MULTIPLE CRITERIA DECISION MAKING

Vol. 11 2016

Gregory Kersten[*]
Ewa Roszkowska[**]
Tomasz Wachowicz[***]

AN IMPACT OF NEGOTIATION PROFILES ON THE ACCURACY OF NEGOTIATION OFFER SCORING SYSTEMS – EXPERIMENTAL STUDY

DOI: 10.22367/mcdm.2016.11.06

Abstract

In this paper an impact of the party's negotiation profile on the misperception of the preferential information provided to the negotiating parties is studied. In particular, the problems with determining an adequate and preferentially correct negotiation offer scoring system is analyzed, when the parties are supported in their decision analyses by means of the SAW technique. In the analyses we use the negotiation data from bilateral negotiation experiments conducted by means of the Inspire negotiation support system. To determine the negotiators' profiles the Thomas-Kilmann Conflict Mode Instrument was used, which allows to describe their general negotiation approach using two dimensions of assertiveness and cooperativeness. The accuracy of scoring systems was defined as the extent to which the negotiator's individual scoring system (agent's system) is concordant to the preferential information provided by the negotiator's superior (principal's system) in the form of verbal and graphical descriptions, and measured by means of ordinal and cardinal accuracy indexes.

Keywords: negotiation profile, Thomas-Kilmann Conflict Mode Instrument, SAW, SMART, negotiation offer scoring system, preferences, electronic negotiation.

[*] Concordia University, Montreal, Canada, e-mail: gregory@jmsb.concordia.ca.
[**] University of Bialystok, Faculty of Economic and Management, Białystok, Poland, e-mail: e.roszkowska@uwb.edu.pl.
[***] University of Economics in Katowice, Department of Operations Research, Katowice, Poland, e-mail: tomasz.wachowicz@ue.katowice.pl.

1 Introduction

Various methods and techniques of operations research, and in particular – of multiple criteria decision aiding (MCDA), play an important role from the viewpoint of measuring the negotiation outcome, its quality and efficiency (Wachowicz, 2010). They are used in prenegotiation phase to help negotiators to elicit their preferences for different resolution levels of negotiation issues and determine the quantitative negotiation offer scoring system, which allows to assign numerical score to each feasible negotiation offer defined within the negotiation template and can be used throughout the whole negotiation process to support the negotiator decisions (Raiffa et al., 2002). In particular, such systems help to measure the scales of concessions, visualize the negotiation progress and conduct the arbitration analysis aimed at finding a fair and balanced negotiation agreement.

Determining the accurate scoring systems that represent the parties' preferences adequately is important from the viewpoint of providing the negotiators with reliable decision support. If the scoring system is inaccurate, the negotiator may falsely interpret the moves of their counterpart, e.g. misinterpret concession as a reverse-concession (or *vice versa*); misevaluate the profitability of alternative offers submitted by both parties during the actual negotiation phase and, finally, accept a contract that does not yield the expected and aspired profits. It is even more important, when the principal-agent context is embodied in the negotiation problem (Spremann, 1987), in which the agents negotiate in the name of their principals, and should be able to prove that their strategies and negotiated contracts are concordant with the principals' requirements and expectations. In this case determining an accurate scoring system, that reflect the preferences of the principal correctly, is of special focus. The agent, having a scoring system discordant with the principal's preference system, may negotiate in good will a contract he will consider to be good (best). Yet, the same contract will be evaluated as poor by the principal, whose preferences were not adequately represented by the agent's scoring system. Thus, it is a key issue to provide the negotiating agents with easy to use and technically accurate decision support tools that would help them to build reliable scoring systems.

Various formal decision support models are implemented in the negotiation support systems (NSS) used in business, research and training, such as OpenNexus (http://en.opennexus.pl/), Inspire (Kersten and Noronha, 1999) or Negoisst (Schoop et al., 2003). In vast majority of situations, it is the simple additive weighting (SAW) method (Churchman and Ackoff, 1954), or its discrete version called SMART (Edwards and Barron, 1994), which are used in the decision support models in negotiation mainly for their simplicity and low cognitive demand.

In discrete negotiation problems they require assigning rating points to each element of the negotiation template assuming that more preferable issues and options obtain higher ratings. Hence, any negotiation offer can be easily evaluated by adding up the ratings of options that comprise this offer. The higher the rating, the better the offer. Naturally, applying SAW or SMART to negotiation support requires the acceptance of the fundamental assumptions that the preferences are additive and preferentially independent (Keeney and Raiffa, 1976).

Even though SAW seems easy, cognitively low-demanding and technically uncomplicated, some of its drawbacks have been recently empirically discovered. For instance, it has been observed (Roszkowska and Wachowicz, 2014) that a majority (57%) of decision makers, when given a choice of the method for defining their preferences, express them qualitatively using linguistic or descriptive labels. If quantitative scores are used, they are usually of ordinal nature. This may suggest that the negotiators may make mistakes when asked to express their preferences by means of cardinal ratings instead of more intuitive qualitative judgements. This seems to be confirmed by initial analyses conducted by us in the negotiation support context (Wachowicz et al., 2015) that reveal significant problems with determining adequate scoring systems. The question, still unanswered, is which factors influence the negotiator's ability to construct the scoring systems precisely, i.e. according to the preferential information provided by their principals.

In this paper we focus on analyzing the prenegotiation process of building a negotiation offer scoring system by means of SAW by the negotiators of various negotiation profiles[1]. These profiles are determined by means of Thomas-Kilmann Conflict Mode Instrument (Kilmann and Thomas, 1977) and describe the negotiator's assertiveness and cooperativeness by means of five different conflict modes: collaborating, competing, compromising, avoiding and accommodating. In our research we analyze a dataset of electronic negotiation experiments conducted in the Inspire system, with a predefined multi-issue bilateral business negotiation case. We study whether the ability of the negotiators to transform correctly the preferential information included in the case description (provided by the principal) into a system of ratings depends on their negotiation profile described by means of five characteristics related to the conflict modes mentioned above. Inspired by earlier research by Vetschera (Vetschera, 2007), we use a negotiation case with precise graphical information about the principal's preferences. It makes possible to measure the scale of potential inaccuracy in determining the negotiation offer scoring systems by means of two separate

[1] This paper is an extension of the conference paper presented during the International Conference on Group Decision and Negotiation 2015, Warsaw (Kersten et al., 2015).

measures of accuracy: an ordinal accuracy and a cardinal accuracy measures. The former one is more general and focuses on measuring the correctness of rank order of issues and options defined in the negotiation template. The latter one refers to measuring the cardinal differences in ratings assigned to options and issues by the negotiators (agents) with the reference ratings that reflect the principal's preferences adequately.

The paper consists of three more sections. In section 2 we describe briefly the experimental setup, i.e. the bilateral negotiation experiment conducted in the Inspire system, the notions of measuring the ordinal and cardinal accuracy of scoring systems determined by the negotiators in our experiment, as well as the Thomas-Kilmann Conflict Mode Instrument (TKI) used to determine the negotiators' profiles. In section 3 we analyze the experimental results and present the key findings regarding the structures of profiles and their impact on the scoring systems' accuracy. In section 4 we present the final conclusions as well as suggest some directions for future research.

2 Experiment setup

2.1 Negotiation case

For the purpose of this paper the bilateral negotiation experiment was organized in the Inspire negotiation support system (Kersten and Noronha, 1999). In this experiment 350 students from Poland, Austria, China, Taiwan, Great Britain, Ukraine and Canada took part. The negotiation case we used in the experiment described in details a bilateral problem of signing a new contract between the entertaining agency (WorldMusic) and the musician (Ms. Sonata). The participants were asked to play the roles of agents of the agency (Mosico) and the musician (Fado) and negotiate for them the best possible contracts. The negotiation problem was defined by means of four issues, for which the feasible resolution levels were predefined in a form of a discrete negotiation template (Table 1).

Table 1: The Mosico-Fado negotiation template

Issues	Salient options
No of new songs	11; 12; 13; 14 or 15
Royalties (%)	1.5; 2; 2.5 or 3%
Contract signing bonus ($)	$125,000; $150,000; $200,000
No of promotional concerts	5; 6; 7 or 8

The negotiators were provided with private information about the goals, expectations and priorities of the principals they represented. The private information on each principal's preferences was accessible only to the agent that represented this principal. This information was both verbal and graphical, the latter one used to illustrate the differences in strength of preferences among issues and options. The graphical preference information was presented in the form of circles (Figure 1).

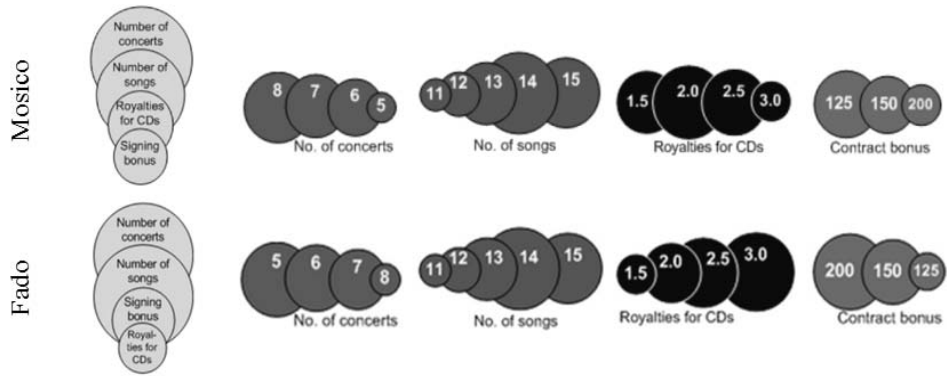


Figure 1. Graphical visualization of preferences for the Mosico-Fado case

As shown in Figure 1, the agents of Fado may learn, for instance, that the issues of number of concerts their principal (Ms. Sonata) would have to perform and number of new songs she would have to write within the contract signed with WorldMusic are the two most important issues. Both agents also knew that the resolution level "14 songs" is the best option for the principal they represent, and is somewhat more important than options: 13 and 15. The latter two are nearly equally preferred, yet 15 seems slightly better than 13.

2.2 Building the offer scoring system with SAW

Having both the negotiation template and the principal's preferences defined, as shown in Table 1 and Figure 1, respectively, the experiment's participants were supposed to determine their individual negotiation offer scoring system in the prenegotiation phase. There are many methods and techniques that can be used to determine such a system if the negotiation template consists of many issues. The problem of scoring the template is similar to an individual discrete multiple criteria decision making problem, hence a range of MCDA approaches can be of use here. A literature review reveals a few examples of using various MCDA techniques to support negotiators in rating the negotiation template, such as: UTA (Jarke et al., 1987), AHP (Mustajoki and Hamalainen, 2000) or TOPSIS

(Roszkowska and Wachowicz, 2015). Yet, the most popular technique that is applied in negotiation support systems used commonly for negotiation training, teaching or real-world problem solving is the one that derives from the multi-attribute utility and multi-attribute value theories (von Neumann and Morgenstern, 1944; Keeney and Raiffa, 1976). Formally defined as SAW, i.e. simple additive weighting (Churchman and Ackoff, 1954), it was later modified and adjusted to various decision contexts, e.g. to elicit preferences in discrete decision making problems as SMAR – simple multiple attribute rating technique (Edwards and Barron, 1994).

When applied to scoring the negotiation template, like in Inspire (Kersten and Noronha, 1999) or NegoCalc (Wachowicz, 2008) systems, SAW consists of two straightforward steps: (1) defining the issue weights (issue ratings); and (2) defining preferences for options within each issue (option ratings). Let us denote by m the number of negotiation issues and by X_j $(j = 1, \dots, m)$ the sets of salient and feasible options for issue j identified within the negotiation template under consideration. The SAW-based process of building the negotiation offer scoring systems, as implemented in the Inspire system, consists of the following steps:

- **Step 1.** Assigning the weights to each of the issues in the form of cardinal ratings so that:

$$\sum_j u_j = 100. \quad (1)$$

- **Step 2.** Assigning the ratings u_{jk} to each option $x_{jk} \in X_j$ within each negotiation issue j so that:

$$u_{jk} \in \langle 0; u_j \rangle, \quad (2)$$

and the most preferred (best) option receives the maximum score resulting from the issue weight (i.e. u_j), while the worst one, the rating equal to 0.

The SAW-based scoring system obtained by means of the above algorithm can be used to evaluate each feasible negotiation offer that can be constructed out of the salient options defined within the negotiation template. The global rating $u(A)$ of the offer A under consideration is determined as the result of additive aggregation of the ratings assigned to each option that comprise this offer, i.e.:

$$u(A) = \sum_{j=1}^{m} \sum_{k=1}^{|X_j|} z_{jk}(A) \cdot u_{jk}, \quad (3)$$

where $z_{jk}(A)$ is a binary multiplier denoting if the option x_{jk} comprises an offer A (1) or not (0).

According to the rating rules described by the SAW-based rating procedure (steps 1 and 2), the best offer within the template, i.e. the one that consists of the most preferred options, will be scored with 100 rating points, while the worst offer – with the score of 0.

2.3 Measuring the accuracy of scoring systems

For the purpose of this experiment two notions of scoring system accuracy were used, both measuring the concordance of the agent's individually built scoring system with the principal's preferential information provided within the case description as private information (see Figure 1) (Roszkowska and Wachowicz, 2015). These were the *ordinal accuracy* and the *cardinal accuracy* measures.

Ordinal accuracy

Ordinal accuracy of the agent's scoring system measures the extent to which the rank order of the preferences defined by this agent for issues and options is concordant with the rank order of the principal's preferences. More precisely, if we consider n alternatives $A_1, A_2, \dots, A_n$ (that represent various options or issues in the negotiation template) ordered according to non-increasing preferences of the principal, the notion of perfect ordinal accuracy requires that cardinal ratings $u(A_i)$ assigned by the agent to each alternative i fulfill the following condition:

$$u(A_1) \geq u(A_2) \geq \cdots \geq u(A_n). \tag{4}$$

Measuring the ordinal accuracy of the whole scoring systems requires different groups of elements of the negotiation template to be considered separately. While building the negotiation offer scoring system by means of the SAW method the agent assigns the scores within a two-step procedure: (1) the issue weights are declared (a ranking of m issues is defined by means of cardinal scores); (2) m series of options are evaluated (one series for each negotiation issue $j = 1, \dots, m$). Consequently, the agent assigns scores to $m + 1$ series of alternatives, i.e. $m + 1$ rank orders defined by the principal's preferences need to be reflected by means of cardinal scores. In our experiment there are five rankings (series of alternatives) that describe the principal's preferences for the complete negotiation template, one for issue weights and four others describing the structure of preferences within each issue: number of concerts, number of songs, royalties and contract signing bonus (see Figure 1).

Consequently, the *ordinal accuracy index* of the agent's negotiation offer scoring system will be defined as a ratio of the number of series of ratings defined by ith agent that are concordant with the rank order of preferences defined by the principal (n_i^{con}) to the total number of series of ratings that the agent needed to define within the negotiation template. In our experiment $m + 1 = 5$, hence the *ordinal accuracy index* of ith agent's scoring system is defined in the following form:

$$OA_i = \frac{n_i^{\text{con}}}{5}. \tag{5}$$

If the agent's individual scoring system represents correctly all the possible rankings resulting from the principal's preferential information, the ordinal accuracy index is equal to 1. If none of the rankings is represented correctly by the ratings assigned by the agent, then $OA_i = 0$.

It seems clear that the ordinal accuracy of the agent's scoring system may be also measured in a more detailed way, e.g. by means of the Kendall or Spearman rank correlation indexes. Yet, in our study we are interested in the most general perception of quality of scoring systems, which summarizes its accuracy at the level of complete rankings.

Cardinal accuracy

The notion of cardinal accuracy of the agent's scoring system is introduced to measure not only if the scores assigned by the agent to the issues and options reflect adequately the order of the principal's preferences, but also the strength of these preferences. To build such a measure a kind of reference rating system needs to be determined, for which we assume that it reflects the principal's preferences precisely, and according to the verbal and graphical preference information provided to the agents. Hence, such a reference rating system for the principal's preferences defines precisely the cardinal ratings describing the issue weights (u_j^{ref}, for $j = 1, \ldots, m$), as well as the ratings of options (feasible resolution levels) defined for each negotiation issue (u_{jk}^{ref}, for $j = 1, \ldots, m$; and $k = 1, .., N_j$, where N_j is the number of options defined for jth issue).

Having the reference scoring system defined, the *cardinal inaccuracy* of ith agent's scoring system may be measured in the following way:

$$CI_i = \sum_j \left| u_j^{\text{ref}} - u_j^i \right| + \sum_j \sum_{k=1,..,N_j} u_j^{\text{ref}} \cdot \left| \bar{u}_{jk}^{\text{ref}} - \bar{u}_{jk}^i \right|, \tag{6}$$

where:

u_j^i – the rating assigned to jth issue by ith agent;

$\bar{u}_{jk}^{\text{ref}}$ – normalized rating of kth option of jth issue in the reference scoring system;

$\bar{u}_{jk}^i$ – normalized rating of kth option of jth issue in the scoring system of ith agent.

The cardinal inaccuracy index is a non-standardized measure. If all agent's ratings are the same as the reference ratings determined on the basis of the principal's preferential information, the cardinal inaccuracy index is equal to 0. The bigger discrepancy between the principal's and the agent's ratings, the bigger the CI_i value.

Please note that measuring the cardinal inaccuracy by means of formula (6) allows to avoid double counting of errors made by the agents when determining their individual scoring system. According to the two-step SAW-based procedure of defining the preferences in the Inspire system (for details see: Kersten and Noronha, 1999; Wachowicz, 2010), the issue rating assigned by the agent in step 1 is a reference value for assigning the option ratings within this issue. The most preferred option of one issue should obtain the rating equal to the weight (rating) of this issue. Hence, any mistake made by the agent at the issue level would be copied to the option level and counted twice, as the difference $u_j^{\mathrm{ref}} - u_j^i$ and then as the difference $u_{jk}^{\mathrm{ref}} - u_{jk}^i$ (where k represents the most preferred option). If the normalization of option ratings is applied, the double-counting of the mistakes at the issue level can be avoided. Yet, the normalized differences in option rating between the agent's and the principal's scoring systems need to be multiplied by the issue weights (u_j^{ref}), to be comparable with the differences determined first at the issue level.

To illustrate the problem of double counting of inaccuracies let us consider the ratings assigned by the Fado agent to the options of issue "Number of promotional concerts" and compare them with the reference ratings by the principal (Ms. Sonata herself). Both the reference and normalized ratings of the agent and the principal are shown in Table 2.

Table 2: Principal's and agent's ratings for options of „Number of concerts"

Option	Principal's ratings (u_{jk}^{ref})	Normalized principal's ratings ($\bar{u}_{jk}^{\mathrm{ref}}$)	Agent's ratings (u_{jk}^i)	Normalized agent's ratings ($\bar{u}_{jk}^i$)	Normalized difference ($\lvert\bar{u}_{jk}^{\mathrm{ref}} - \bar{u}_{jk}^i\rvert$)	Final inaccuracy ($u_j^{\mathrm{ref}} \cdot \lvert\bar{u}_{jk}^{\mathrm{ref}} - \bar{u}_{jk}^i\rvert$)
5	32	1.00	17	1.00	0.00	0.00
6	25	0.78	10	0.58	0.20	6.40
7	21	0.66	5	0.29	0.37	11.84
8	0	0.00	0	0.00	0.00	0.00

From the agent's ratings displayed in Table 2 we read that he underestimated the rating (weight) of the whole issue. Instead of scoring the issue importance at the level of 32 rating points (the principal's reference rating), he assigned to it 17 rating points only. This issue-level inaccuracy will be accumulated within the first summand of formula (3), i.e. $\left|u_{\mathrm{concerts}}^{\mathrm{ref}} - u_{\mathrm{concerts}}^i\right| = 32 - 17 = 15$. However, from the viewpoint of option-level accuracy, the best option (5 concerts) was correctly recognized by the agent, and he assigned to it the highest possible rating resulting within all options of this issue. He cannot be penalized for assigning 17 rating points to the option "5 concerts" instead of 32, since the specificity of the SAW algorithm does not allow him to operate with 32 rating points, if in step 1 the pool of 17 rating points was used to indicate the importance of this issue. Accumulating the non-normalized differences $u_{jk}^{\mathrm{ref}} - u_{jk}^i$ would result in counting the previous mistake one more time here. However, the normalized differences $\left|\bar{u}_{jk}^{\text{ref}} - \bar{u}_{jk}^{i}\right|$ would not indicate any problem at the option-level in such a situation. Yet, if the inaccuracies in ratings appear for the remaining options, the normalized differences would allow us to capture the scale of the problem. For instance, the option "6 concerts" assures 78% of rating points (25 out of 32) assigned by the principal to the option of "5 concerts". When we look at the agent's rating we will find that his evaluation of "6 concerts" is inaccurate. He assigned 10 rating points to this option, which amounts to 58% of the value of the best option (17 rating points were assigned to "5 concerts"). This is a difference of $0.78 - 0.58 = 0.22$ percentage points and should be included in the global value of scoring system inaccuracy. Yet, the cardinal inaccuracy index is measured in rating points, thus this percentage-based inaccuracy must be recalculated using the reference value of the rating assigned to this issue, and hence it will be equal to $u_{\text{concerts}}^{\text{ref}} \cdot \left|\bar{u}_{\text{concerts},6}^{\text{ref}} - \bar{u}_{\text{concerts},6}^{i}\right| = 32 \cdot$ $\cdot\, |0.78 - 0.58| = 6.4$ rating points.

One technical issue needs also to be raised while determining the principal's reference scoring system in Inspire experiments. Since the graphical preferential information was provided by means of circles, there is a question of how to measure the circle sizes that would reflect the final rating values of issues and options. These can be determined either by measuring the circles' radiuses, or the circles' areas. The reference scoring systems determined by measuring radiuses and areas are shown in Table 3. There is no objective rationale that we could use to support our choice of radiuses or areas, hence in this paper we will measure the inaccuracies with respect to two references scoring systems and obtain two cardinal inaccuracy indexes for each agent: radius-based cardinal inaccuracy index (CIR), and area-based cardinal inaccuracy index (CIA).

Table 3: The reference scoring systems for Fado and Mosico determined by measuring radiuses and areas

Party	Reference rates															
	No. of concerts				**No. of songs**					**Royalties for CDs**				**Contract bonus**		
	5	6	7	8	11	12	13	14	15	1.5	2.0	2.5	3.0	125	150	200
Radius-based reference ratings																
Mosico	0	21	26	32	0	7	16	28	21	13	23	16	0	17	10	0
Fado	32	25	21	0	0	8	20	32	24	0	7	12	16	0	15	20
Area-based reference ratings																
Mosico	0	22	30	39	0	5	15	30	20	10	20	13	0	11	6	0
Fado	38	27	22	0	0	6	20	38	26	0	4	7	9	0	10	15

2.4 Thomas-Kilmann Conflict Mode Instrument

The Thomas-Kilmann Conflict Mode Instrument (Kilmann and Thomas, 1977) is a questionnaire-based psychometric test that has been widely used for analyzing the conflict attitudes of people in various contexts and problems (Rahim, 1983; Hignite et al., 2002). It consists of 30 questions regarding the surveyed person's attitude toward conflict and conflict solving. Each question consists of two statements, each describing the examples of different behavior in conflict and related to one of the five conflict modes: competing, collaborating, compromising, avoiding, and accommodating. All five modes are positioned in two-dimensional space described by the intensity of two personal characteristics that play a key role in conflict: assertiveness and cooperativeness, as shown in Figure 2.

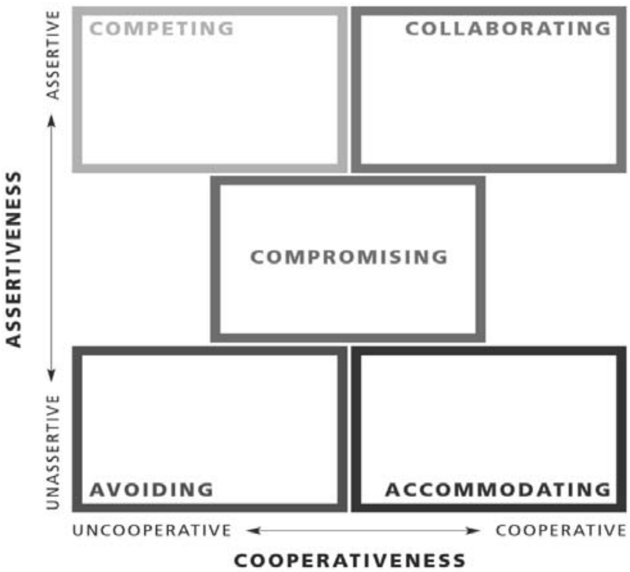


Figure 2. Thomas-Kilmann conflict modes

Source: (www 1).

The competing mode represents a high concern for self, low concern for others; collaborating – high concern for self and others, compromising – moderate concern for self and for others; accommodating – low concern for self and high concern for others, and avoiding – low concern for self and low concern for others.

Within each of 30 questions the responder chooses one of these two statements that describes their behavior better. The intensity of each mode for the responder is determined in the form of raw scores, as a total number of sentences corresponding to this mode chosen by the responder in the TKI test. Since each of these modes can be evaluated on a 0-12 scale (there are in total 12 sentences in the TKI test corresponding to each mode), the results may be represented as the percentage rates of the maximal possible scores. For instance, if the negotiator's raw scores are the following: competing − 6, collaborating − 11, compromising – 4, avoiding − 4, and accommodating – 5, we find collaborating as a leading mode, with very high intensity at the level of 92%. We would also call this person to be little compromising and avoiding at the level of 33%. The TKI results are, however, also interpreted in a relative way by comparing the responder's answers to the typical results obtained by other responders of similar profession or background that form a norm sample (Thomas et al., 2008). Such an interpretation is shown in Figure 3, where the same raw scores were compared to the reference group of 8,000 people preselected to ensure representative numbers of people by organizational level and race/ethnicity.

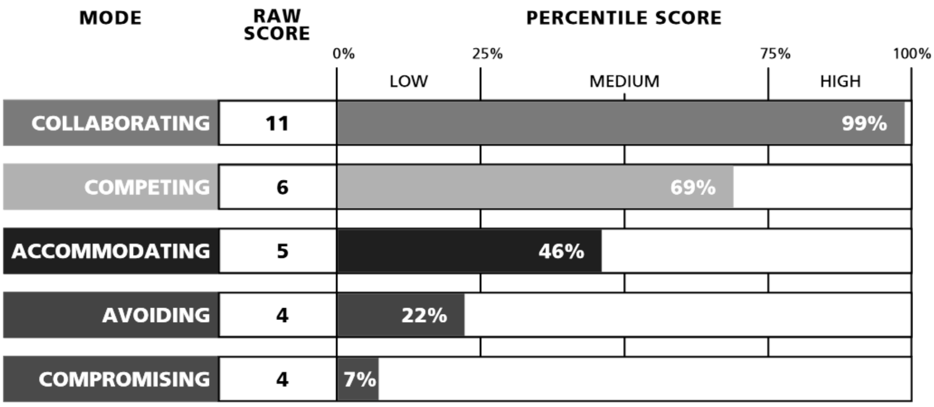


Figure 3. TKI percentile score for a preselected norm sample (reference group)

Source: (www 1).

In our experiment, however, the negotiation profile of each participant will be described by means of a non-relative numerical description of conflict modes represented in the form of raw scores, since − to the best of our knowledge − there is no experimental research that defines the reference percentiles for the intensities of the conflict modes for the international bachelor and master students.

3 Results

3.1 General findings on the inaccuracy of scoring systems depending on the agent's role

The results of the analysis of the scale of the inaccuracy in defining the scoring systems by the agents playing different roles confirm our earlier findings from the pilot studies (Roszkowska and Wachowicz, 2014; Kersten et al., 2015; Roszkowska and Wachowicz, 2015). Out of 176 representatives of WorldMusic (Mosico agents) only 31 (18%) were able to build the scoring systems that were fully concordant with the principal's (WorldMusic) structure of preferences, i.e. for which $OA = 1$. The percentage of fully accurate Fado agents (representatives of Ms. Sonata) was a little higher and equal to 22% (38 out of 174). Yet, the fraction test does not allow to reject the hypothesis on equal proportions of fully accurate Mosico and Fado agents ($p = 0.243$). The structures of ordinal accuracy indexes for representatives of both negotiation parties are shown in Figure 4.

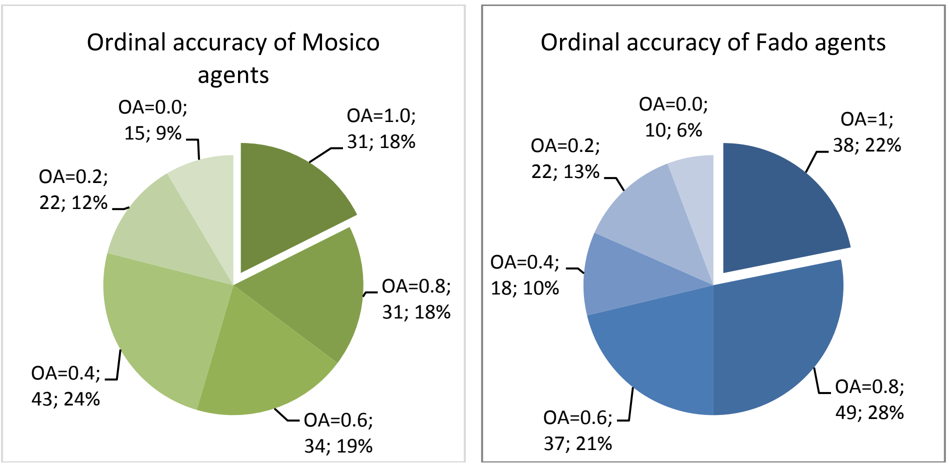


Figure 4. Structures of ordinal accuracy for Mosico and Fado agents

Despite the fact that the percentages of Mosico and Fado agents with $OA = 1$ do not differ significantly, the charts suggest the global difference in structure accuracy between the negotiation roles. Indeed, the chi-squared test determined for the contingency table representing the frequencies illustrated in Figure 4 allows to reject the hypothesis on equal structure of accuracy indexes for Mosico and Fado agents at $p = 0.007$. Hence, analyzing the charts we may conclude that Fado representatives were in general more accurate than Mosico agents.

The next issue we investigated while analyzing the general inaccuracy of agents' negotiation offer scoring systems, were the potential links between the ordinal accuracy (OA) and cardinal inaccuracy indexes (CIR and CIA). In other words, we aimed at discovering whether there is any relationship between the number of mistakes made at the ordinal level and the cardinal scale of these mistakes. To measure this relationship, the Pearson correlation coefficient was used (see Table 4).

Table 4: Relationship among OA, CIR and CIA indexes for Mosico and Fado agents

Pearson correlation coefficients	**Mosico**			**Fado**		
	CIR	CIA	OA	CIR	CIA	OA
CIR	1	.959**	-.708**	1	.924**	-.643**
CIA	.959**	1	-.602**	.924**	1	-.602**
OA	-.708**	-.602**	1	-.643**	-.602**	1

** Correlation significant at 0.01 (two-tailed).

The results shown in Table 4 confirm the existence of a very strong relationship between both cardinal inaccuracy indexes. This suggests that we could use only one of these indexes in our further analyses to make it more clear and the results easier to interpret. A strong negative relationship is also observed between OA and each of the cardinal inaccuracy indexes. The results are statistically significant and similar for both negotiation parties.

Knowing the strong relationship between CIR, CIA and OA and deriving from previous findings on different structures of ordinal accuracy for Mosico and Fado agents we may analyze the differences in average values of accuracy indexes for both roles at ordinal and cardinal levels. The results are shown in Table 5.

Table 5: Average accuracy and inaccuracy indexes for Mosico and Fado agents

Agents	**Indexes**		
	CIR	CIA	OA
Mosico	92.4	83.1	0.556
Fado	64.6	74.5	0.638
Significance (p)	0.000	0.089	0.012

The indexes' values confirm the previous findings on different structures of ordinal accuracy for the negotiation parties. The inaccuracy indexes are larger, while the ordinal accuracy index is smaller, for Mosico than for Fado agents. All the differences are significant at the level no worse than 0.089. Thus the question arises, what could influence such a difference in scoring systems accuracy between the negotiation roles in our case. To answer this question we analyzed first the detailed structures of accuracy in the representation of the principal's rank order for issues' weights and options shown in Table 6.

Table 6: Numbers of Mosico and Fado agents with accurate ratings for the subsequent elements of negotiation template

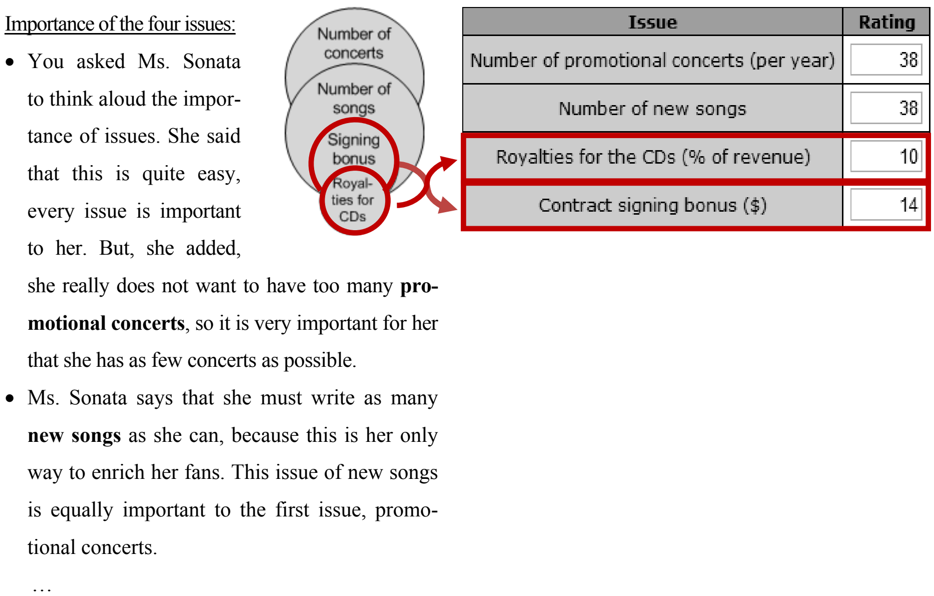
Agent	**Number (%) of agents with ratings concordant with the principal's rank order for:**				
	(1) options of no. of concerts	(2) options of no. of songs	(3) options of royalties	(4) options of contract signing bonus	(5) issue weights
Mosico	133 (76%)	51 (29%)	65 (37%)	120 (68%)	120 (68%)
Fado	135 (77%)	98 (56%)	125 (72%)	137 (79%)	60 (34%)

Cells in grey correspond to the elements of the negotiation template, for which the principal's preferences were non-monotonous.

As can be seen from Table 6, the frequencies of accurate ratings are not homogenous across all elements of the negotiation table and negotiation roles. Some elements of the negotiation template seemed to make bigger problems for agents with assigning concordant ratings than others. For three elements of the negotiation template the principals had defined non-monotonous preferences, i.e.: for options of "No. of concerts" (both principals), and for options of "Royalties" (WorldMusic) − marked in grey in Table 6. For these three elements the percentage of agents with correct ratings is lower (29-56%) than for all the remaining elements of template (68-79%), but one (Fado issue weights). These differences are statistically significant at $p = 0.000$, which was confirmed by the Cochran n-fraction test for dependent samples ($Q = 162.93$ for Mosico agents with 5 samples: elements 1–5; $Q = 42.49$ for Fado agents with 4 samples: elements 1-4). Simultaneously, the Cochran tests determined for both agents separately, and for elements described by monotonous preferences only (for Mosico – elements: 1, 4 and 5; for Fado – elements: 1, 3 and 4) did not allow to reject the H_0 hypothesis on equal fractions ($p = 0.115$ and $p = 0.084$ respectively). This allows us to confirm the following hypothesis: *The accuracy in defining the scoring system for a selected element of the negotiation template depends on the structure of preferences defined by the principal for this element, and is higher when the principal's preferences are monotonous, and lower for non-monotonous ones.*

Yet, there is still a difference in rating accuracy between Fado and Mosico agents for the same element of the negotiation template, i.e. options of number of songs. The structures of the principal's preferences defined for this element by both agents are the same, but the agents' accuracies differ significantly. Thus, it seems there must be also other factors that influence this accuracy, related to the general psychological or demographical profile of the negotiators, that still need to be discovered and studied.

There is another element in the negotiation template for which the fraction of accurate ratings is also very low, i.e. the series of issue weights defined by Fado agents. Only 60 out of 174 (35%) of agents were able to assign the ratings that correctly represented Ms. Sonata's preferences. This is the lowest fraction for Fado agents and the McNemar test confirms the significance of differences in fractions of correct ratings between issue weights and other elements of the negotiation template ($p = 0.000$). However, it is not an issue of non-monotonous preferences that could play a role here. For this element of the template an order of issues described in preferential information by the principal was different than an order used in step 1 of the SAW-based preference elicitation procedure. More precisely, in preferential information the issues were described in the following order: "No. of concerts", "No. of songs", "Contract signing bonus", "Royalties" (as shown in Figure 1), while in the preference elicitation process the two least important issues were reversed in order, i.e. "Royalties" came before "Contract signing bonus" in the list (see Figure 5).



Preferencial information SAW-based rating procedure

Figure 5. Differences in issue lists – preferential information vs. SAW-based rating procedure

It appears that such a technical change in listing the template elements subjected to evaluation had a stronger impact on rating accuracy than non-monotonous preferences. Here, however, it is not a specificity of preferences, but rather the agents' oversight that made the potential difficulties in accurate mapping of the principal's preferences. This allows us to formulate a general hypothesis for future research, the confirmation of which requires a deeper and more detailed analysis, that *there could be some user-specific heuristics and perception errors that affect agents' accuracy in defining their individual negotiation offer scoring systems*.

Summarizing the above results, we find that the problem of determining an accurate scoring system by agents is quite common, and hence decided to find whether the conflict/negotiation profile can differentiate among the agents with respect to the scale of scoring system inaccuracy.

3.2 Comparing profiles of Fado and Mosico agents

In the next stage our analysis was focused on the comparison of the differences in profiles of our negotiators depending on the role they were playing (Mosico or Fado). Because of the size of our research sample we could not consider different profiles taking into account the full range of potential results for each conflict mode (a 0-12 scale). We grouped the results into three classes depending on the intensity of each conflict mode: (1) low – raw scores from 0 to 4; (2) medium – raw scores from 5 to 8; and (3) high – raw scores from 9 to 12. Figures 6 and 7 show the structures of conflict modes for Mosico and Fado agents, respectively.

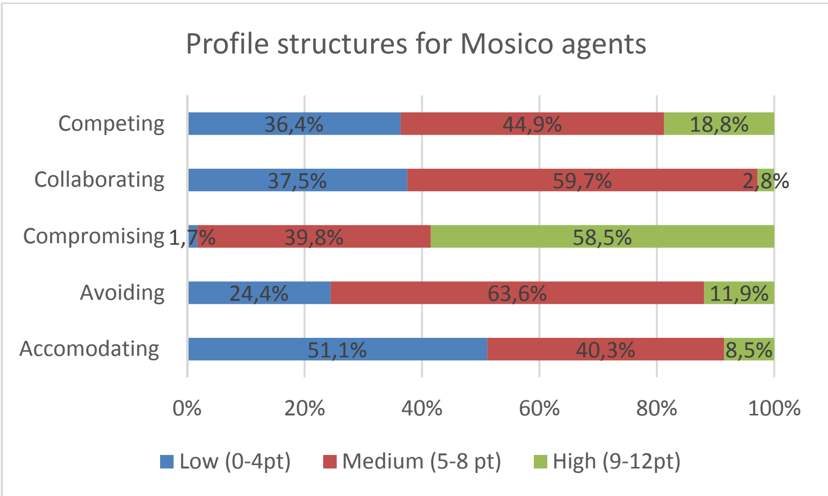


Figure 6. The levels of conflict modes intensities for Mosico agents

Comparing the structures of profiles of Mosico and Fado agents we find that for both parties the compromising mode was the most intensive. For more than 97% of agents this mode was medium or high. The fraction test confirms that the structure of compromising mode is similar for both agents ($\chi^2 = 3.48$, for two degrees of freedom, $p = 0.175$) and significantly different from the structures of other conflict modes.

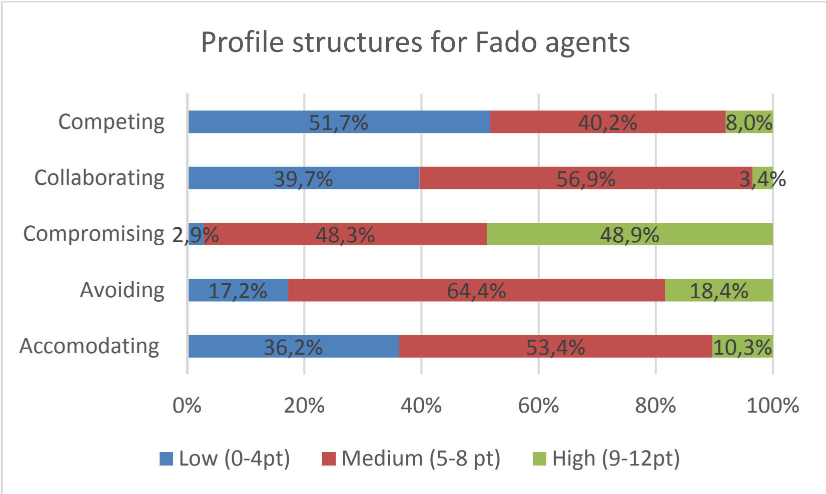


Figure 7. The levels of conflict modes intensities for Fado agents

These observations are confirmed by the analysis of the profiles conducted at the level of average profile values for both agents. The Kolmogorov-Smirnov test confirms that the distribution of intensity of each conflict mode for each agent separately is normal ($p < 0.005$), hence the parametric t-Student test can be applied to measure the significance in differences of average mode values within the profiles. The average profiles are shown in Figure 8.

As can be seen from Figure 8, Fado agents are lower in competing and compromising modes, and higher in avoiding and accommodating modes than Mosico agents. The differences for all these four modes are significant for $p < 0.038$. There is only one mode in both profiles that refers to the collaborating behavior, which does not differentiate significantly between the agents. The t-Student test does not allow to reject the hypothesis on equal average scores for the collaborating mode for Mosico and Fado agents at the level $p = 0.388$.

It is worth noting that competing and compromising modes are the ones that indicate the negotiator's medium and high assertiveness, while avoiding and accommodating ones are characteristic to unassertive decision makers (see Figure 2). Hence, we decided to perform the confirmatory factor analysis using the raw scores of conflict modes to find whether it is possible to determine a single factor describing the assertiveness level across the whole sample. As the result we obtained the following table of loadings that allow us to interpret the factor as the "unassertiveness level" (see Table 7).

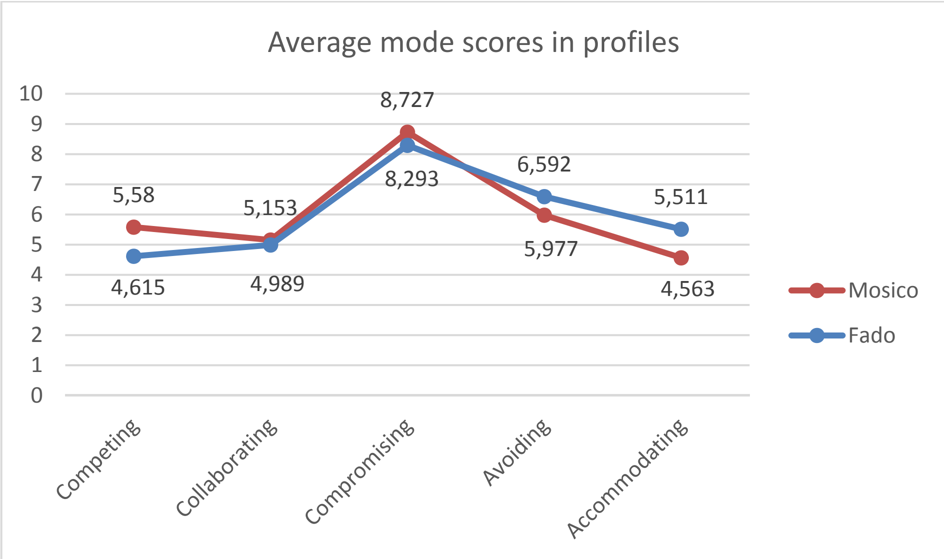


Figure 8. Average negotiation profiles for Mosico and Fado agents

Table 7: Modes' loadings to a single factor

Modes	**Factor 1** "unassertiveness"
Competing	-.837
Collaborating	-.470
Compromising	–
Avoiding	.573
Accommodating	.701

In Table 7 only the loadings greater than 0.3 are shown. They confirm our interpretation of the average profiles from

Figure 8 and the potential impact of selected modes on the intensity of unassertive behavior of agents. In our experiment high competing and collaborating scores load negatively to the unassertiveness level (yet, the collaborating mode loads distinctly less than competing), while high avoiding and accommodating ones load in a positive way.

Using the regression approach we determined the factor values for all agents in our experiment and then calculated the average unassertiveness level for Mosico and Fado agents, which are equal to -0.201 and 0.203, respectively. They seem to be good aggregates of the profiles presented in Figure 8, since Mosico agents were higher in competing and compromising, and lower in avoiding and accommodating, which we interpreted as more assertive behavior. For Fado agents the relation between modes and the unassertive factor is converse. The difference between average unassertiveness levels for Mosico and Fado agents is significant for $p = 0.000$.

On the basis of the average profile analysis conducted in this section, which leads us to the conclusion that the agents differ significantly in the structure of negotiation profiles, and general ordinal accuracy of scoring systems analyzed in section 3.1, that confirmed differences in structures of OA indexes for the agents, we may expect that the negotiation profiles (or at least some of their elements or aggregates, such as the level of unassertiveness) influence the general ordinal accuracy of scoring systems determined by the agents.

3.3 Conflict modes, profiles and scoring systems accuracy

In order to find the direct links between the intensities of conflict modes and the accuracy of scoring systems we determined the Pearson correlation coefficient among the raw scores of modes, unassertiveness level, OA, CIR and CIA separately for each negotiation role. The results are shown in Table 8.

Table 8: Pearson coefficients for accuracy indexes and conflict modes

Conflict mode	Mosico			Fado		
	CIR	CIA	OA	CIR	CIA	OA
Competing	.014	-.009	-.022	.123	.121	-.007
Collaborating	.023	.018	-.053	-.085	-.081	.045
Compromising	-.209**	-.193*	.134	-.105	-.127	-.039
Avoiding	.038	.044	-.011	-.094	-.093	.078
Accommodating	.099	.132	-.030	.078	.094	-.058
Factor 1: unassertiveness	.018	.039	.022	-.047	-.042	-.011

* Correlation significant at 0.05 (two-tailed).
** Correlation significant at 0.01 (two-tailed).

The results reveal no correlation or very weak correlation between selected conflict modes of agents and the selected accuracy measure of the scoring systems they built. The highest relationship was observed for the raw scores of the compromising mode and cardinal inaccuracy indexes only for Mosico agents. This relationship is weak, yet statistically significant at $p = 0.01$ for CIR and $p = 0.05$ for CIA. More precisely, we could conclude that being more likely to compromise results in higher inaccuracy in the scoring system determined by Mosico agents. Unfortunately, such a relationship is not confirmed for Fado agents. Consequently, we are not able to build any convincing regression model that would explain the relationship between the series of conflict modes and the final concordance of the agent's scoring system with the preferential system declared by the principal. Hence, we may hypothesize that there may be some other demographical or sociological characteristics that may affect the negotiators' ability to determine an accurate scoring system. Unfortunately, the experimental data we gathered do not allow us to perform such an in-depth analysis of the demographical profiles of the experiment participants.

Being unable to find the direct links between raw scores of a mode and accuracy levels we performed the cluster analysis for the whole negotiation profiles described by means of a series of all conflict modes in a profile for our further analysis. We applied the k-means clustering procedure, for which an optimal number of clusters was determined as a result, with an authorial approach for measuring the classification quality. We aimed at determining the smallest possible number of homogenous clusters with respect to the negotiation profiles of our agents. Hence, the clustering evaluation procedure required analyzing various classifications obtained for consecutive numbers of classes (starting from the smallest possible, i.e. two) and determining for each classification results the Kruskal-Wallis test to measure, if the distribution of mode raw scores within each conflict mode is significantly different within each of the clusters (for $p < 0.05$). The first classification that meets such requirements was "optimal". For both agents the optimal number of clusters was five. The clustering results for Mosico and Fado agents are shown in Tables 9 and 10, respectively.

Table 9: Average profiles and accuracy indexes for Mosico clusters

Cluster	Average profile					OA	CIR	CIA
	Competing	Collaborating	Compromising	Avoiding	Accommodating			
Cl1 (N = 40)	8.300	4.400	7.700	7.000	2.600	0.595	80.884	72.290
Cl2 (N = 40	5.500	3.900	8.800	5.100	6.700	0.550	102.521	92.873
Cl3 (N = 44)	2.114	5.045	8.886	7.818	6.136	0.477	104.008	95.588
Cl4 (N = 19)	9.211	7.684	7.526	2.474	3.105	0.474	97.701	83.103
Cl5 (N = 33)	4.909	6.273	10.364	5.364	3.091	0.667	75.376	67.687
K-W sign.	.000	.000	.000	.000	.000	.056	.032	.087
All Mosicos	5.58	5.15	8.73	5.98	4.56	.556	92.365	83.097

K-W sing. – p value for the Kruskal-Wallis test.

Table 10: Average profiles and accuracy indexes for Fado clusters

Cluster	Average profile					OA	CIR	CIA
	Competing	Collaborating	Compromising	Avoiding	Accommodating			
Cl1 (N = 11)	11.091	5.364	5.909	4.545	3.091	0.618	74.394	82.101
Cl2 (N = 42	4.286	4.643	9.119	8.333	3.619	0.629	55.634	63.603
Cl3 (N = 53)	1.660	5.000	8.094	7.189	8.057	0.619	65.415	76.909
Cl4 (N = 43)	5.302	5.698	9.395	4.465	5.140	0.651	67.576	77.583
Cl5 (N = 25)	7.400	4.160	6.480	6.960	5.000	0.680	68.595	79.054
K-W sign	.000	.003	.000	.000	.000	.876	.954	.858
All Fados	4.61	4.99	8.29	6.59	5.51	.638	64.613	74.500

K-W sing. – p value for the Kruskal-Wallis test.

From section 3.2 we know that for the average Mosico and Fado profiles a higher accuracy of scoring systems was observed for Fado agents (see last rows in Tables 9 and 10). Now we try to find whether the accuracy varies for different clusters of profiles within each group of agents. The Kruskal-Wallis test confirms the differences in accuracy levels for Mosico agents to be significantly different across all clusters for $p < 0.087$. Yet, the differences in accuracy do not differ significantly for Fado agents ($p > 0.858$). What is more, it is not easy to identify the cluster of negotiators with highest scoring system accuracy within each group of agents. For Mosico Cl5 seems to have the highest OA index and the lowest CIA and CIR values. Yet, the Mann-Whitney test does not confirm these values to differ significantly from the ones determined for Cl1 ($p < 0.288$), that seems to be the second best. The situation is even worse in the case of Fado agents. Here there is no single cluster that outperforms others with respect to OA, CIA and CIR simultaneously. Since the differences in OA values are really insignificant, we may try to identify the cluster with best accuracy using CIA and CIR values only. Using this rationale Cl2 will be considered as best. Yet, the situation is similar to the one we encounter in the case of Mosico agents, when we compare Cl2 of Fado agents with the second best accurate cluster, i.e. Cl3. The Mann-Whitney test will not allow to consider these two clusters as significantly different with respect to CIA and CIR values. Hence, our comparison of the most accurate clusters of Mosico and Fado agents presented below is only illustrative.

Comparing the average profiles of Mosico Cl5 and Fado Cl2 we find that no general conclusion may be drawn regarding a single profile that is most likely to generate the most accurate scoring systems, yet some regularities may be indicated for further investigation (see Figure 9).

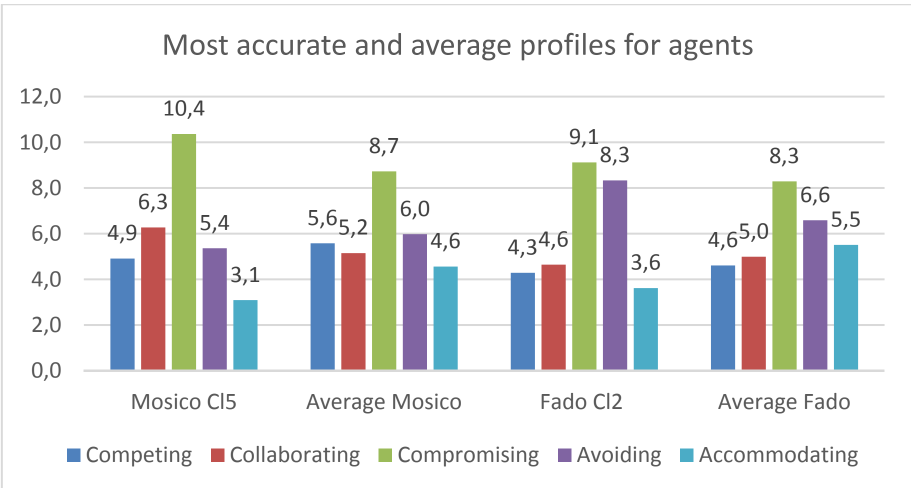


Figure 9. Comparison of most accurate (Mosico Cl5, Fado Cl2) and average profiles

The most accurate profiles of both agents are characterized by the competing level slightly lower than the average within each role and second lowest in each group. No univocal conclusions may be drawn with regard to the collaborating mode. The most accurate profile of Mosico agents is characterized by second highest raw rate of collaboration across all Mosico agents (higher than average). The situation is opposite for Fado agents. The profile of the most accurate Fados is characterized by second lowest collaborating mode, lower than average within the Fado group. Both accurate Mosico and Fado profiles have a very high level of compromising, which for Mosico is the highest across all profiles, and for Fado, second highest. The most accurate Mosico and Fado profiles differ entirely with regard to the level of the avoiding mode. They are, however, second lowest with respect to the accommodating mode, significantly lower than the average for each of the roles.

3.4 Clustering the agents with respect to accuracy indexes

In the last stage of our analysis we changed the perspective used previously in analyzing the relationship between the negotiation profiles and scoring system accuracy. We decided to conduct a more general analysis using the whole dataset without the distinction between the roles and the accuracy measures introduced. Therefore we decided to build a single inaccuracy measure (SIM) that would combine all three indexes: OA, CIA and CIR. We used exploratory factor analysis with regression-based aggregation of factors to determine the potential number of factors and loadings values with an eigenvalue threshold equal to 1 as a discriminant value for the final factor number and the varimax rotation. This analysis proved that the factor model is best fitted for only one factor and allows to explain 84% of the variance measured by three variables considered in the analysis. The loading values of OA, CIA and CIR calculated by means of the principal component method are equal to: -0.846; 0.953 and 0.952, respectively. Hence, the higher the SIM value the bigger inaccuracy of the scoring system under consideration.

Having determined the SIM values for the scoring systems of all experiment participants we identified three classes of participants that differ significantly with respect to SIM values using two-step cluster analysis and Bayesian Information Criterion. For each cluster we calculated the average profiles and SIM values (see Table 11).

Table 11: Average profiles and accuracy for three clusters of negotiators

Negotiators	**Average mode values in profiles**					**Average assertiveness level**	**Average *SIM* value**
	competing	collaborat-ing	compromis-ing	avoiding	accommo-dating		
Highly accurate (N = 188)	4.99	5.16	8.54	6.35	4.96	.003	-.720
Medium accurate (N = 105)	5.01	4.94	8.79	6.33	4.92	.034	.303
Little accurate (N = 57)	5.61	5.02	7.91	5.98	5.47	-.077	1.816
K-W significance	.440	.537	.050	.511	.255	.888	.000

As can be seen from Table 11, there are again no significant differences among most of conflict modes for the profiles described as highly, medium and little accurate. The only mode for which the difference can be considered as significant at $p = 0.05$ is compromising. Yet, it is difficult to draw unambiguous conclusions out of the average values of this mode. It seems that highly compromising negotiators (average raw score of 7.91) are on average less accurate than others. However, the highest intensity of the compromising mode (8.79) does not lead to the most accurate scoring system. It is a medium level of 8.54 that describes the negotiators of highest accuracy in defining the negotiation offer scoring systems. This confirms in some way the previous findings for individual agents (see Figure 8), where Fado agents, being more accurate than Mosico ones, were less compromising, but still at the average level above 8.00. Similarly, there are no significant differences between the clusters with respect to the general assertiveness levels.

4 Summary and conclusions

In this paper we tried to analyze the scale of inaccuracy in defining the scoring systems by the negotiator and its potential links with their negotiation profile, describing the negotiators' attitude and behavior in conflict situations. In our analyses, we used the dataset of bilateral electronic negotiations conducted in the Inspire system, for which a predefined negotiation problem was defined (the Mosico-Fado case). Within the negotiation problem applied, the agent-principal context was embodied, and the preferences of the principal were clearly described both verbally and graphically. Despite such a detailed preferential information, the students that played the roles of Mosico and Fado agents appeared to be relatively inaccurate in defining their scoring systems. Less than one third of all agents built their scoring systems in complete concordance with the principal's preferences (i.e. with $OA = 1$). We observed, however, that the accuracy differed with respect to the agents' roles. Fado agents (the buyers) were on average more accurate than Mosico agents (the sellers). The difference in accuracy seemed to be linked to the structure of the principal's preferences, i.e. non-monotonous preferences made bigger problems for agents to handle them accurately. The effect of heuristic thinking (fast thinking, not paying attention to differences in issue lists) has also affected the ordinal accuracy of assigning the issue weights. What is interesting, the average profiles of both agents also differed significantly. Fado agents, being more accurate in building their scoring system, were also less assertive than Mosicos, i.e. they had lower levels of competing and compromising modes and higher levels of avoiding and accommodating behaviors.

Unfortunately, the in-depth analyses of both the whole dataset and the agent's subsamples did not lead us to any further binding conclusions. The correlations among accuracy indexes and conflict modes appeared to be very weak; hence, it was impossible to build any regression model that would be able to describe the relationship between the negotiators' profiles and their accuracy at the satisfying level of determination and significance. Even though we succeeded in clustering the agents into classes of significantly different profiles, we were unable to prove that these classes differ significantly with respect to the scoring system accuracy, no matter which notion of accuracy was used. A converse approach that amounted to clustering the agents with respect to a single inaccuracy measure did not lead to a better explanation of the problem. It allowed only to formulate a conclusion on the desired level of compromising mode required to determine the most accurate scoring systems. The negotiators with intermediate level of compromising behavior were also the most likely to build the most accurate scoring systems. This general conclusion was also confirmed partially by correlation analysis, where for the Mosico party the compromising mode was the only one that was significantly (yet, weakly) correlated with the selected accuracy measure.

We need to emphasize that the findings and general results we obtained from the experimental analysis are focused on the enriching of the general knowledge on the use and usefulness of the decision support tools applied in negotiation support and the potential factors that influence their use and usefulness rather than on providing any additional support directly to the negotiation parties (asymmetric negotiation support) in the negotiation process. Usually, the parties do not know each other so well or are unable to investigate the profiles of the counterparts based on public information to be able to determine the detailed negotiation profile of their counterparts and derive from them additional information on their accuracy and the potential misinterpretation of the negotiation moves and concessions made. The information about the negotiation profiles of both parties is confidential and may be accessible only by a third party, such as a negotiation support system or a mediator. These third parties can use it to model the best ways, methods and tools to support the negotiation parties in the best possible way, taking into account the negotiators' cognitive limitations, skills and expert knowledge. The last one is actually a part of our ongoing project, and the research presented in this paper was focused only on selected behavioral issues that can be studied when analyzing the general profile of the parties.

As the initial results confirmed the differences in accuracy depending on the role the participants played in our experiment, this may suggest that there are other characteristics of the negotiators that may have an impact on their accuracy in determining the scoring systems, different from the ones described in the TKI test. There may be some demographical or sociological characteristics or also background issues (such as educational level or field) that may affect the results. Therefore, in our future research we will conduct an exploratory analysis of other potential factors that could be used to describe the negotiator's profiles in a different way. We will investigate the applicability of other tests, such as Rational-Experiential Inventory (Handley et al., 2000) or Scott-Bruce (Scott and Bruce, 1995) tests that allow to measure the decision making profile of the respondents.

Acknowledgements

This research was supported by the grant from Polish National Science Centre (2015/17/B/HS4/00941).

References

Churchman C.W., Ackoff R.L. (1954), *An Approximate Measure of Value*, Journal of the Operations Research Society of America, 2(2), 172-187.

Edwards W., Barron F.H. (1994), *SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement*, Organizational Behavior and Human Decision Processes, 60(3), 306-325.

Handley S.J., Newstead S.E., Wright H. (2000), *Rational and Experiential Thinking: A Study of the REI*, International Perspectives on Individual Differences, 1, 97-113.

Hignite M.A., Margavio T.M., Chin J.M. (2002), *Assessing the Conflict Resolution Profiles of Emerging Information Systems Professionals*, Journal of Information Systems Education, 13(4), 315.

Jarke M., Jelassi M.T., Shakun M.F. (1987), *MEDIATOR: Towards a Negotiation Support System*, European Journal of Operational Research, 31(3), 314-334.

Keeney R.L., Raiffa H. (1976), *Decisions with Multiple Objectives: Preferences and Value Trade-offs*, Wiley, New York.

Kersten G.E., Noronha S.J. (1999), *WWW-based Negotiation Support: Design, Implementation, and Use*, Decision Support Systems, 25(2), 135-154.

Kersten G.E., Roszkowska E., Wachowicz T. (2015), *Do the Negotiators' Profiles Influence an Accuracy in Defining the Negotiation Offer Scoring Systems?* The 15th International Conference on Group Decision and Negotiation Letters, B. Kamiński, G.E. Kersten, P. Szufel, M. Jakubczyk, T. Wachowicz, Warsaw School of Economics Press, 129-138.

Kilmann R.H., Thomas K.W. (1977), *Developing a Forced-choice Measure of Conflict-handling Behavior: The "MODE" instrument*, Educational and Psychological Measurement, 37(2), 309-325.

Mustajoki J., Hamalainen R.P. (2000), *Web-HIPRE: Global Decision Support by Value Tree and AHP Analysis*, INFOR J, 38(3), 208-220.

Von Neumann J., Morgenstern O. (1944), *Theory of Games and Economic Behavior*, Princeton University Press.

Rahim M.A. (1983), *A Measure of Styles of Handling Interpersonal Conflict*, Academy of Management Journal**,** 26(2), 368-376.

Raiffa H., Richardson J., Metcalfe D. (2002), *Negotiation Analysis: The Science and Art of Collaborative Decision Making*, The Balknap Press of Harvard University Press, Cambridge (MA).

Roszkowska E., Wachowicz T. (2014), *Defining Preferences and Reference Points − A Multiple Criteria Decision Making Experiment* [in:] *Group Decision and Negotiation. A Process-Oriented View*, eds. P. Zaraté, G.E. Kersten, J.E. Hernández, Springer, 180, 136-143.

Roszkowska E., Wachowicz T. (2014), *Inconsistencies in Defining Preferences and Their Impact on the Negotiation Outcome of eNS Inspire Users*, InterNeg Research Seminar, Concordia Universrity, Montreal.

Roszkowska E., Wachowicz T. (2015), *Application of Fuzzy TOPSIS to Scoring the Negotiation Offers in Ill-structured Negotiation Problems*, European Journal of Operational Research, 242(5), 920-932.

Roszkowska E., Wachowicz T. (2015), *Inaccuracy in Defining Preferences by the Electronic Negotiation System Users*, Lecture Notes in Business Information Processing, 218, 131-143.

Schoop M., Jertila A., List T. (2003), *Negoisst: A Negotiation Support System for Electronic Business-to-business Negotiations in e-commerce*, Data & Knowledge Engineering, 47(3), 371-401.

Scott S.G., Bruce R.A. (1995), *Decision-making Style: The Development and Assessment of a New Measure*, Educational and Psychological Measurement, 55(5), 818-831.

Spremann K. (1987), *Agent and Principal. Agency Theory, Information, and Incentives*, Springer, 3-37.

Thomas K.W., Fann Thomas G., Schaubhut N. (2008), *Conflict Styles of Men and Women at Six Organization Levels*, International Journal of Conflict Management, 19(2), 148-166.

Vetschera R. (2007), *Preference Structures and Negotiator Behavior in Electronic Negotiations*, Decision Support Systems, 44(1), 135-146.

Wachowicz T. (2008), *NegoCalc: Spreadsheet Based Negotiation Support Tool with Even-Swap Analysis* [in:] *Group Decision and Negotiation 2008: Proceedings – Full Papers*, eds. J. Climaco, G.E. Kersten, J.P. Costa, Coimbra, INESC 323-329.

Wachowicz T. (2010), *Decision Support in Software Supported Negotiations*, Journal of Business Economics and Management, 11(4), 576-597.

Wachowicz T., Kersten G.E., Roszkowska E. (2015), *The Impact of Preference Visualization and Negotiators' Profiles on Scoring System Accuracy*, 27th European Conference on Operational Research EURO 2015, University of Strathclyde, Glasgow.

(www 1) http://www.kilmanndiagnostics.com/sites/default/files/TKI_Sample_Report.pdf.