5.25%	0.0410	0.0115
	33.4%	6.07%	0.0096	0.0040
	18.1%	4.42%	0.0503	0.0125
	14.7%	7.95%	0.0321	0.0118
	$\bar{x} = 20.6$	$\bar{x} = 5.92$	$\bar{x} = 0.0333$	$\bar{x} = 0.0100$
W.		s_1 *max* = 20		s_1 *max* = 20
	11.9%	3.33%	0.0188	0.0065
	28.9%	14.2%	0.0032	0.0056
	12.7%	3.16%	0.0209	0.0080
	11.7%	6.67%	0.0169	0.0091
	$\bar{x} = 16.3$	$\bar{x} = 6.84$	$\bar{x} = 0.0150$	$\bar{x} = 0.0073$

Hence, the agents in our further development of the present ABM should include rules to stop listing properties accordingly. Thus, the most relevant range of s_1 for the outputs of the ABM is from the beginning of the listing process ($s_1 = 1$) to the value of s_1 that maximizes the given ABM's outputs (i.e., the simulated curves, see dashed lines in Figure 2). Given that we do not exactly know at which simulated value of s_1 agents will actually stop listing properties, and to be on the safe side, we think that a good stopping limit will be the value of s_1 past the one which maximizes the given simulated output. As Table 3 shows for those ranges of s_1 (see columns 3 and 5), the fit is much better, with a MAPE that is half or less the value of the MAPE for all the points. The same conclusions can be reached when analyzing the RMSEs. For RMSEs, the differences between RMSEs for all points and the ones for the relevant s_1 ranges are generally even bigger than for MAPEs. The ratios of RMSEs for all points to the ones for the relevant s_1 ranges go from 2.1 to 3.8, whereas the ratios for the corresponding MAPEs span from 1.7 to 3.5.

DISCUSSION/CONCLUSIONS

As we stated in this paper, the PLT is widely used across psychology, but little is known about what people are doing when they perform the task. By viewing the task as a communicative process and applying CAT, we were able to mathematically model part of it, and then developed an ABM that models the PLT from a subject's point of view. The fit obtained between real and ABM´s data is a first step in trying to model and better understand the PLT. Of course, these are preliminary results, given that we must assess whether we will obtain the same good fit to other PLT data. Additionally, given that the fit between real and ABM's data do not guarantee per se the validity of the model (Macy and Willer 2002), we must still work on validation. However, the effort in developing the ABM by finding relevant theory and practice and then merging it in a coherent model, allowed us to assess that the model may be plausible and to gain valuable insights regarding the PLT. Also, the appropriate fit of the real data to the outputs of the ABM encourages us to continue developing and using the ABM to unravel the subject's mental processes behind the PLT. We think that the development of the ABM can take two different but complementary ways. The ABM suggests that the core mechanisms embedded in the rules that govern agents are part of a probabilistic sampling process. Thus, we could refine that part of the ABM to gain further comprehension of that process and then use it to better model that part of the PLT, perhaps with a probabilistic mathematical model. Also, per our vision that subjects performing the PLT are maximizing the relation between $p(a1)$, $p(a2)$, $p(a1) + p(a2)$, or $p(a1) - p(a2)$ (i.e., maximize $p(a1)$ and minimize $p(a2)$) and the number of properties they list (s_1), we could develop an extension to the present ABM, so that agents have rules that allow them to stop listing properties when they reach the maximum of those curves. Of course, many issues must be solved to do that, like how agents will calculate $p(a1)$ and $p(a2)$ in their minds and without actually explicitly knowing the properties listed by other agents. Hence, agents will need to perform something like a PLT in their minds, using information from the property frequency distributions of concepts they have in their minds and guessing when they might have achieved the sought property listing stopping condition. Although that refinement of the ABM seems plausible, but also difficult, it will help us to better understand the PLT. The work we have done until today shows that it is possible.

ACKNOWLEDGMENT

This work was supported by CONICYT, FONDECYT (Fondo Nacional de Ciencia y Tecnologia of the Chilean Government) grant Nr. 1150074 to the first two authors.

REFERENCES

Axelrod, R. 1997. "Advancing the art of simulation in the social sciences". In *Simulating Social Phenomena 1997*, R. Conte, R. Hegselmann, and P. Terna (Eds.). Springer, Berlin, 21-40.

Canessa, E. and S. Chaigneau. 2016. "When are concepts comparable across minds?" *Quality & Quantity* 59, No. 3, 1367-1384.

Chaigneau, S.; E. Canessa; C. Barra, C.; and R. Lagos. 2017. "The role of variability in the property listing task". *Behavior Research Methods*, DOI 10.3758/s13428-017-0920-8.

Chaigneau, S.E.; E. Canessa; and J. Gaete. 2012. "Conceptual agreement theory". *New Ideas in Psychology* 30, No. 2, 179-189.

Cree, G. S. and K. McRae. 2003. "Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns)". *Journal of Experimental Psychology: General* 132, 163-201.

Devereux, B. J.; L.K. Tyler; J. Geertzen; and B. Randall. 2014. "The Centre for Speech, Language and the Brain (CSLB) concept property norms". *Behavior Research Methods* 46, No. 4, 1119–1127.

Grimm, V. et al. 2006. "A standard protocol for describing individual-based and agent-based models". *Ecological Modelling* 198, 115-126.

Macy, M.; R. Willer. 2002. "From factors to actors: Computational Sociology and Agent-Based Modeling". *Annual review of sociology* 28, 143-166.

McNorgan, C.; R.A. Kotack; D.C. Meehan; and K. McRae. 2007. "Feature–feature causal relations and statistical co-occurrences in object concepts". *Memory & Cognition* 35, No. 3, 418–431.

McRae, K.; G.S. Cree; M.S. Seidenberg; and C. McNorgan. 2005. "Semantic feature production norms for a large set of living and nonliving things". *Behavior Research Methods* 37, No. 4, 547-559.

Santos, A.; Chaigneau, S. E.; Simmons, W. K.; and Barsalou, L. W. 2011. Property generation reflects word association and situated simulation. *Language and Cognition* 3, No. 1, 83-119.

Wilensky, U. 1999. NetLogo. http://ccl.northwestern.edu/netlogo/, Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Wu, L.L. and Barsalou, L.W. 2009. "Perceptual simulation in conceptual combination: Evidence from property generation." *Acta Psychologica* 132, 173-189.

AUTHOR BIOGRAPHIES

ENRIQUE CANESSA is associate professor at U. Adolfo Ibáñez, Chile. He holds a PhD in MIS/CIS (2002), a Certificate of Graduate Studies in Complex Systems (2001) and an MBA (1991) from the University of Michigan, USA. His research interests include the study of sociology and cognitive psychology using ABM.

SERGIO E. CHAIGNEAU is full professor at U. Adolfo Ibáñez, Chile and holds a PhD in Cognitive and Developmental Psychology (2002) from Emory University, USA, and a M.A. (1995) in Experimental Psychology from the University of Northern Iowa, USA. His research interests include the study of causal categorization and inter-subjective agreement.

CARLOS BARRA holds a postgraduate degree in Integrated Political Science (2005) from the Chilean Naval Academy, an MBA (1999) from IEDE, Chile, and a M.Sc. in Computer Engineering (1996) from UTFSM, Chile. His research interests include the study of complex systems.