Psicológica (2012), 33, 631-654.

Functional Measurement: An Incredibly Flexible Tool

Etienne Mullet[*a], Guadalupe Elizabeth Morales Martinez[b], Ioannis Makris[c], Bernadette Rogé[d], Maria Teresa Munoz Sastre[d]

Institute of Advanced Studies (EPHE), France[a]; Universidad Autonoma de Nuevo Leon, Mexico[b], Athens Technological Institute, Greece[c]; Mirail University, France[d]

Functional Measurement (FM) has been applied to a variety of settings that can be considered as "extreme" settings; that is, settings involving participants with severe cognitive disabilities or involving unusual stimulus material. FM has, as instance, been successfully applied for analyzing (a) numerosity judgments among children as young as 3 years, (b) area judgment among children and adolescents blind from birth, (c) moral judgment among persons with autism and persons with learning disabilities, (d) performance judgments among completely illiterate peasants living at the border of the Sahara, (f) esthetic emotion associated with music excerpts, and (g) ethical thinking among elderly people who were about to die. The methodological aspects of these studies are presented in great detail and the main findings are discussed. The reasons why FM was successful in uncovering new, sometimes unexpected findings in these settings are discussed, notably: (a) no verbalizations/justifications were asked for; that is, judgment processes were not viewed as intrinsically connected with language, and (b) very concrete, daily life material (concrete scenarios) likely to be understood by everyone was used.

As is well-known among IIT researchers, functional measurement (FM) has been applied to a great variety of areas of everyday life psychology (Anderson, 2008). These areas include person cognition, theory of attitudes, moral algebra, group dynamics, developmental theory, consciousness analysis, person science, learning theory, functional memory, cognitive judgment-decision, language, psychophysics, psychological measurement, and analytic gestalt theory.

Less known, however, is the fact that FM has been applied to a variety of settings that can be considered as "extreme" settings; that is,

[*] Corresponding author. E-mail address: etienne.mullet@wanadoo.fr

settings involving participants with severe cognitive disabilities (e.g., Down syndrome) or settings involving complex stimulus material (e.g., musical excerpts). For example, FM has been successfully applied for analyzing (a) numerosity judgments among children as young as 3 years, (b) area judgment among children and adolescents blind from birth, (c) moral judgment among persons with autism and persons with learning disabilities, (d) performance judgments among completely illiterate peasants living at the border of the Sahara, (f) esthetic emotion associated with music excerpts, and (g) ethical thinking among elderly people who were about to die.

We will present in great detail the methodological aspects of these "extreme" studies and discuss the main findings. We will also tentatively present the reasons why, from our viewpoint, FM was, in each case, successful in uncovering new, sometimes unexpected findings.

NUMEROSITY JUDGMENTS AMONG CHILDREN AGED 3

In the early 80s, it was the common view among developmental psychologists that young children use only the length dimension in numerosity judgments of large arrays of objects. This view was based on findings from studies using discrete methodologies (e.g., Gelman, 1972). Using Functional measurement, Cuneo (1982) re-examined the way in which children as young as 3 judged numerosity of arrays of beads as a function of their length and density. The arrays of beads she designed were constructed by application of a Length x Density, 3 x 3 design. The three lengths were 6cm, 10cm, and 14cm, and the three density levels were ½, 1, and 2 beads by cm.

Instead of a yes-no response technique, she used a 19-point rating scale. This scale consisted of a series of 19 white circles (1/2cm in diameter) spaced along a black wood bar (1/2cm apart). There were two faces at the ends of the scale: a sad, frowning face on the left side and a happy, smiling face on the right side. The circles along the bar were said to stand for faces showing intermediate levels of sadness-happiness. The participants' task was to indicate, by pointing with the index, which intermediate level of sadness-happiness they associate with each situation.

The participants were familiarized to the use of the response scale in a very careful way. Firstly, they were presented with seven brown circles representing cookies of increasing size. The smallest (biggest) cookie was designated as the one that would make the little girl (boy) represented by the faces the saddest (happiest). The young participants were instructed to rate each of the remaining five cookies by pointing the index at the appropriate place on the scale. During this familiarization phase, feedback was provided to children who did not immediately understand the task.

There were two familiarization phases, one in which the five cookies were rated by increasing size, and one in which they were presented randomly.

The children who had demonstrated an understanding of the response scale were subsequently presented with rows of blue plastic beads. They were explained that these beads could be used by a little girl for making a necklace for her mother, and that the more beads in the row, the happier the little girl would be. Task instructions were designed in order to facilitate the making of intuitive estimates. Participants completed the task twice. Results are shown in Figure 1.

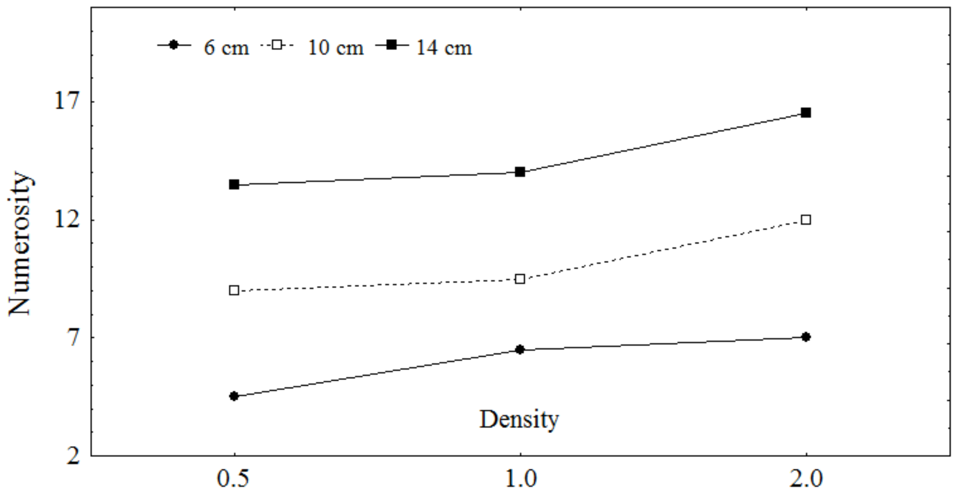

Figure 1. Children aged three years' numerosity judgments as a function of length and density of bead arrays (Cuneo, 1982).

The mean numerosity ratings are on the y-axis. The three levels of density are on the x-axis. Each curve corresponds to one level of length. The slope of the curves is ascending, which indicates an effect of density. The curves are clearly separated, which indicates an effect of length. They are roughly parallel, which indicates that the combination rule is of an additive-type. Individual analyses were conducted. They showed that the Length + Density rule holds for each young participant (except one out of 12 who showed a more complex pattern of results). In other words, this additive-type rule was not the byproduct of a situation in which, for judging, some participants would use the length information only and the remaining participants would use the density information only.

In summary, Cuneo (1982), using FM, demonstrated that children aged only 3 were able to combine length and density information for judging numerosity at a time when all developmental psychologists strongly endorsed the view that young children could only rely on one kind of information – the length of the row – for estimating total number of beads (see also Wilkening, 2007).

AREA JUDGMENT BY THE BLIND FROM BIRTH CHILDREN AGED ABOUT 6

In the early 80s, it was also the common view among developmental psychologists that young children use only one dimension for judging quantity, and this phenomenon was called "centration" (e.g., Piaget & Szemiska, 1952). Anderson and Cuneo (1978a,b; see also Wilkening, 1979) had, however, demonstrated that for judging the area of rectangles, children aged 5 did not focus on the more salient dimension of the figure (the length or the width) but, instead, were able to additively integrate both pieces of information. Nevertheless, there remained some doubts about the exact nature of the cognitive operation. Miroux re-examined the issue by studying the way blind children estimate areas (Mullet & Miroux, 1996). He identified the gestural strategies that blind children use for judging, and coupled these gestural strategies with area judgment patterns.

What can be learned from gestural strategies? Let us start in the situation of an additive rule. If children apply the principles of the Length + Width rule, then the behavior seen should involve measurement of one of the two dimensions using one of the two hands, for example by following one of the edges of the rectangle, or by placing the thumb on one of the corners and sliding one of the fingers until it reaches the end of this dimension. This first measurement technique should be repeated for the second dimension. At this time, children should want to give their answer.

If, however, children apply the Perimeter rule that was suggested by Leon (1982), it would be more likely to see that children seek the edge of the rectangle with the finger of one of the two hands, and then follow it, clockwise or anti-clockwise, until returning to the starting point. If children apply the Maximal Extent rule that was also suggested by Leon (1982), their gestural strategy would be an attempt at analyzing in terms of the diagonal or at least something one-dimensional providing a fairly direct idea of the maximal extent of the area. In concrete terms, one or both hands could be used to measure the distance between the two most opposite points of the rectangle.

Finally, if children use an area rule, they should explore the area itself with the palm of the hand or fingertips until they obtain a sufficient estimation to provide a response. This is not an absurd option owing to the fact that, because of their very blindness, young children might already have direct access to areas because of the everyday burden they must bear of the tactile exploration of objects, and hence of areas in particular. A strategy involving exploration of the interior of the area should be clearly distinguishable from a simple strategy of exploration of the contours, even if the two strategies may share a common starting point initially, that is, the identification of borders.

Participants were eight 6-year-old children blind from birth. Their visual acuity was less than 1/20, i.e., they were "legally" blind according to the international classification of the World Health Organization. Material consisted of 11 wooden rectangles approximately 1 cm thick. The dimensions of nine of these rectangles were 7 cm, 9 cm and 11 cm (as width or height). The dimensions of the other two rectangles were 5 cm by 5 cm and 13 cm by 13 cm.

The response device consisted of a wooden board with anchors at each end. These two anchors enclosed a runner containing a sliding mobile cursor. The response device was tested previously in a sample of three blind 15-year-olds. They stressed the need of being able to determine by touch the length of the response scale as represented by the runner and of being able at all times to identify the position of the cursor in relation to the two ends of the runner. The runner was first presented as a machine for expressing happiness. The left anchor of the runner corresponded to "Very unhappy" and the right anchor to "Very happy".

The 5 cm x 5 cm rectangle was the first presented. It was described as a cookie given to a very hungry child. Will this child be happy or not very happy with such a small cookie? Subjects answered that the child would certainly not be very happy. The investigator agreed and pointed out that this was the smallest cookie possible; the cursor was positioned at the left anchor. The 13 cm x 13 cm cookie was then presented. Subjects generally answered that the child would be happy. The investigator agreed and pointed out that this was the largest cookie possible; the cursor was positioned at the right anchor. The other nine rectangles were then presented in random order, and the subject placed the cursor at the place that felt most appropriate along the length of the runner.

There were four phases in all. The first, familiarization phase has just been described. The other three phases were the actual experimental phases. These followed the same procedure as the familiarization phase with the exception that the role of the investigator was limited to presenting the eleven rectangles and encouraging the subject when necessary. During phase 2, the cursor was systematically returned to the left anchor. During phase 3, the cursor was systematically returned to the right anchor. In phase 4, it was left in its previous position.

The subjects had no possibility of estimating the weight of the stimuli, which would have led to confusion between mass and area. The correlation between responses obtained for each of the three initial position of the cursor (left anchor, right anchor or previous position) was calculated first, and it proved to be very strong, about .90.

The pattern of judgments is shown in Figure 2. The mean area judgment ratings are on the y-axis. The three levels of height are on the x-axis. Each curve corresponds to one level of width. The slope of the curves is ascending, which indicates an effect of height. The curves are separated, which indicates an effect of weight. They are not parallel but converging on the right, which indicates that the combination rule is of a disjunctive-type.

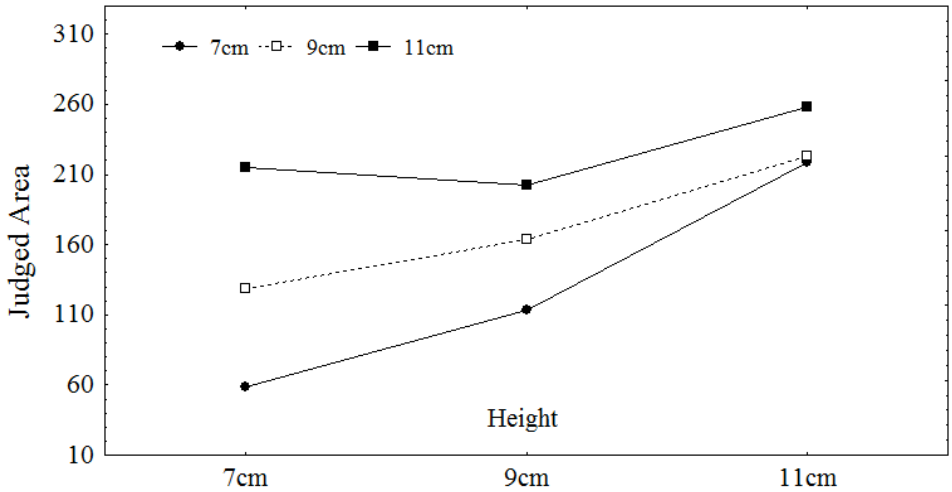

Figure 2. Blind children aged 6 years' area judgments as a function of length and width (Mullet & Miroux, 1996).

Acquisition of information concerning figures appeared to follow a fairly strict order for each subject. Typically, the subject first estimated height by placing the thumb and a finger (not necessarily the index) at the bottom and top of the rectangular object. Width was then estimated by placing the thumb and finger, in the same way, to the left and right of the object. In certain cases, the thumb was blocked on one of the edges and the finger slid along the dimension which was being estimated. This technique for the acquisition of information proved to be extensively predominant throughout the experiment and did not vary from one phase to another.

Once this process of information acquisition was completed, the subject used the cursor to provide his/her response. No systematic exploration of perimeter, surface or diagonal, with the aim of directly estimating the value, was detected.

Miroux concluded that (a) integration of information took place during the judgment process; that is, the perimeter rule and the maximal extent rule can be discarded, (b) the strategy of acquisition and integration of information was a very markedly dimensional strategy, a strategy based on length and width, and (b) the integration process obeys an additive type rule, with greater weight attributed to the larger dimension. This kind of integration rule had already been observed in studies conducted by Lautrey (Lautrey, Mullet & Pâques, 1989; see also Rulence-Pâques & Mullet, 1998). More generally, Miroux demonstrated that, once an adequate setting is provided, studies on the psychology of perception are possible with children blind from birth.

BLAME ATTRIBUTION AMONG PERSONS WITH AUTISM, AND AMONG PERSONS WITH LEARNING DISABILITIES

Autism. Blaming a person for his/her harmful actions naturally depends on the damage that has been caused but also depends on this person's intent (Anderson, in press; Hommers & Anderson, 1991). As persons with autism (PWA) have been shown to have trouble at understanding and interpreting others' intentions (e.g., D'Entremont & Yazbek, 2007), which has been associated with reduced 'theory of mind' abilities (Young, Cushman, Hauser, & Saxe, 2007), examining the way in which PWA use intent information for making moral judgments has been viewed as an important issue. Using FM, Rogé re-examined the blame schema among persons suffering from autism, paying particular attention to the use of intent in conjunction with severity of consequences for judging the appropriate level of blame (Rogé & Mullet, 2011).

The participants were ten children aged 7-14, nine adolescents aged 15-18, and six adults aged 22-36 who had received a diagnosis of autism by a qualified psychiatrist. The material was composed of six scenarios composed according to an Intent x Severity of consequences, 2 x 3 design. An example scenario (no intent, severe consequences) is the following: "Paul uses John's portable radio. Accidentally, Paul lets the radio fall down on the ground. Because of the shock, the radio is broken. It cannot be repaired. In your opinion, what level of blame does Paul deserve"?

The experimenter explained to each participant what was expected from him/her during the familiarization phase. The participant read out loud the six stories in which a person committed a harmful act against another person. After each story was read, the experimenter reminded them of the critical items of information. The participants then rated the degree of blame they would attribute to the offender. They were allowed to compare their responses and change them. In a second, experimental phase, the same six stories were presented again (in different order for each participant), and the participants provided their ratings. It was no longer possible to compare responses or to go back and make changes. Testing was individual, and there was no time limit. Importantly, no verbalization was asked from the participants.

The results are shown in Figure 3. The mean blame ratings are on the y-axis. The three levels of severity are on the x-axis. Each curve corresponds to one level of intent. Each panel corresponds to one age level. The slope of the curves is ascending, which indicates an effect of severity. The curves are separated, which indicates an effect of intent. In the right panel, curves are more separated than they are in the left panel. The intent effect was significantly weaker among children than among adults.

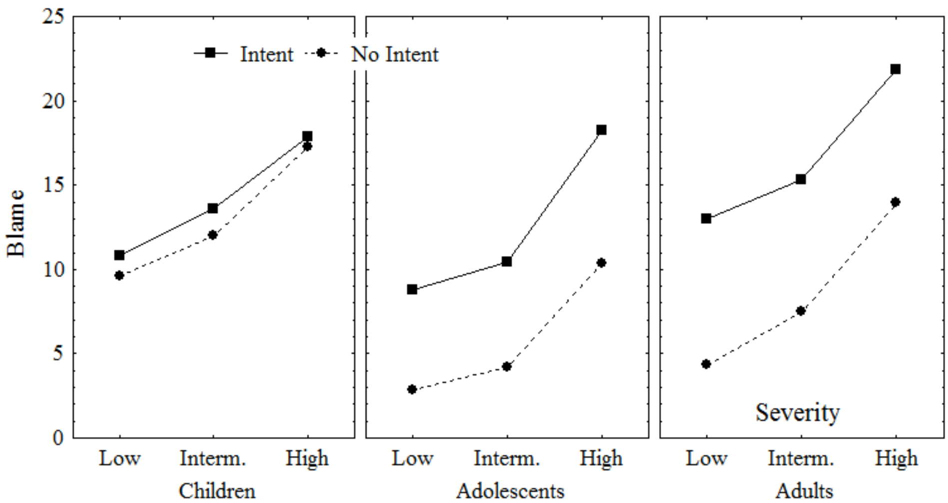

Figure 3. Persons with autism's blame judgments as a function of intent and severity of consequences (Rogé & Mullet, 2011).

In summary, Rogé was able to show that, indisputably, PWA are able to use both intent and severity information for inferring level of blame. This finding supported the view expressed by Grant et al. (2005) suggesting that PWA are able to use intent information, and with findings by Leslie and Mallon (2006) showing that basic moral judgment is substantially intact in PWAs. In other words, it would be an exaggeration to state that PWA lack an appropriate 'theory of mind'.

Learning Disabilities. Using basically the same kind of scenarios, Morales Martinez (2011) examined, for the first time, the blame schema among persons with learning disabilities (Down syndrome). Unlike Rogé, the main challenge Morales Martinez faced was finding an appropriate means for allowing participants to quantify their responses. After several unsuccessful attempts, she finally found a way using what she subsequently called the moral frog story.

Firstly, Morales Martinez familiarized her participants with daily life stories such as "Everyday Juan helps her mother to put things back at their place. Is it a good thing or is it a bad thing?", and "Very often Clara fights with her daughter. Is it a good thing or is it a bad thing"? Morales then told her participants that sometimes people do things that are bad but not that bad and at other times people can do things that are really very bad.

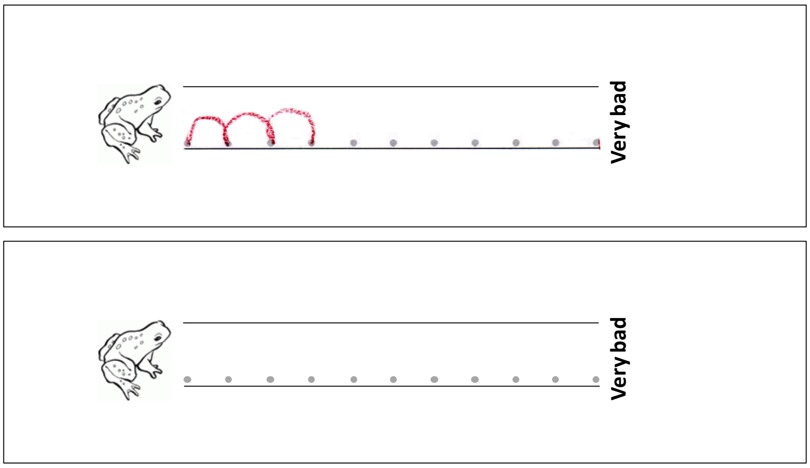

Figure 4. The way Morales Martinez explained the use of the response scale to her participants with Down syndrome is shown in the top of the figure. The scale that was used during the experiment is on the bottom of the figure.

Once the participant agreed with this view, Morales Martinez showed the frog (Figure 4) and explained that this frog indicates whether the offense that was committed was not very severe, moderately severe or very severe (by moving from its initial position to the end of the path where a sign stands, which bears the expression Very bad). If the offense is not severe, the frog jumps only one or two times. If the offense is very severe the frog may jump many more times, until eventually reaching the sign. Finally, Morales Martinez introduced the six stories by reading them aloud and showing appropriate pictures.

The results of Morales Martinez's pilot study are encouraging. Out of five persons she contacted, four were able to complete the task (25-47 years). Their results are shown in Figure 5. Effects of intent and severity of consequences are clearly present. A study using a similar device will be launched at the end of 2011. It will include samples of persons with and without Down syndrome matched by age and gender.

Using FM, Morales Martinez was able to conduct the first study on the moral judgment of persons with Down syndrome. The results of this type of study have evident ethical implications. If persons with Down syndrome are able to demonstrate the same type of moral judgment as other persons, they may deserve the same rights (see also Morales Martinez, Lopez Ramirez, Esterle, Munoz Sastre & Mullet, 2010).

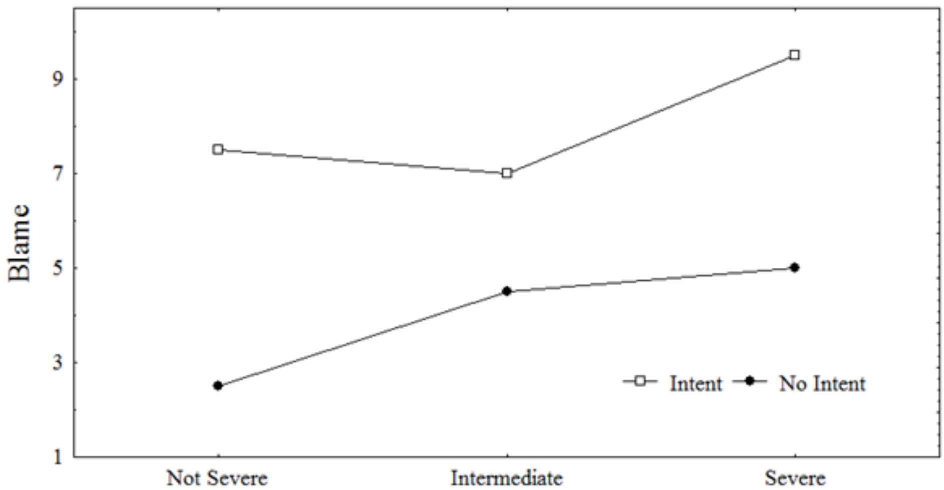

Figure 5. Persons with Down syndrome's blame judgments as a function of intent and severity of consequences (Morales Martinez, 2011).

PERFORMANCE JUDGMENTS BY ILLITERATE AFRICAN FARMERS

Anthropologists are, most of the time, reluctant at using quantitative methodologies with people living in rural areas in Africa or Asia. Breaking with this solidly established tradition, Ouedraogo, using FM, examined the performance schema (Anderson & Butzin, 1974) in a sample of illiterate farmers living in the northern part of Burkina Faso, Western Africa (Ouedraogo & Mullet, 2001). In Burkina Faso traditional culture, ability and motivation are not the only determinants of success (or failure). Success or failure is seen as partly depending on natural forces (ability and motivation), and partly depending on supernatural forces (the "wak"). In Burkina Faso farmers' views, anyone performing a difficult task can be helped by an efficient "wak" (which can be bought from the village's witch). If directly asked whether a "wak" can exert any effect on success or failure, a farmer would resist publicly declaring (e.g., to a white anthropologist) that a good "wak" can be as efficient as ability or motivation. However, in his everyday life, the farmer is intimately convinced of the reality of supernatural powers. He will choose his behaviors so as to conform to the supernatural powers' expectations.

One of Ouedraogo's many experiments in this village was precisely aimed at studying the "wak" effect in combination with ability and motivation. The material consisted of 16 stories. Twelve stories corresponded to the 12 combinations resulting from orthogonal crossover of the Ability, Motivation, and Wak factors, and four stories depicted extreme situations. There were two conditions in which performance judgments were recorded: an exam condition and a plowing condition. In the plowing condition, the two ability levels were: "physically strong and well proficient" (high ability), and "physically not very strong and moderately proficient" (low ability). The two motivation levels were: "has worked hard" (high motivation), and "has not worked hard" (low motivation). The "wak" factor had three levels: low efficiency, average efficiency, and high efficiency. The following is a typical scenario: "Ousmane is not physically very strong and is moderately proficient, but he has worked hard and he is said to have bought a very efficient wak."

Considering that most of the participants had never had any formal education and were illiterate, it was, as in the study with participants suffering from Down's syndrome reported before, necessary to construct a special device making quantitative responses possible for them. Interviews with participants about the way measures and quantities are usually expressed led Ouedraogo to construct a 60cm wooden scale with a cursor. In these populations, quantities are frequently expressed in a way that is gestural and the distance between the two hands is a common way for sharing measure information. A reasonably large wooden scale made from the branch of a tree that grew in the village, appeared thus as a convenient substitute for spontaneous measurement by hands.

Each of the 20 male farmers who accepted to participate in the study was visited individually at home. The experiment involved three phases. The first phase was a familiarization phase. The first, extreme scenario low ability-no effort-cheap wak was presented first, using the participant's own language (Moré). The experimenter explained that the right end of the wooden scale corresponded to a very large area of plowed field, and the left end of the scale to a very small area of plowed field. The response provided by the participant to the first extreme scenario was generally “very small area”. The investigator agreed and asked the participant to answer by moving the cursor to the left end of the response scale. The high ability-extremely sustained effort-excellent wak scenario was presented next, in the same way. The response offered was generally “very large area”, and the cursor moved to the right end of the response scale. Two other extreme stories were also rated in the same way.

One of the 12 scenarios making up the experimental design was then presented. The farmer was asked to move the cursor to the appropriate point on the response scale. The other stories were presented in succession, according to the individual timing of each participant. After this familiarization session, the farmer had the possibility of going back, comparing responses and making changes if necessary.

A rest period was followed by the second phase, which was the actual experimental phase. The procedure was the same as in the first phase, with the exception that the investigator no longer intervened and the participant was not allowed to go back. The third phase followed another rest period and was identical to the second. The first names used were all different from one story to another and from one phase to another.

The main results are shown in Figure 6. The mean performance ratings are on the y-axis. The two levels of ability are on the x-axis. Each curve corresponds to one level of motivation. Each panel corresponds to one level of wak efficiency. The slope of the curves is ascending, which indicates an effect of ability. The curves are separated, which indicates an effect of motivation. The set of curves in the right panel is more elevated than the set of curves in the left panel; which indicates an effect of wak. In the left panel, curves are converging to the right, but in the two other panels they are slightly converging to the left. The motivation effect was significantly weaker when "wak" efficiency was high rather than low. It was, however, significant at each of the three "wak" levels. The ability effect was significantly weaker when "wak" efficiency was average rather than high or low. It was, however, significant at each of the three "wak" levels.

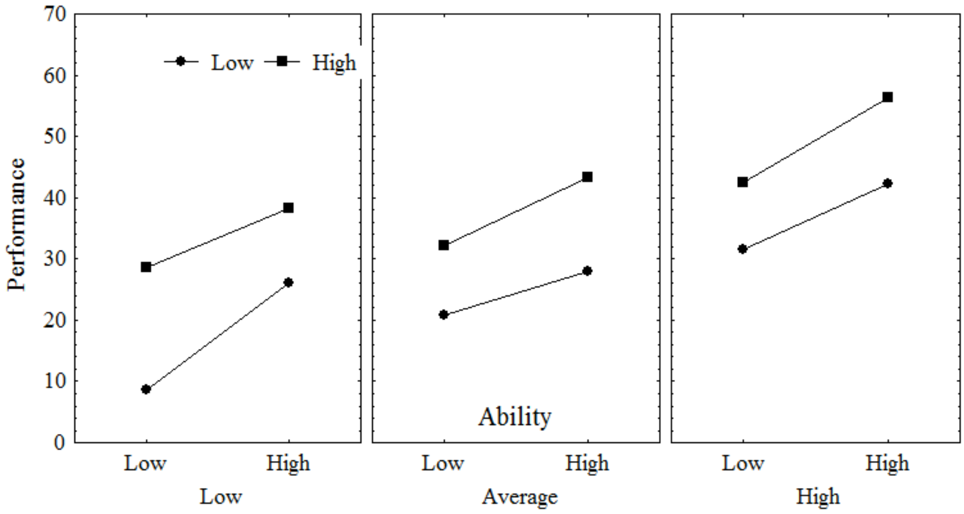

Figure 6. African farmers' judgments of performance as a function of ability, motivation, and wak efficiency (Ouédraogo & Mullet, 2001).

Ouedraogo concluded that her findings have non-negligible implications for people engaged in rural development in these geographic areas. Firstly, supernatural factors play an important role in the population's performance prediction but do not completely replace ability and motivation factors: The combined effect of the natural factors was higher than the effect of the supernatural factor. As a consequence, it can be reaffirmed that working at training people's abilities and at motivating people remain valuable objectives, despite what is sometimes said by volunteer technicians in the field who are frequently disappointed by the poor progress registered. Secondly, the faith in supernatural powers encountered in people could be positively employed by the volunteer forces rather than resisted. Before engaging in a new project in a village, the volunteer technician himself could consult the local "wakmaker" and ensure that the spirits are not hostile to the project. As strange as it may seem, such "irrational" behavior from the volunteer technician could, in many cases, considerably enhance people's motivation. It can be seen as the rational use of the irrational.

JUDGING THE PLEASANTNESS OF MUSIC EXCERPTS

Psychologists have investigated music perception since the early times of scientific psychology (Helmholtz, 1863). As a result, music perception constitutes, nowadays, a relatively rich field of investigation. Many diverse aspects of music perception have been examined, and participants in music perception studies have been faced with many different tasks: discriminating differences in rhythm, discriminating differences in chords, discriminating differences in interval, discriminating change in texture, judging consonance and dissonance, judging loudness, judging the simplicity of melodies, judging the conventionality of melodies, judging the similarity of melodic contour, differentiating melodies (that differ in contour), judging the completeness of a musical phrase, identifying melodies, recalling melodies, recalling rhythmic groups, writing melodies after having listened to them (written recall task), and even associating body movements to various musical intervals. Despite this relative abundance of studies in music perception, however, works specifically aimed at examining the way pleasantness is assessed are practically inexistent. This is surprising in view of the usual definition of music as “the art of combining vocal or instrumental sounds (or both) to produce beauty of form, harmony, and expression of emotion” (Allen, 1990, p. 781.).

Makris examined the way people judge the pleasantness of musical excerpts; that is, he examined the way in which people integrate the diverse aspects of a musical excerpt - timbre, melodic contour, rhythm and pitch level - in an overall pleasantness judgment (Makris & Mullet, 2003). In other words, Makris’ study was not about contour *per se* or pitch *per se* or rhythm *per se* or timbre *per se*, as in most previous studies, but about the conjoint effect of these factors on experienced pleasantness.

Makris faced a challenge: creating music excerpts by orthogonally combining levels of the four factors. The main difficulty was dissociating rhythm and contour in an orthogonal way. For doing so, Makris (a music composer and director of orchestra) composed an original theme. Then, he composed a second theme that was exactly symmetrical to the first theme. Then, from the first theme, he composed a third theme by simply changing the rhythm, in such a way that the contour was exactly the same. Finally, from the second theme, he composed a fourth theme that was exactly symmetrical to the third theme. As a result, the contour and the rhythm factors were totally independent. These four basic themes were played by three different instruments (guitar, flute and violin) at two different pitch levels. The 24 different compositions (2 x 2 x 3 x 2) were played by the same performer, and were recorded; that is, differences in interpretation of the music excerpts were made minimal.

The set of 24 musical items was first presented in random order for a familiarization phase. The participant was asked to listen to each item, to associate with each item a liking value, and to translate it onto the response scale. “Like much” and “Not like much” were explained as referring to a dimension of aesthetic pleasantness. Some minutes after, the same set of items was presented again in a different random order for the first experimental phase. The participants’ task was the same as in the familiarization phase. Finally, the same set of items was presented again in the reverse order as the one used in this first experimental phase. The participants’ task in this second experimental phase was the same as in the other two phases. There was no time limit. The experiment was self-paced.

The number of times each of the two levels of each factor was the preferred level was computed. Overall, 75 participants preferred the guitar, 51 preferred the flute, and the remaining 37 preferred the violin. Regarding the contour factor, participants were evenly divided. Finally, a majority of participants preferred the basic rhythm over the syncopated rhythm and preferred the low pitch level over the high pitch level.

In order to take into account these individual differences (in particular, preferences regarding the timbre), and as a result be in a better position to model the rule employed, data were reordered for each participant. For each factor, an optimal level was determined (i.e., the one associated with the highest rating.) A contra-optimal level, the one associated with the lowest rating was then determined, and an intermediate level was also determined for the timbre factor. For each participant, the 48 responses were filled in a 3 x 2 x 2 x 2 hypermatrix, composed in such a way that the order of the mean values for the rows and columns of the matrix would be monotonically increasing. From one participant to another, the preference order was slightly different, that is to say, from one participant to another, the structure of the matrices was not always exactly the same. Averaging over participants was done within these matrices; that is, averaging was done once the data had been properly reordered.

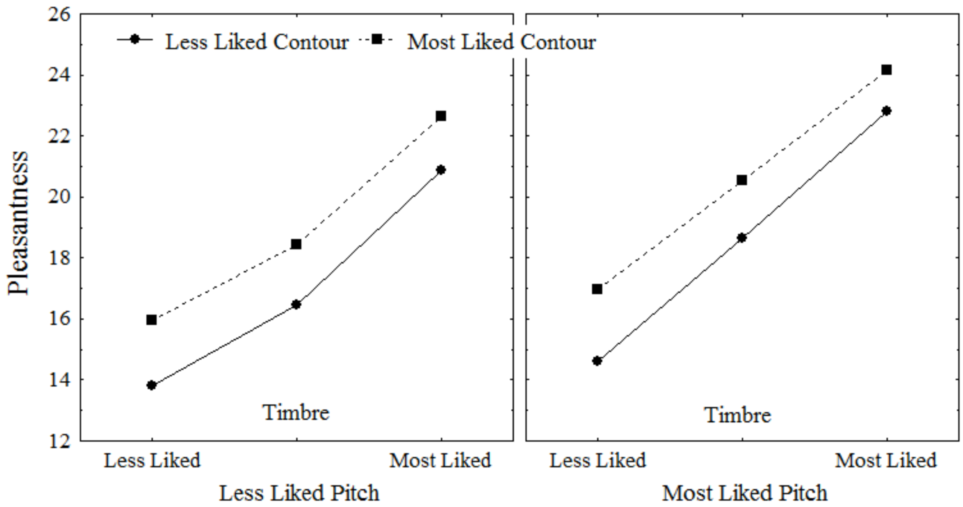

Figure 7. Pleasantness judgments about musical excerpts as a function of timbre, contour and pitch (Makris & Mullet, 2003).

The main results are shown in Figure 7. The mean likeness ratings are on the y-axis. The three levels of timbre are on the x-axis. Each curve corresponds to one level of contour. Each panel corresponds to one level of pitch. The slope of the curves is ascending, which indicates an effect of timbre. The curves are separated, which indicates an effect of contour. The set of curves in the right panel is more elevated than the set of curves in the left panel; which indicates an effect of pitch. Also, the rhythm factor was significant (not shown). All curves were roughly parallel. As a result, the following equation can synthesize the data: Pleasantness = f (Timbre + Contour + Rhythm + Pitch).

The present findings may have practical implication for music composition. Pleasantness being an additive function of the many aspects of a composition (at least the one included in the present studies), when composing a musical piece aimed at seducing listeners, composers are thus free to work in an independent way to the diverse aspects of the composition. The choice of a determined timbre for a determined theme will impact on the level of pleasantness experienced by the listeners; however, this impact will not much depend on the theme or on the overall pitch. Similarly, the choice of a certain musical contour for a melody will impact on the level of pleasantness experienced by the listeners but this impact will not much depend on the rhythm chosen for the melody: rhythm and contour may thus be separately worked on. In addition, the direction of the impact mostly depends on the listener. As shown in Study 1, there were strong individual differences in timbre preferences. Choosing violin over flute may increase the pleasantness experienced by some listeners and decrease the pleasantness experienced by some other listeners. These views about music composition are not foreign to music theoreticians, even though they were never tested before.

ACCEPTABILITY OF END OF LIFE PROCEDURES AMONG SEVERELY ILL PERSONS

Recent technological advances have transformed the act of dying by making it possible not only to alleviate pain but also to extend life. The resulting possibility of being maintained on life support for months, and, in some cases for years, has engendered anxiety among elderly patients. The patients, their families and physicians, as well as the public and policy makers have recently had to face difficult questions such as "Should a terminally-ill patient be allowed to die"?

Antonini examined the acceptability of physician-assisted suicide among persons whose lives were currently in danger (Frileux, Munoz Sastre, Antonini, Mullet, & Sorum, 2004). Among them, eight had heart disease, and seven had cancer (mean age = 70 years).

Antonini used scenarios that had already been used in former studies (e.g., Frileux, Lelièvre, Munoz Sastre, Mullet, & Sorum, 2003). These scenarios were created according to a Patient's Age x Curability of the illness x Pain x Mental health x Patient's Request, 3 x 2 x 2 x 2 x 3 design. One example scenario was the following: "Mr. Dupuis is 85 years old. He has a serious illness, totally incurable given current knowledge. He is currently receiving the best possible treatment. He suffers atrociously; pain medication cannot completely relieve his suffering. He is in good mental health. He has asked clearly and repeatedly to resort to euthanasia or physician-assisted suicide. To what extent do you think that physician-assisted suicide is an acceptable solution in this case"? The response scale was a large 35-cm scale with a left-hand anchor of "Not acceptable at all" and a right-hand anchor of "Completely acceptable."

What is remarkable about the participants in the study was that all of them were willing to contribute to the advancement of knowledge about ethical issues. In other words, despite their chronic suffering and the fact that they knew their life expectancy was short, they were interested in the study and none of them complained about the length of the scenarios. It was as if, for the first time, someone offered them the opportunity to express their opinion about one of the most important issues in life, an issue about which few interlocutors had been available.

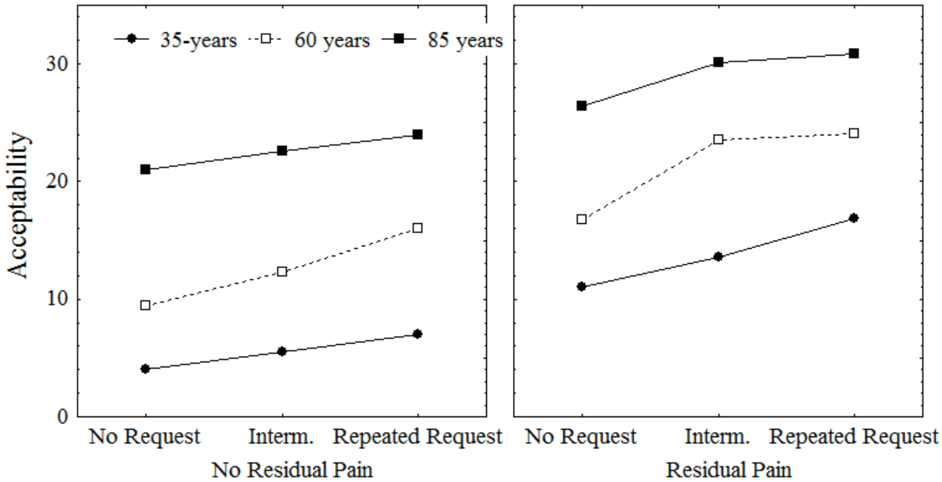

Figure 8. Elderly patients' judgments of acceptability of end of life procedures as a function of patient's request, physical suffering, and incurability of the illness (Frileux, Munoz Sastre, Antonini, Mullet, & Sorum, 2004).

The main results are shown in Figure 8. The mean acceptability ratings are on the y-axis. The three levels of request are on the x-axis. Each curve corresponds to one patient's age level. Each panel corresponds to one residual pain level. The slope of the curves is ascending, which indicates an effect of request. The curves are separated, which indicates an effect of patient's age. In the right panel, the set of curves is more elevated in relation to the y-axis than it is in the left panel, which indicates an effect of pain. The following equation can be offered as a synthesis of Antonini's findings: Acceptability = *f* (Patient's Age + Level of Pain + Incurability of the illness + Patient's Request).

In judging the acceptability of physician-assisted suicide, none of these elderly patients expressed extreme views, despite the fact that they were presented with what could be considered as a frightening situation. They took into account, in a sensible way, the concrete circumstances in which such a procedure may occur for elaborating their acceptability judgments. It can be said that, probably without being aware, they conformed to the virtue of prudence or practical wisdom that was advocated by Aristotle in his Nicomachean Ethics (1955). Antonini certainly demonstrated that elderly people can willingly contribute to the advancement of knowledge, even when they are in bad health and they know their life expectancy is short, on the condition that they are presented with meaningful material.

WHY WAS FM REPEATEDLY SUCCESSFUL IN UNCOVERING NEW, SOMETIMES UNEXPECTED FINDINGS?

No verbalizations/justifications are asked for. One reason seems to reside in the fact that in a typical IIT experiment, no verbalizations/justifications are asked for. The view IIT researchers have about judgment processes is that these processes are not intrinsically connected with language. In daily life, judgment may occur in a snap of the fingers, even when the material to be evaluated is complex. This is not to say that deliberative judgment never occurs. However, most of the time, judgment occurs without conscious deliberation, even when the instructions given and the material that is presented are verbal in nature (scenarios). This view allows, for example, IIT researchers to study judgment processes among persons who, for various reasons, are unable to verbalize (e.g., Morales Martinez, 2011), too young to do so (e.g., Cuneo, 1982), not trained to do so (Makris & Mullet, 2003) or even reluctant to do so (e.g., Rogé & Mullet, 2011). For IIT researchers, the only cognitive requirements for judging are the ones that are indicated in the well-known Anderson's information integration diagram (Anderson, 2008): functional thinking (the valuation function) and integrative capacity (the integration function).

Functional thinking is probably one of the most basic properties of mind (animal as well as human mind). As most laws in the environment (e.g., physics laws, biology laws) may be expressed in terms of functional relations between events, organisms' ability to detect these continuous relations (and learn them) has necessarily had strong adaptive value. As a result, minds (human and animal) have very probably been shaped in a way that allows detection of functional relations (in addition to detection of simple associations between pairs of discrete events). Anderson (1981) spoke of it in terms of basic metric sense.

Integrative capacity is also probably a basic property of mind. As most phenomena in the environment depend on multiple factors, an organism's ability to react to combinations of stimuli rather than reacting to only one stimulus at a time has also necessarily had strong adaptive value. As a result, minds have very probably been shaped in a way that allows integration of multiple aspects of a situation, hence the simple algebraic-type integration laws repeatedly encountered in IIT studies. As integration capabilities probably preceded the apparition of language by millions of years (see Farley & Fantino, 1978), IIT researchers have carefully avoided making the study of judgment processes among people too dependent on their mastery of language.

No preexisting normative (mathematical) models are needed. A second reason probably resides in the fact that IIT allows for the study of human judgment in everyday circumstances in daily life. In other words, IIT researchers do not limit themselves at studying judgment processes in situations for which normative (mathematical) models already exists. As most judgments of everyday life lack normative standards, confining judgment processes research to situations in which standards exist amounts to turning our backs on most everyday life situations. In addition, these mathematical models are bound to be inappropriate for the description of natural judgment processes owing to the fact that they have been created for complementing judgment processes each time these natural, inherited processes have shown their limits; that is, in areas in which the human innate capacities have been shown to be insufficient for solving specific problems (e.g., cutting meat, combining uncertainties, see Mullet, 2012).

For example, Reverent Bayes intended to create a mean for proving/disproving God's benevolence towards humans (Bayes, 1731) from the many aspects on earth that seem to support God's benevolence (e.g., love between parents and children) and the many aspects on earth that seem to disprove God's benevolence (e.g., wars and illnesses). He probably devised his theorem once he realized that there were so many aspects to be considered for proving/disproving God's benevolence that integrating them by simply using his human mind was unfeasible. In other words, Reverent Bayes probably devised his theorem once he realized that his inherited mind was not naturally Bayesian.

As a result, asking whether human mind obeys Bayes' theorem is largely nonsensical. It would amount to asking whether human teeth function as a flint biface or human legs function as a bicycle. Flint bifaces, bicycles and mathematical models have been created because human teeth are not as efficient as the flint biface for cutting skin and meat, because human legs do not run as efficiently as a bicycle and because the human mind has not the integrative capacity of a mathematical model (or a computer), respectively. If human teeth or human hands were as efficient as flint bifaces, there would have had no point at creating and using blint bifaces and similar tools. If human mind was naturally Bayesian, Reverent Bayes himself would not have felt the necessity to create his theorem.

Human mind has, however, been found to be able to integrate chance information in an efficient, if not Bayesian, way, as shown by Birnbaum & Mellers (1983), and to be able to integrate multiplicatively chance and value information as shown by Shanteau (1974). In other words, the cognitive ground on which both mathematical models have been built is already present in inherited human mind.

This does not mean that IIT findings are bound to lack structure and constraints. As illustrated in most studies that were reported above, there is a checking device for internal coherence that is built into FM procedures that allows the researcher to know when the experiment has been successful or not successful. This checking device resides in the patterns of results themselves (e.g., Hofmans & Theuns, 2008, 2010; Gamelin, Munoz Sastre, Mullet & Sorum, 2006). Simple geometric forms are usually found in these patterns (e.g., parallelism). These forms testify to the fact that a cognitive process was at work at the time of responding.

Concretely, the parallelism shown in Rogé's findings, Figure 6, testifies to the fact that persons with autism were able to combine intent and severity of consequences for attributing blame. In other words, it testifies to the fact that persons with autism were able to use a cognitive rule for judging. If a recognizable pattern had not been observed, the results would have been doubtful. In the same vein, the parallelism shown in Makris findings testifies to the fact that the participants in his study were judging pleasantness of music in a coherent, rule-governed way.

In IIT studies, structures are found in participant's responses (internal structures) and not in constrained laboratory situations (external structures). As surprising as it may be, the structures found in IIT, although being pure psychological structures, can be equivalent, in terms of metric level, to those in physics (Masin, 2007). These structures can be viewed nice illustrations of the way Nature (here, cognition) self-organizes when complexity arises (here, when several informers are present). These structures can be considered as the product of what modern physicians call an emergence (see Laughlin, 2005).

Very concrete, daily life material liable to be understood by everyone is used. A third reason, which is a consequence of the second one, resides in the fact that basic experiments on IIT use, most of the time, very concrete, daily life material (concrete scenarios) liable to be understood by everyone. For example, Ouedraogo's farmers were presented with a task where one of their companions, a strong peasant, was plowing a field with more or less enthusiasm, and were instructed to infer the amount of work done over one day. No participant objected to the reality of the situation. Cuneo's very young children were presented with beads that could be used for making a necklace. Antonini's elderly persons were presented with situations they too well know.

As advocated by ethnologists, and more generally by scientists working on animal behavior (e.g., Goodall, 1998), the view one can have about an animal's intelligence and sociality can considerably vary whether the animal is observed in the laboratory (in circumstances that are unnatural to it) or the animal is observed in its natural environment (in circumstances in which it has evolved). As human participants are also animals, this principle may also apply to them. If Ouedraogo's farmers had been presented with an abstract problem about blue cabs, green cabs, and proportions, it is likely that they would have performed poorly, independently of the problem examined by the investigator. If Antonini's elderly participants had been presented with generic questions of the type "Are you favorable to euthanasia", without specifying the context, their responses would have probably been different.

It is probably the insistence with which IIT researchers attempt to make every aspect of the experimental situation clear to the participants (e.g., Cuneo, 1982) that is the key to the success of IIT methodology. In IIT, it is the rule that data gathering can only begin once everything has been done to ensure the participants' full understanding of the task. In other fields of psychology, it is sometimes the reverse approach that applies. Experimental social psychologists sometimes dedicate much effort and ingenuity for deceiving participants. One common assumption in this field is that for learning anything important about human nature, one has to trap/trick human participants in one or the other way, hence, the importance given to debriefing after the fact. With such a research philosophy at work, it is no wonder that participants in these studies are repeatedly found to be plagued with cognitive biases and other deficiencies.

From a creationist viewpoint, these many biases and deficiencies make perfect sense. Each bias can be interpreted as an evidence of Man's imperfection in relation to his creator. From an evolutionist viewpoint, however, this set of (unrelated) biases and deficiencies is difficult to reconcile with the idea that Man's mind is the product of millions of years of evolution, that is, with the idea that the human mind has been shaped by the very environment in which human lives, and in which human's predecessors (e.g., mammals) have lived. It remains to be proved that Man would have fared better, as a species, if he had been genetically equipped with Bayes' theorem or with the expected utility model. As these rigid models only apply to a very narrow set of circumstances (e.g., when a priori probabilities are known), it is possible that a species that would have been genetically equipped with such sophisticated mental tools (instead of being equipped with rough information integration rules) would have perished a long time ago, despite being, in theory, able to be successful at solving cab problems in the psychology laboratory. These models are useful human creations that can be profitably employed by humans. They are in no way ideal models of mind, no more than flint bifaces are ideal models of teeth or wheels are ideal models of legs.

REFERENCES

Allen, R. E. (Ed.) (1990). *The concise Oxford dictionary* (8th edition). Oxford: Clarendon Press.

Anderson N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anderson N. H. (2008). *Unified social cognition*. New York: Psychology Press.

Anderson N. H. (in press). *Issues in moral science*. New York: Psychology Press.

Anderson, N. H., & Butzin, C. A. (1974). Performance = Motivation x Ability: An integration-theoretical analysis. *Journal of Personality and Social Psychology, 30*, 598-604.

Anderson, N. H., & Cuneo, D. (1978a). The height plus width rule in children's judgment of quantity. *Journal of Experimental Psychology: General, 107*, 335-378.

Anderson, N. H., & Cuneo, D. (1978b). The height plus width rule seems solid: Reply to Bogartz. *Journal of Experimental Psychology: General, 107*, 388-392.

Aristotle (1955). *The Nicomachean ethics.* London: Penguin Books.

Bayes, T. (1731). *Divine benevolence, or an attempt to prove that the principal end of the divine providence and government is the happiness of his creatures*. London: John Noon.

Birnbaum, M. H., & Mellers, B. A. (1983). Bayesian inference: Combining base rates with opinions of sources who vary in credibility. *Journal of Personality and Social Psychology, 45,* 792-804.

Cuneo, D. O. (1982). Children's judgment of numerical quantity: A new view on early quantification. *Cognitive Psychology, 14,* 13-44.

D'Entremont, B., & Yazbeck, A. (2007). Imitation of intentional and accidental actions by children with autism. *Journal of Autism and Developmental Disorders, 37,* 1665-1678.

Farley, J. & Fantino, E. (1978). The symmetrical law of effect and the matching relation in choice behavior. *Journal of the Experimental Analysis of Behavior, 29,* 37-60.

Frileux, S, Lelièvre, C., Muñoz Sastre, M. T., Mullet, E., & Sorum, P. (2003). Deciding to end a patient's life. *Journal of Medical Ethics, 29,* 330-336.

Frileux, S., Muñoz Sastre, M. T., Antonini, S., Mullet, E., & Sorum, P. C. (2004). Acceptability for French people of physician-assisted suicide. *Death Studies, 28,* 941-953.

Gamelin, A., Muñoz Sastre, M. T., Sorum, P. C., & Mullet, E. (2006). Eliciting utilities using functional methodology: People's disutilities for the adverse outcomes of cardiopulmonary resuscitation. *Quality of Life Research, 15,* 429-439.

Gelman, R. (1972). Logical capacity of very young children: Number invariance rules. *Child Development, 43,* 75-90.

Goodall, J. (1998). Learning from the chimpanzees: A message humans can understand. *Science, 282,* 2184-2185.

Grant, C. M., Boucher, J., Riggs, K. J., & Grayson, A. (2005). Moral understanding in children with autism. *Autism, 9,* 317-331.

Helmholtz, L. F. von (1863, 1954). *On the sensation of tone as a physiological basis for the theory of music* (A. J. Ellis, Ed. & Trans.). New York: Dover.

Hofmans, J., & Theuns, P. (2008). On the linearity of predefined and self-anchoring Visual Analogue Scales. *British Journal of Mathematical & Statistical Psychology*, *61*, 401-413.

Hofmans, J., & Theuns, P. (2010). Testing the impact of predefined and self-defined end anchors on the linearity of the category rating scale. *Quality & Quantity*, *44*, 397-408.

Hommers W., Anderson N. H. (1991). Moral algebra of harm and recompense. In: Anderson N. H., editor. *Contributions to information integration theory*, Vol. II: Social (pp. 101-141). Hillsdale: NJ: Lawrence Erlbaum.

Laughlin, R. B. (2005). *A different universe.* New York: Basic Books.

Lautrey, J., Mullet, E. & Paques, P. (1989). Judgment of quantity and conservation of quantity: The area of the rectangle. *Journal of Experimental Child Psychology, 47*, 193-209.

Leon, M. (1982). Extent, multiplying, and proportionality rules in children's judgments of area. *Journal of Experimental Child Psychology, 33*, 124-141.

Leslie, A. M., & Mallon, R. (2006). Transgressors, victims, and cry babies: Is basic moral judgment spared in autism? *Social Neuroscience, 1,* 270-283.

Makris, I., & Mullet, E. (2003). Judging the pleasantness of contour-rhythm-pitch-timbre combinations. *American Journal of Psychology, 116,* 581-616.

Masin, S. C. (2007). Functional measurement in physics. *Teorie & Modelli, 12*, 277-280.

Morales Martinez, G. E. (2011). *Blame judgment among persons with Down syndrome.* Unpublished manuscript.

Morales Martinez, G., Lopez Ramirez, E. O., Esterle, M., Muñoz Sastre, M. T., & Mullet, E. (2010). Judging the acceptability of sexual intercourse among persons with learning disabilities: A Mexico-France comparison. *Sexuality & Disability, 28,* 81-91.

Mullet, E. (2012). The superfluous postulate of human rationality. *Rivista Internazionale de Scienze Sociali* [Special Issue], in press.

Mullet, E., & Miroux, R. (1996). Judgment of rectangular areas in children blind from birth. *Cognitive Development*, 11, 123-139.

Ouédraogo, A., & Mullet, E. (2001). Prediction of performance among West African farmers: Natural and supernatural factors. *International Journal of Psychology, 36,* 32-41.

Piaget, J. & Szemiska, A. (1952). *The child's conception of number.* New York: Norton.

Rogé, B., & Mullet, E. (2011). Blame and forgiveness judgments among children adolescents and adults with autism. *Autism, 15,* 702-712.

Rulence-Pâques, P. & Mullet, E. (1998). Perception of Surface and Inference of Surface: The case of the Rectangle. *Journal of Experimental Child Psychology, 69,* 1-27.

Shanteau, J. (1974). Component processes in risky decision-making. *Journal of Experimental Psychology, 103,* 680-691.

Wilkening, F. (1979). Combining of stimulus dimensions in children's judgments of area: An information integration analysis. *Developmental Psychology, 15*, 25-33.

Wilkening, F. (2007). Demystifying three orthodox views of cognitive development via functional measurement. *Teorie & Modelli, 12*, 41-52.

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 8235-8240.

(Manuscript received: 28 September 2011; accepted: 9 February 2012)