

RESEARCH PAPER / ARTÍCULO DE INVESTIGACIÓN.

Measurement Scales on Investment Decision-making. Empirical Evidence Based on MiFID Questionnaires. Escalas de Medida en la Toma de Decisiones de Inversión. Evidencia Empírica Basada en Cuestionarios MiFID.

Raquel González-Pozo

Universidad Complutense de Madrid, Madrid, Spain.

Contact email: raqueg18@ucm.es

ABSTRACT

The MiFID European directive regulates the provisions of investment services and improves the protection of retail investors. To know the financial knowledge and experience of investors, this directive uses basically two types of questionnaires: appropriateness and suitability tests in which are employed different measurement scales. Since some sub-fields of Behavioral Finance are focused on the study of perception biases, the main objective of this paper is to analyze some advantages and drawbacks of numerical and verbal rating scales, and to examine their effects on investors' perceptions. To do that, it was carried out an empirical study based on different measurement scales: verbal, numerical rating scale, and verbal scale considering proximities. In this contribution, the concept of ordinal proximity measure is employed to collect the individuals' perceptions about ordinal scales.

RESUMEN

La Directiva Europea MiFID regula la prestación de servicios de inversión e intenta mejorar la protección de los inversores minoristas. Para conocer los conocimientos y la experiencia financiera de los inversores, esta directiva utiliza fundamentalmente dos tipos de cuestionarios: los de conveniencia e idoneidad los cuales se emplean diferentes escalas de medición. Dado que algunos subcampos de Finanzas Conductuales se centran en el estudio de los sesgos de percepción, el objetivo principal de este artículo es analizar algunas ventajas y desventajas de las escalas de calificación numérica y verbal, y examinar sus efectos en las percepciones de los inversores. Para ello, se llevó a cabo un estudio empírico basado en diferentes escalas de medida: verbal, escala de calificación numérica y escala verbal considerando proximidades. En esta contribución se emplea el concepto de medida de proximidad ordinal para recoger las percepciones de los individuos sobre las escalas ordinales.

PAPER HISTORY

Received: 27-06-2023

Accepted: 28-02-2024

PALABRAS CLAVE

Cuestionarios MiFID

Percepciones

Escalas

Test de los de conveniencia e idoneidad

Medidas de proximidad ordinal.

KEYWORDS

MiFID questionnaires

Perceptions

Scales

Suitability and

appropriateness tests

Ordinal

proximity measures

Agradecimientos y/o financiamiento

The financial support of the research project PID2021-122506NB-I00 funded by MICIU/AEI /10.13039/501100011033 and ERDF, EU, is acknowledged.

1. Introduction

The Markets in Financial Instruments Directive (MiFID I) was adopted on 21 April 2004 and came into force on 1 November 2007 to regulate several aspects of financial instruments and markets. Afterwards, in 2018, this directive was replaced by a revised set of rules, known as MiFID II (Aghanya et al., 2020; Casey and Lannoo, 2009; Gomber et al., 2018; Jansen et al., 2016).

This latter regulation is focused on governing the provision of investment services in financial instruments considering three main pillars: transparency, investor protection, and competition. MiFID II is applied to financial companies, banks, investment firms, broker dealers, credit institutions, among others.

The main objective of MiFID II is to make European financial markets more transparent and to strengthen the investor protection. To do that, MiFID II provides that financial services companies conduct some questionnaires to determine the financial knowledge of their customers, as well as their investment objectives. These assessments are based on the responses of two types of questionnaires: suitability and appropriateness tests (Miloş, and Miloş, 2019; Prorokowski, 2015).

MiFID II does not establish specific questionnaires to determine the above aspects. Each financial services company can employ its own tests using different questions and measurement scales: numerical or verbal rating scales (Warr et al., 1979). However, psychological research has shown that verbal rating scales are more appropriate than numerical scales, since, among other things, representing individuals' opinions by exact numerical values in situations of vagueness and imprecision is quite difficult (Budescu et al., 1988; Wallsten et al., 1993; Windschitl and Wells, 1996; Zimmer, 1983).

In fact, most of financial services companies and investment firms use qualitative questionnaires, where each response is assigned to a numerical value. Then, a total score is calculated to determine an investment profile considering these numerical values (Busch, 2017; Jansen et al., 2016).

In daily practice, this methodology is very common in the field of Social Sciences. Nevertheless, a problem derived from the transformation of qualitative information into quantitative is that we can obtain different results depending on how the response categories are coded which can sometimes lead to a distorted interpretation of the results (Franceschini et al., 2004; Gadrich et al., 2015; Stevens, 1946).

On the other hand, another factor that can influence the analysis qualitative data is the way in which individuals perceive qualitative scales, because, sometimes, the semantic value of the words, the cultural context or even the individual's personality can cause that verbal rating scales are perceived as non-uniform

(García-Lapresta and Pérez-Román, 2015), that is, individuals can perceive different psychological proximities between the consecutive response categories of the scales. Thus, for example, the ordered qualitative scale used by Credit Suisse to measure risk tolerance: {"Low", "Moderate", "Medium", "Enhanced", "High"}, it can be considered as non-uniform if investors appreciate that the response category "Moderate" is closer to "Medium" than to "Low", or if the category "Enhanced" is closer to "High" than to "Medium".

In addition, one of the main objectives of Behavioral Finance is to analyze finance from a psychological point of view. Thus, Behavioral Finance studies how perceptions and biases affect the financial behaviors of investors (Kahneman et al., 1991; Mitroi and Oproiu., 2014; Sadi, 2011; Sjöberg and Engelberg, 2009; Tversky and Kahneman, 1981). In this sense, several authors have demonstrated that individuals' perceptions affect the financial decisions (Aspara, 2013; Dew and Xiao, 2011; Stein, 1996; Vlaev et al., 2009.) For that reason, when a financial services company tries to identify the investor profiles of their customers by a questionnaire, it is very important to consider how investors perceive the measurement scales used in suitability and appropriateness tests.

Since the publication of the directive MiFID, a large part of the financial and accounting literature has been focused on its implementation. However, most of these reports do not address the contents and characteristics of these suitability and appropriateness tests. Many of them have evaluated the implementation of the MIFID regulations during the banking crisis of recent years (Miloş, and Miloş, 2019; Prorokowski, 2015; Loonen and Janssen, 2023, among others), while other reports have focused on highlighting its European regulations (Islam and Khan, 2019). For that reason, the objective of this paper is to present the main different between suitability and appropriateness tests and to analyze some advantages and drawbacks of numerical and verbal rating scale on investors' perceptions. Likewise, to evaluate the possible influence of the different measurement scales on individuals' assessments, it was carried out an empirical study in which a group of participants used different measurement scales (verbal, numerical rating scale, and verbal scale considering proximities) to assess a set of financial instruments. The results evidenced the importance of considering investors' perception about the measurement scales.

The rest of the paper is organized as follows. Section 2 collects the main objectives of MiFID II and describes the principal characteristics of appropriateness and suitability tests. Section 3 presents the main measurement scales used in MiFID questionnaires focuses on their advantages and disadvantages. Section 4 includes an empirical case study and provides a discussion of the relevant results. Finally, Section 5 shows some concluding remarks.

2. MiFID questionnaires

To protect retail investors, the Market in Financial Instruments Directive (MiFID) is applied in all EU member countries. The directive was introduced in 2007 and it was updated (MiFID II) in 2018 (Bellofatto et al., 2018; Casey and Lannoo, 2009; Fang et al., 2020).

The MiFID principal objectives are:

- To collect information from customers to assess their knowledge of financial products and investment objectives, thereby determining their investment profile and providing personalized advice.
- To offer suitable products and advice for the investors considering their investment knowledge and experience.
- To provide information on the risk investment before contracting financial products.
- To achieve these objectives, the directive MiFID establishes two types of tests: appropriateness and suitability tests.

Appropriateness tests

These tests assess the investors' knowledge and experience with respect to a specific or particular financial product. They are usually employed when financial services companies provide investment services without advice, for instance, execution orders. The aim of this test is to determine which type of financial instruments and services are appropriate for each customer. Some questions included in these tests are:

- Types of financial services.
- Information about the number of operations and volume carried out in capital markets.
- The customer's level of education and profession.

Suitability tests

These tests assess the investors' knowledge and experience, as well as their financial capacity and investment. Suitability tests are used when firms or financial services companies provide investment advice or portfolio management, and they are required to apply to professional and non-professional investor.

Some questions included in these tests are:

- Financial situation.
- The customer' assets, including liquid assets, investments, and real property.

- The ability to bear losses.
- Investment objectives and horizon.
- The customer's risk preferences.
- Information about the number of operations and volume carried out in capital markets.
- The customer's level of education and profession.

The aim of the above tests is to improve investor protection. However, suitability and appropriateness tests will be carried out considering the investor's categorization, as well as on the service rendered by financial companies. Therefore, not all customers of investment firms fill out both tests.

MiFID II does not establish a specific questionnaire to use in suitability and appropriateness tests, but each financial institution can elaborate its own tests including the questions and the measurement scales that it considers convenient (Fünfgeld and Wang, 2009; Loix et al., 2005). In a way that, the identification of suitable products for their customers, in accordance with the criteria of knowledge and experience, depends heavily on how the questionnaires are structured.

Some of the scales included in the questionnaires are numerical rating scales, while other are verbal rating scales formed by a set of linguistic terms or statements. That is the case of the following banks: HSBC Bank, Credit Suisse or Raiffeisen Bank International, among others.

Verbal or qualitative scales are more appropriate to the numerical ones to express opinions and preferences, that why some financial services companies employ verbal rating scales. For instance, Credit Suisse Investment Banking collects some financial information of its customers by means of different verbal rating scales like the following: {"Low", "Moderate", "Medium", "Enhanced", "High"}. This verbal or ordered qualitative scales is used by the firm to determine the risk tolerance of its investors in suitability and appropriateness tests.

3. Measurement scales: Numerical and verbal rating scales

As we have seen, the use of questionnaires with a set of questions evaluated by means of a numeric or verbal scale is very common in suitability and appropriateness tests. For that reason, this section is devoted to present some advantages and disadvantages of above measurement scales and to propose new techniques and procedures to overcome their drawbacks.

- Numerical rating scales: they are the most popular rating scales. They require that the individuals provide their responses by means of numerical values. They present good psychometric properties, and they allow the implementation of statistical procedures.
- Numerical rating scales are usually used to manage and measure qualitative data, such as pain, opinions, customer satisfaction, feelings, likelihood to recommend, etc. (Coelho and Esteves, 2007; Ferreira-Valente et al., 2011; Hjermstad et al., 2011).
- Likert type-scales are an example of numerical rating scales (Likert, 1932). These scales are used to show the respondents' attitude or predisposition towards a specific issue. Likert scales assign a numerical value to each linguistic term considering the number of responses.
- Verbal rating scales also known as verbal description scales are qualitative or categorical ordinal scales composed of four or five levels of adjectives or word expressions to describe emotions, opinions, or different levels of pain. The number of response categories can vary from four to seven elements, but in general, most of verbal rating scales are formed by five response categories symmetrically distributed. (Chien et al., 2013; Lozano et al., 2008; Menold et al., 2014; Preston and Colman, 2000).
- Verbal rating scales present good psychometric properties, and they are an appropriate instrument to collect some abstract concepts which are very difficult to express through numbers. However, in the context of categorial or qualitative scales, it is very usual the assignment of consecutive numerical values to the response categories of the scales. This procedure is not adequate due to it may suggest that the difference levels of the scale are equal, which is not necessarily true.

On the other hand, although verbal rating scales are easily understandable by respondents, the interpretation of their response categories can be influenced by various factors, such as age, sex, culture, or education (Krishnan and Beena, 2009; Loix et al., 2005; Streiner et al., 2014). In these cases, the verbal rating scales can be considered as non-uniform (García-Lapresta and Pérez-Román, 2015), in the sense that respondent can perceive different psychological proximities between the consecutive response categories of the scales.

To overcome the main problems of verbal rating scales, it was proposed recently a method that manages and analyzes qualitative information in an ordinal way, this is the concept of ordinal proximity measure introduced by García-Lapresta and Pérez-Román (2015). Ordinal proximity measures consider how respondents perceive the proximities between the response categories of the scales by means of a set of ordinal degrees of proximity. It is important to highlight that these degrees of proximity are not numbers, they

are abstract objects that represent different degrees of proximity. Thus, δ_1 and δ_h represent the maximum and minimum degrees of proximity, respectively.

The main advantage of the concept of ordinal proximity measure is that it can be applied in the context of non-uniform ordered qualitative scales. Moreover, it follows an ordinal approach to manage qualitative data and avoids the assignment of arbitrary numerical values to the response categories of the scales.

Ordinal proximity measures have been used to devise a multi-criteria procedure in new product development (García-Lapresta et al., 2021) and they have been applied in political, sociological and culture contexts in which verbal rating scales are employed to collect the individuals' opinions (González del Pozo et al., 2020, González del Pozo and García-Lapresta 2019). In addition, the concept of ordinal proximity measure was implemented in the of tourism management by means of the smiley-face scale used by the website Booking.com (González del Pozo and García-Lapresta, 2021).

As will be shown in the following section, ordinal proximity measures can be illustrated by intuitive and simple figures. Likewise, they can be represented mathematically through so-called proximity matrices where δ_1 represents the maximum proximity and δ_h , the minimum.

4. Case study

To analyze the effect of numerical and verbal scales on investor perceptions, this section presents an empirical case study based on the responses of a set of students and professionals in the field of finance. Likewise, in this study, the perceptions of the participants on the verbal rating scale are also considering through the concept of ordinal measure introduced in above section.

Data

The objective of the empirical study was to examine whether the different measurement scales may influence on the decision-making procedures of retail investors. To analyze these effects, in the case study participated a group of 46 students of the different degrees and master's programmes of the Complutense University of Madrid. The participants were a representative mix of both sex (20 males and 26 females) whose ages range from 18 to 40. The size of the sample was considered sufficient for the case study, because in previous reports and studies related to the implementation of MiFID II, the number of participants in the surveys was between 30 and 41 participants (Loonen and Janssen, 2023; Valiante and Assi, 2011).

The data was collected during the academic year 2021/2022. The participants of the study were final

year students of two undergraduate programmes: Bachelor's Degree in Business Administration and Bachelor's Degree in Finance, Banking and Insurance, and one master programme: Master in Finance.

To examine the effect of the measurement scales used in MiFID questionnaires on investors' responses. To do that, participants were asked about their knowledge about the six financial instruments contained in Table 1. This question was obtained from an appropriateness test of the Renta 4 Banco. This financial institution is a firm of a group of companies focused on investment services and asset management. In addition, Renta 4 Banco is the most popular bank in Spain specialized in wealth management, capital markets and investment services.

Table 1. Financial instruments used in MiFID questionnaires.

Public fixed income
Private fixed income
Stocks listed on regulated markets
Harmonized financial investment funds
Listed investment funds

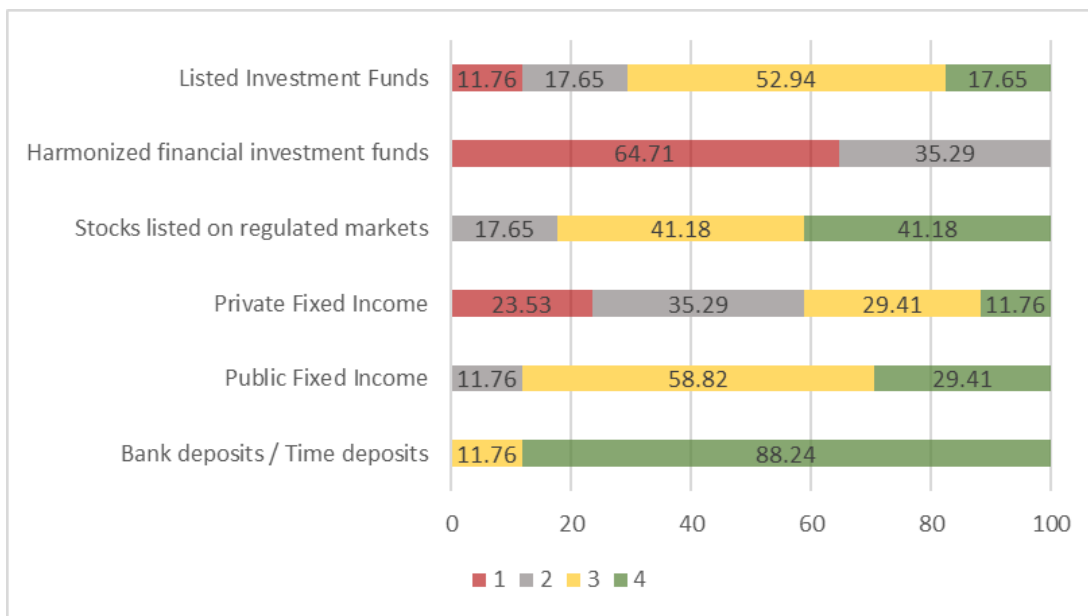
Source: Renta 4 Banco. Appropriateness test.

Procedure and results

The case study was divided into three steps which were conducted in three different moments:

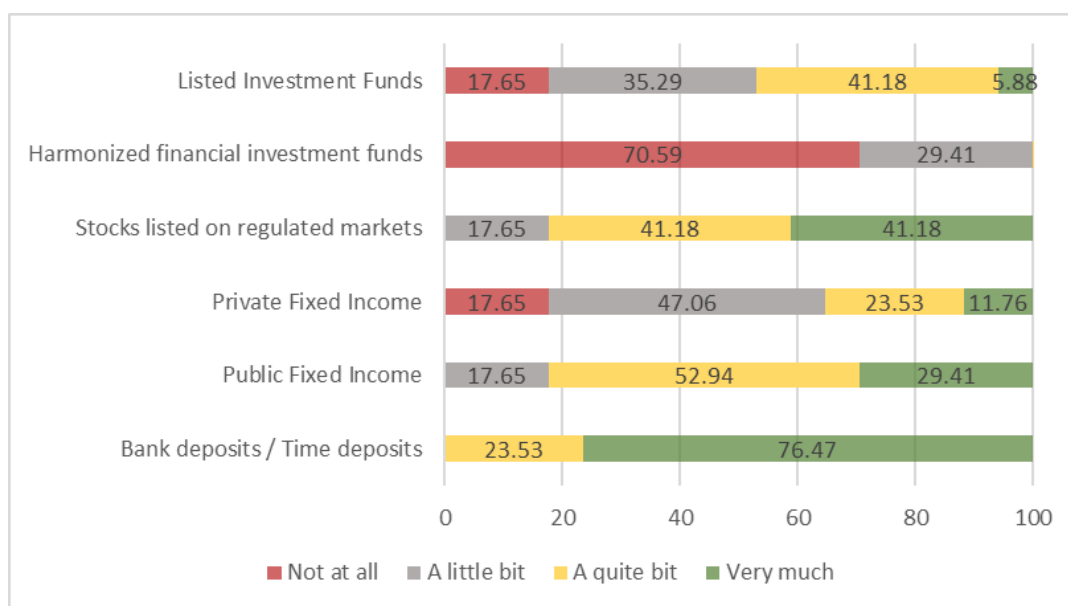
- 1) First, participants used a numerical scale from 1 to 4 to indicate their level of knowledge about the six financial instruments contained in Table 1. The results obtained considering this numerical rating scale are presented in Graph 1.

Graph 1. Participants' knowledge about financial instruments expressed as percentage. Scale 1 - 4.



- 2) In the second step, participants made use of the original verbal scale from the Renta 4 Banco's appropriateness test to show their knowledge. The scale was formed by four response categories: {"Not at all", "A little bit", "Quite a bit", "Very much"}. In Graph 2 are contained the results obtained considering the above verbal rating scale.

Graph 2. Participants' knowledge about financial instruments expressed as percentage. Original verbal rating scale.



3) Finally, to know how participants perceive the verbal rating scale, they complete an on-line survey about the proximities between the linguistic terms of the Renta 4 verbal rating scale. The survey is based on an algorithm that enables to generate the ordinal proximity measure for each respondent (García-Lapresta et al., 2018). The results showed three different ways to perceive the verbal scales (see Figures 1, 2 and 3).

According to the results, more than 82% of the respondents perceived the ordered qualitative scale as non-uniform, while the proportion of participants who considered equal proximities between the consecutive linguistic terms of the scale was only 17.05%.

Figure 1. Ordinal proximity measure associated with the proximity matrix A_{232} . Non-uniform case (1) perceived by 47.05% of participants.

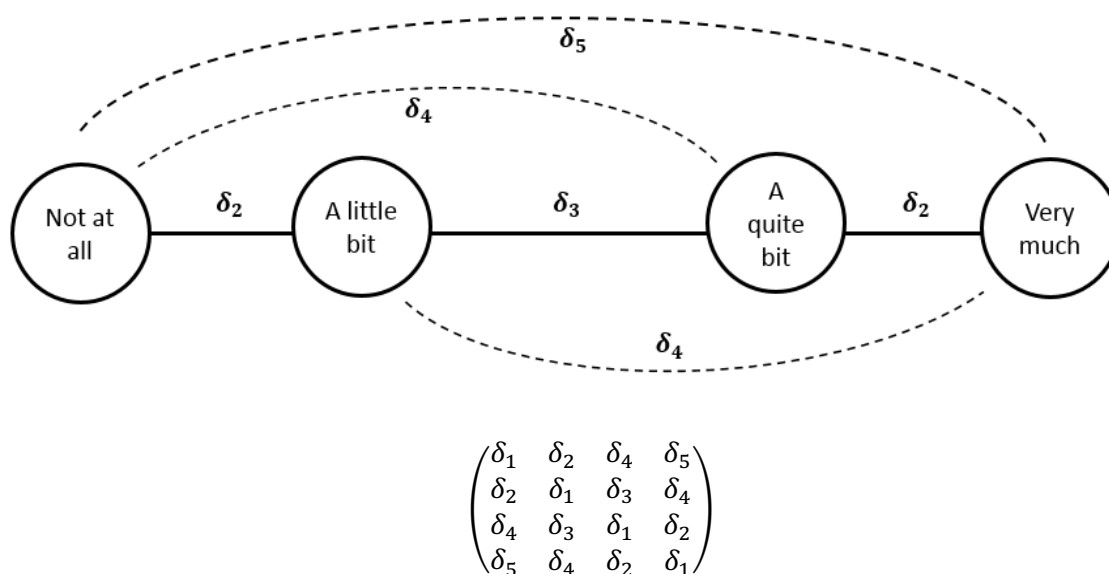


Figure 2. Ordinal proximity measure associated with the proximity matrix A_{332} . Non-uniform case (2) perceived by 35.30% of participants.

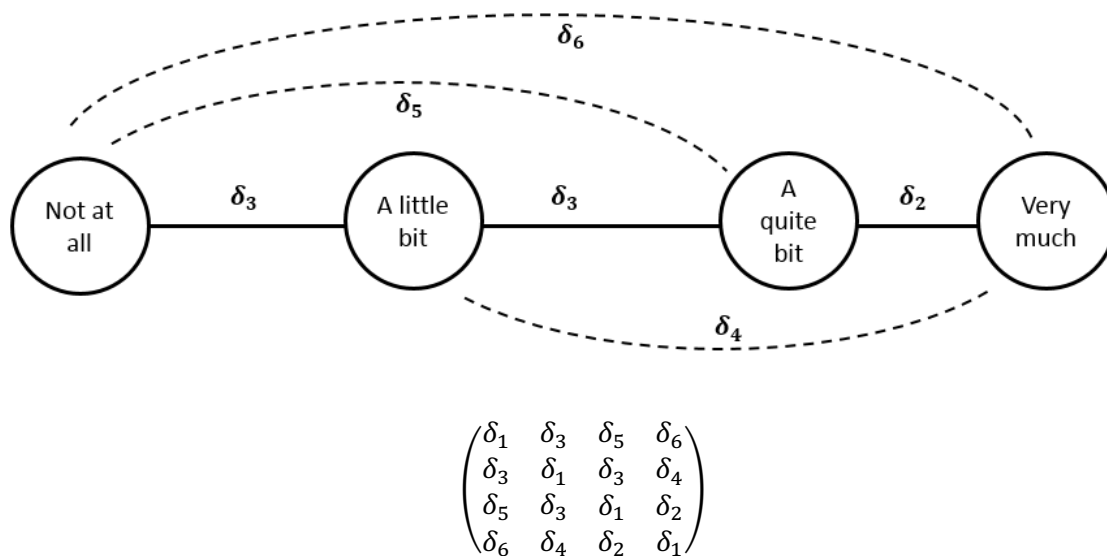
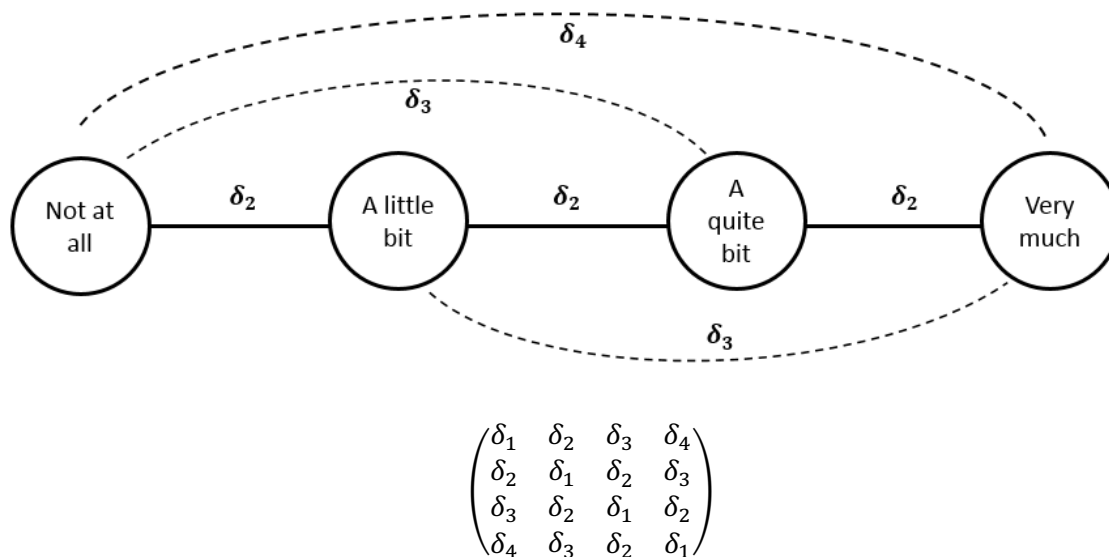
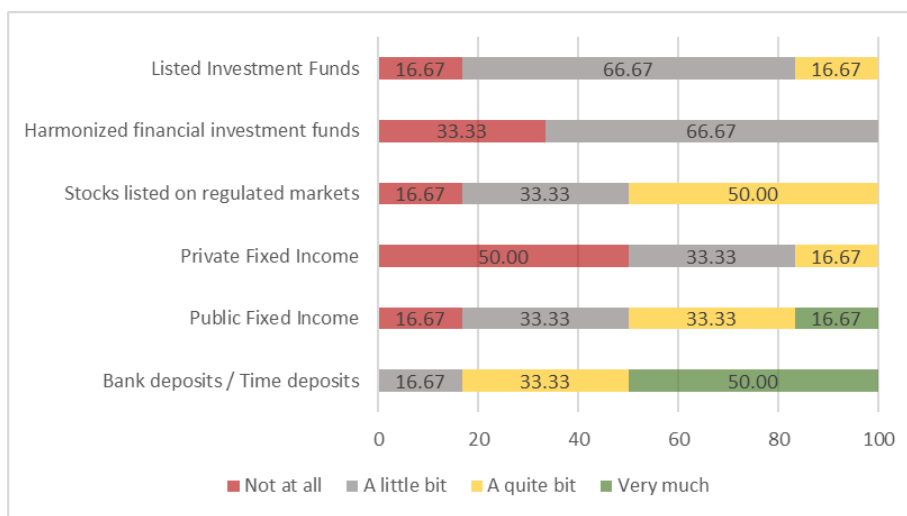


Figure 3. Ordinal proximity measure associated with the proximity matrix A_{222} . Uniform case perceived by 17.65% of participants.

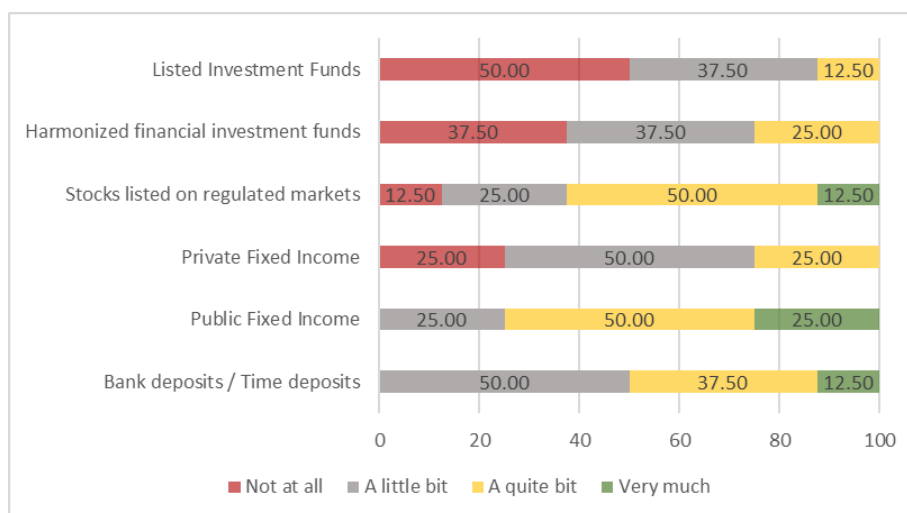


Then, once each participant was aware of how they perceived the proximity between the consecutive terms of the scale, they assessed again their knowledge about the six financial products. In Graphs 3, 4 and 5 are shown the level of knowledge of the participants about the above products considering the three ordinal proximity measures: A_{232} , A_{332} and A_{222} .

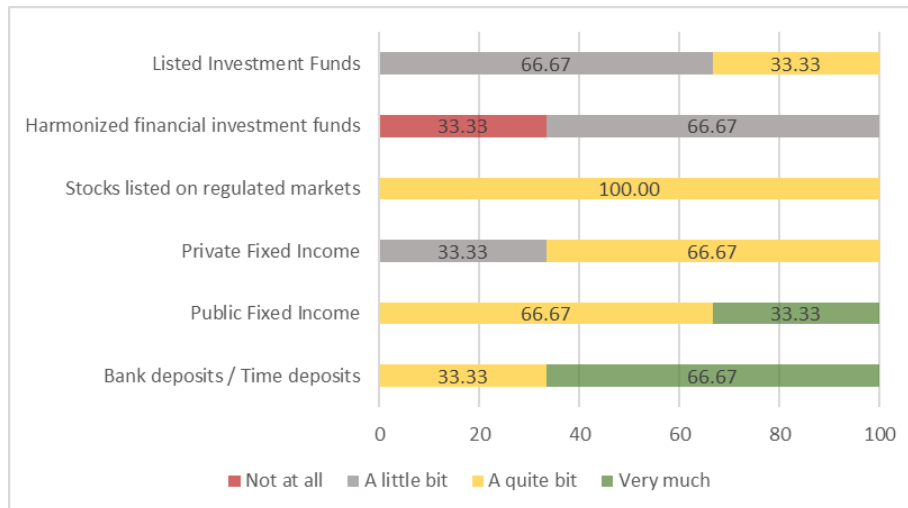
Graph 3. Participants' knowledge about financial instruments expressed as percentage considering the ordinal proximity measure associated with the proximity matrix A_{232} .



Graph 4. Participants' knowledge about financial instruments expressed as percentage considering the ordinal proximity measure associated with the proximity matrix A_{332} .



Graph 5. Participants' knowledge about financial instruments expressed as percentage considering the ordinal proximity measure associated with the proximity matrix A_{222} .



5. Discussion and Conclusions

As we have seen previously, suitability and appropriateness tests use a lot of different measurement scales to determine investors' investment objectives, as well as their financial knowledge and experience.

In order to analyze the effects of numerical and verbal rating scales on investors assessments, in the empirical case study carried out, a group of participants showed their knowledge about six financial instruments by means different types of scales: a numerical rating scale, a verbal scale formed by four linguistic terms and a verbal scale considering different proximities, which were collected in three different ordinal proximity measures represented in Figures 1, 2 and 3.

First, when we compare the results obtained from the numerical and verbal rating scales (see Graphs 1 and 2), the percentages of both scales are quite similar, so there are hardly any differences when the numerical and verbal rating scales. This can be a sign that participants are associated the numerical scale 1- 4 with the four linguistic terms of the scale. However, it is important to mention that the numerical values in categorical scales do not represent a quantitative relationship between response categories, therefore, the outcomes obtained from performing arithmetic operations with these values may be distorted.

On the other hand, it is interesting to comment the results obtained once the concept of ordinal proximity measure has been applied. As seen in the previous section, when the participants responded to the on-line survey about how they appreciate the proximities between the four linguistic terms, three different

ordinal proximity measures were generated. Most of the participants (82.35%) consider that the scale used in the Renta 4's appropriateness tests can be perceived as non-uniform.

Likewise, the fact that the participants are more conscious of their perceptions about the scale, it allows them to give more exact and adjusted assessments. Thus, for example, if we focus on the financial instrument: "Bank deposits/ Time deposits" we see that no participants selected the option "2" (numerical scale) or the option "A little bit" (verbal scale). Nevertheless, in the non-uniform cases associated with the proximity measures associated with the matrices A_{232} and A_{332} , 16.67% and 50% of participants, respectively, selected the response category "A little bit". This fact can also be observed in the assessments of other financial products such as "Public fixed income" and "Stocks listed on regulated markets".

Theoretical contributions

The MiFID is a European directive that regulates the provisions of investment services and improves the protection of retail investors. MiFID employs questionnaires to determine the financial knowledge and experience of investors (Jansen et al., 2016; Miloş and Miloş, 2019).

As we seen, the regulation and the implementation of the MiFID was addressed in several reports. However, there are relatively few studies in the context the suitability and appropriateness tests. In fact, most of the information on this type of tests is hazily available on the websites of banks or financial institutions. For that reason, this paper contributes to this area trying to highlight the main different between suitability and appropriateness tests, as well as the principal questions included in these tests. Likewise, this papers also reviews and analyzes the advantages and disadvantages of the most popular measurement scales, whose effects on the decision-making procedures are not always considered in the fields of economics or finance.

On the other hand, one of the main objectives of Behavioral Finance is to study about how psychological behavior of the investors affects finances. The importance of considering how investors perceive the measurement scales had not been addressed in the economic, financial, or accounting literature. Therefore, this paper also contributes to this issue by means an empirical study in which respondents used three different measurement scales (verbal, numerical rating scale, and verbal scale considering proximities) to assess a set of financial instruments.

Limitations and avenues for future research

While the approach presented in this paper tries to tackle some important issues related to suitability and appropriateness tests the which are not previously considered by other authors, it does present some limitations.

The main objective of the case study was examined the perceptions of the respondents regarding the measurement scales, and other important attributes such as gender, age or culture were not considered. Another limitation of the study can be the origin of respondents. In this case, the participants were students of the different degrees and master's programmes of the Complutense University of Madrid. Nevertheless, the results obtained from this case study have shown significant differences depending on the measurement scale employed, regarding future research, it could be interesting to replicate this idea with a larger sample in which age, gender and culture will be analyzed.

6. References

- Aghanya, D., Agarwal, V., and Poshakwale, S. (2020). Market in Financial Instruments Directive (MiFID), Stock Price Informativeness and Liquidity. *Journal of Banking & Finance*, 113, 105730. doi: 10.1016/j.jbankfin.2019.105730
- Aspara, J. (2013). The Role of Product and Brand Perceptions in Stock Investing: Effects on Investment Considerations, Optimism and Confidence. *Journal of Behavioral Finance*, 14(3), 195-212. doi: 10.1080/15427560.2013.819803
- Bellofatto, A., D'Hondt, C., and De Winne, R. (2018). Subjective Financial Literacy and Retail Investors' Behavior. *Journal of Banking & finance*, 92, 168-181. doi: 10.1016/j.jbankfin.2018.05.004
- Budescu, D. V., Weinberg, S., and Wallsten, T. S. (1988). Decisions Based on Numerically and Verbally Expressed Uncertainties. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 281-294. doi: 10.1037/0096-1523.14.2.281
- Busch, D. (2017). MiFID II and MiFIR: Stricter Rules for the EU Financial Markets. *Law and Financial Markets Review*, 11(2-3), 126-142. doi: /10.1080/17521440.2017.1412060
- Casey, J. P., and Lannoo, K. (2009). *The MiFID Revolution*. Cambridge University Press. doi: 10.1017/CBO9780511770470
- Chien, C. W., Bagraith, K. S., Khan, A., Deen, M., and Strong, J. (2013). Comparative Responsiveness of Verbal and Numerical Rating Scales to Measure Pain Intensity in Patients with Chronic Pain. *The Journal of Pain*, 14(12), 1653-1662. doi: /10.1016/j.jpain.2013.08.006
- Coelho, P. S., and Esteves, S. P. (2007). The Choice between a Fivepoint and a Ten-point Scale in the Framework of Customer Satisfaction Measurement. *International Journal of Market Research*, 49(3), 313-339. doi: 10.1177/147078530704900
- Dew, J., and Xiao, J. J. (2011). The Financial Management Behavior Scale: Development and Validation. *Journal of Financial Counseling and Planning*, 22(1), 43-60. doi: 0.1207/s15427579jpfm0604_3

- Fang, B., Hope, O. K., Huang, Z., and Moldovan, R. (2020). The Effects of MiFID II on Sell-side Analysts, Buy-side Analysts, and Firms. *Review of Accounting Studies*, 25(3), 855-902. doi: 10.1007/s11142-020-09545-w
- Ferreira-Valente, M. A., Pais-Ribeiro, J. L., and Jensen, M. P. (2011). Validity of Four Pain Intensity Rating Scales. *Pain*, 152(10), 2399-2404. doi: 10.1016/j.pain.2011.07.005
- Franceschini, F., Galetto, M., and Varetto, M. (2004). Qualitative Ordinal Scales: The Concept of Ordinal Tange. *Quality Engineering*, 16(4), 515-524. doi: 10.1081/QEN-120038013
- Fünfgeld, B., and Wang, M. (2009). Attitudes and Behaviour in Everyday Finance: Evidence from Switzerland. *International Journal of Bank Marketing*, 27(2), 108-128. doi: 10.1108/02652320910935607
- Gadrich, T., Bashkansky, E., and Zitikis, R. (2015). Assessing Variation: A Unifying Approach for all Scales of Measurement. *Quality & Quantity*, 49(3), 1145-1167. doi: 10.1007/s11135-014-0040-9
- García-Lapresta, J. L., and González del Pozo, R. (2019). An Ordinal Multi-Criteria Decision-making Procedure under Imprecise Linguistic Assessments. *European Journal of Operational Research*, 279(1), 159-167. doi: 10.1016/j.ejor.2019.05.015
- García-Lapresta, J. L., González del Pozo, R., and Pérez-Román, D. (2018). Metrizable Ordinal Proximity Measures and Their Aggregation. *Information Sciences*, 448, 149-163. doi: 10.1016/j.ins.2018.03.034
- García-Lapresta, J. L., Moreno-Albadalejo, P., Pérez-Román, D., and Temprano-García, V. (2021). A Multi-criteria Procedure in New Product Development Using Different Qualitative Scales. *Applied Soft Computing*, 106, 107279. doi: 10.1016/j.asoc.2021.107279
- García-Lapresta, J. L., and Pérez-Román, D. (2015). Ordinal Proximity Measures in the Context of Unbalanced Qualitative Scales and Some Applications to Consensus and Clustering. *Applied Soft Computing*, 35, 864-872. doi: 10.1016/j.asoc.2015.02.035
- Gomber, P., Clapham, B., Lausen, J., and Panz, S. (2018). The Impact of MiFID II/MiFIR on European Market Structure: A Survey Among Market Experts. *The Journal of Trading*, 13(2), 35-46. doi: 10.3905/jot.2018.13.2.035
- Gonzalez del Pozo, R., Dias, L. C., and García-Lapresta, J. L. (2020). Using Different Qualitative Scales in a Multi-criteria Decision-making Procedure. *Mathematics*, 8(3), 458. doi: 10.3390/math8030458
- González del Pozo, R., and García-Lapresta, J. L. (2019). Tratamiento Ordinal de las Escalas Cualitativas Utilizadas por el Centro de Investigaciones Sociológicas. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 28, 124-142. doi: 10.46661/revmetodoscuanteconomia.3788
- González del Pozo, R., and García-Lapresta, J. L. (2021). Managing the Smiley Face Scale used by Booking. com in an Ordinal Way. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(2), 128-139. doi: 10.4067/S0718-18762021000200110
- Hjermstad, M. J., Fayers, P. M., Haugen, D. F., Caraceni, A., Hanks, G. W., Loge, J. H., and European Palliative Care Research Collaborative (EPCRC). (2011). Studies Comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for Assessment of Pain Intensity in Adults: A Systematic Literature Review. *Journal of Pain and Symptom Management*, 41(6), 1073-1093. doi: 10.1016/j.jpainsymman.2010.08.016

- Islam, S.T., Khan, M.Y.H. (2019). Evaluating the changes in the European Banking Regulation – MiFID and its possible effects on the Global Economy: A Theoretical Study. *Financial Markets, Institutions and Risks*, 3(4), 24-31. doi: 10.21272/fmir.3(4).24-31.2019
- Jansen, R., Kilian, A., and Loonen, T. (2016). MiFID II: Suitability and Appropriateness Practical guidelines for Investment Services. Technical Report. University Amsterdam. doi: 10.13140/RG.2.2.13193.29288
- Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *Journal of Economic Perspectives*, 5(1), 193-206. doi: 10.1257/jep.5.1.193
- Krishnan, R., and Beena, F. (2009). Measurement of Conformity to Behavior Finance Concepts and Association with Individual Personality. *The IUP Journal of Behavioral Finance*, 6, 25-40.
- Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22, 1-55.
- Loonen, T., and Janssen, R. (2023). Implementation of MiFID II investor protection provisions by private banks within the European Union. *Journal of Financial Regulation and Compliance*, 31(1), 1-15. doi: 10.1108/JFRC-10-2021-0087
- Loix, E., Pepermans, R., Mentens, C., Goedee, M., and Jegers, M. (2005). Orientation toward finances: Development of a Measurement Scale. *The Journal of Behavioral Finance*, 6(4), 192-201. doi: 10.1207/s15427579jpfm0604_3
- Lozano, L. M., García-Cueto, E., and Muñiz, J. (2008). Effect of the Number of Response Categories on the Reliability and Validity of Rating Scales. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 4(2), 73-79. doi: 10.1027/1614-2241.4.2.73
- Menold, N., Kaczmirek, L., Lenzner, T., and Neusar, A. (2014). How do respondents attend to verbal labels in rating scales?. *Field Methods*, 26(1), 21-39. doi: 10.1177/1525822X13508270
- Miloş, M. C., and Miloş, L. R. (2019). Challenges Regarding the Implementation of MiFID II. In *International conference Knowledge-Based Organization*, 25(2), 158-162. doi: 10.2478/kbo-2019-0073
- Mitroi, A., and Oproiu, A. (2014). Behavioral Finance: New Research Trends, Socionomics and Investor Emotions. *Theoretical and Applied Economics*, 18(4), 153-166.
- Preston, C. C., and Colman, A. M. (2000). Optimal Number of Response Categories in Rating Scales: Reliability, Validity, Discriminating Power, and Respondent Preferences. *Acta Psychologica*, 104(1), 1-15. doi: 10.1016/S0001-6918(99)00050-5
- Prorokowski, L. (2015). MiFID II Compliance—are We Ready?. *Journal of Financial Regulation and Compliance*, 23(2), 196-206. doi: /10.1108/JFRC-02-2014-0009
- Sadi, R., Asl, H. G., Rostami, M. R., Gholipour, A., and Gholipour, F. (2011). Behavioral Finance: The explanation of Investors' personality and Perceptual Biases Effects on Financial Decisions. *International Journal of Economics and Finance*, 3(5), 234-241. doi: 10.5539/ijef.v3n5p234
- Sjöberg, L., and Engelberg, E. (2009). Attitudes to Economic Risk Taking, Sensation Seeking and Values of Business Students Specializing in Finance. *The Journal of Behavioral Finance*, 10(1), 33-43. doi: 10.1080/15427560902728712

- Stein, J. C. (1996). Rational Capital Budgeting In An Irrational World. *The Journal of Business*, 69(4), 429-455. doi: 10.3386/w5496
- Stevens, S. (1946). On the Theory of Scales of Measurement. *Science*, 103, 677-680. doi: 10.1126/science.103.2684.677
- Streiner, D. L., Norman, G. R., and Cairney, J. (2014). Scaling Responses. In: *Health Measurement Scales: A Practical Guide to their Development and Use*, 38-73. Oxford University Press, USA. doi: 10.1093/med/9780199685219.003.0004
- Tversky, A., and Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211, 453-458. doi: 10.1126/science.7455683
- Valiante, D. and Assi, B. (2011). MiFID Implementation in the midst of the financial crisis: results of an ECMI survey. ECMI Research Report No. 6.
- Vlaev, I., Chater 1, N., and Stewart, N. (2009). Dimensionality of Risk Perception: Factors Affecting Consumer Understanding and Evaluation of Financial Risk. *The Journal of Behavioral Finance*, 10(3), 158-181. doi: 10.1080/15427560903167720
- Wallsten, T. S., Budescu, D. V., and Zwick, R. (1993). Comparing the Calibration and Coherence of Numerical and Verbal Probability Judgments. *Management Science*, 39(2), 176-190. doi: 10.1287/mnsc.39.2.176
- Warr, P., J. Cook, and T. Wall. (1979). Scales for the Measurement of Some Work Attitudes and Aspects of Psychological Well-Being. *Journal of Occupational Psychology*, 52 (2),129-48. doi: 10.1111/j.2044-8325.1979.tb00448.x
- Windschitl, P. D., and Wells, G. L. (1996). Measuring Psychological Uncertainty: Verbal versus Numeric Methods. *Journal of Experimental Psychology: Applied*, 2(4), 343-364. doi: 10.1037/1076-898X.2.4.343
- Zimmer, A. C. (1983). Verbal vs. Numerical Processing of Subjective Probabilities. In: *Decision Making under Uncertainty*. (R.W. Scholz, ed), 159-182. Elsevier. doi: 10.1016/S0166-4115(08)62198-6