



The effect of readability and language choice in management accounting reports on risk-taking: an experimental study

Ewgenij Besuglov¹ · Nils Crasselt¹

Published online: 3 April 2020
© The Author(s) 2020

Abstract

Concisely written financial documents facilitate information processing and help to improve decision-making. However, the role of readability seems to be still neglected in accounting practice. Readability concerns become even more relevant in the light of the fact that English is commonly used as the corporate language in multinational corporations throughout the world. Management reports prepared in a foreign language may be less readable than reports written in the preparer's mother tongue. With this study we set out to experimentally investigate the effects of report readability and choosing either managers' mother tongue or a foreign language as the reporting language on risk-taking in a management accounting context. We were able to show that a low level of readability significantly reduces the willingness to accept beneficial risks, a phenomenon which can be explained by the cognitive load effect. In line with the foreign language effect, we provide tentative evidence that the use of a foreign language in combination with a high readability level enhances the willingness to undertake beneficial risks. Further investigation revealed that the foreign language effect is present only for a medium–high language proficiency level and can be best explained by the reduced emotionality account.

Keywords Cognitive load effect · Foreign language effect · Management accounting reporting · Readability · Risk-taking

JEL Classification M16 · M40

✉ Ewgenij Besuglov
besuglov@wiwi.uni-wuppertal.de
Nils Crasselt
crasselt@wiwi.uni-wuppertal.de

¹ Schumpeter School of Business and Economics, University of Wuppertal, Wuppertal, Germany

1 Introduction

Readability is an omnipresent concern not only in accounting practice but also in contexts as wide-ranging as the military, healthcare, and the law (Bonsall et al. 2017). The Plain English Rule 421(d), passed by the US-American Securities and Exchange Commission (SEC) in 1998, was initiated to obligate the issuers of financial disclosures to adhere to plain English principles (Rennekamp 2012). Such appreciation by the legislation underlines the importance of readability for financial disclosures. The consensus in the accounting literature is that less readable financial documents cause negative consequences for a decision-maker such as increased processing difficulty (Tan et al. 2014), which can lead to impeded understanding and less willingness to extract relevant information (Bloomfield 2002). Most importantly, even if the understanding is not hampered, there is a chance that lower levels of readability will modulate judgmental processes by making a decision-maker more prone to rely on heuristic cues like sentiment (Tan et al. 2014).

Today, English is the lingua franca in almost all spheres of our globalized world such as politics, culture life, and business (Tietze 2008). For example, some multinational corporations like Nissan or Honda, despite being based in a non-Anglophone country, have already implemented or plan to install English as their official corporate language (Borzykowski 2017). This means that people in these organizations have to carry out their official communication in a foreign language, and, therefore, make their decisions on the basis of a foreign language. A uniform corporate language brings many benefits, but, at the same time, there are some potential problems as well. On the one hand, recipients of information could struggle with information processing in a foreign language due to a lack of language proficiency. On the other hand, employees who prepare the information, e.g., management accountants, in a non-native language are oftentimes challenged to provide highly readable documents. Recent cognitive-psychological studies show that, considered alone, both cognitive load (e.g., through lower levels of readability) and foreign language use can influence risk-taking behavior. The aim of this study is to jointly scrutinize the effect of readability and foreign language use on ‘smart’ risk-taking, i.e., beneficial risks that are worthwhile to undertake, within a management accounting context.

According to the *dual-process model* of cognition, mental processes are operated by two distinct types of systems, so-called Systems 1 and 2 (Frankish 2010; Sloman 1996; Stanovich and West 2000). System 1 is characterized as: automatic, unconscious, heuristic, and affective; System 1 is executed in the human brain by default. In contrast, the following features are attributed to System 2: control, consciousness, analysis, and high demand for cognitive resources; System 2 needs to be activated consciously (for an extensive list of attributes attached to the two systems, see Evans 2008).¹

In line with the dual-process model, one side effect of a low readability level is a higher distraction of System 2 due to intensified working memory activities and a higher level of cognitive load (Seufert et al. 2017). Under cognitive load, people tend to be more risk-averse and this phenomenon is explained by the help of the dual-process

¹ It has been empirically demonstrated that System 1 is systematically prone to certain biases (e.g., affect substitution, framing effect, anchoring effect, etc.) which can be corrected when System 2 is activated, see Kahneman (2011) and Stanovich and West (2008).

model. When cognitive capacities are burdened, there remain no sufficient cognitive resources to control and override decisions passed by the affective and bias-prone System 1 (Deck and Jahedi 2015).

The *foreign language effect* constitutes a positive effect of foreign language use on the promotion of rational decisions. Extant cognitive-psychological studies demonstrated that people who process information in a foreign language are less superstitious (Hadjichristidis et al. 2017a), show higher levels of self-regulation (Klesse et al. 2015), and—most importantly for our study—are less loss- and risk-averse (Hadjichristidis et al. 2017b).

One of the leading explanations for the foreign language effect builds upon a lower level of emotional reactivity when a foreign language is used (Pavlenko 2012). In the case of risk decisions, if emotionality is hampered, and given that risks are associated with rather negative feelings (Loewenstein et al. 2001), there will be more room left for rational arguments of System 2 in favor of risky alternatives implying monetary benefits. Other potential explanations for the foreign language effect are based on the psychological distance mechanism and the disfluency effect; we discuss these explanations in more detail in our theoretical underpinnings in Sect. 2.2.

In the context of foreign language use, it is important to consider the role of language proficiency. On the one hand, there will be a negligible difference when language proficiency approaches native-like levels. On the other hand, lower levels of language proficiency may impose too much cognitive load and, thus, take away the possibility to engage in the deliberative thinking mode of System 2. Following from this, the foreign language effect is most likely to emerge only for a medium–high level of language proficiency (Costa et al. 2017).

Our study was designed to capture the negative effect of cognitive load (through a low level of readability) and a positive effect of foreign language use on ‘smart’ risk-taking in a management accounting context. As we simultaneously expose some subjects of our experiment to added cognitive load through a low readability level, we predict to find the foreign language effect only in the condition with a high readability level. Further, we take a deeper look at the nature of the foreign language effect by scrutinizing the role of language proficiency and by testing the most discussed psychological mechanisms behind it.

Our study was set up as a 2×2 between-subject experiment with the main independent variables: language (foreign vs. native) and readability (high vs. low). The participants of our experiment took the role of a manager within a principal-agent framework with only hypothetical financial consequences for themselves and received a fictitious management accounting report either in their native tongue (German) or a foreign language (English). The report comprised two alternative purchase offers for the firm’s production. One offer was a safe option and the other one represented a risky but beneficial option in terms of expected value of revenues. Readability was manipulated by means of linguistic and formatting choices according to the SEC’s Plain English Handbook (SEC 1998). We asked the subjects to choose between the two purchase offers first and, hereafter, to provide their individual certainty equivalents for the risky option, with the latter representing our main dependent variable.

Our experiment revealed a highly significant and robust effect of readability. Consistent with the predictions, participants in the low readability condition were significantly

less willing to accept ‘smart’ risks than their counterparts in the high readability condition. We were also able to detect a positive effect of foreign language use on ‘smart’ risk-taking in combination with a high readability level. However, this effect was only slightly significant and not very robust. Further analysis showed that participants with a medium–high language proficiency level and within the high readability condition exhibited substantially higher willingness to undertake ‘smart’ risks and were significantly less affected by positive feelings towards the option with the safe revenue relative to other language proficiency groups. Hence, our study offers support in favor of the reduced emotionality account as the explanation for the foreign language effect.

Our findings may be helpful for management accounting practitioners in two ways. First, by showing a significant negative effect of low readability on the willingness to undertake beneficial risks, we draw the attention of management accountants to this phenomenon, which may not be obvious at first glance. Second, we provide some evidence that the use of a foreign language can have beneficial effects on the willingness to accept ‘smart’ risks within a typical management accounting context. However, the necessary prerequisite for this effect are: a medium–high language proficiency level of the decision-maker, and a highly readable management accounting report.

We advance the existing literature as follows. First, we bring together two streams of research by jointly investigating the influence of cognitive load (through readability manipulations) and foreign language use on ‘smart’ risk-taking behavior. Second, our study is—to our best knowledge—the first one in which the leading explanations for the foreign language effect are simultaneously tested and language proficiency is measured by means of a standardized test (C-Test).

The rest of this paper proceeds as follows: Sect. 2 provides the theoretical background. In Sect. 3, we develop our hypotheses. Section 4 sets forth our experimental method, and, in subsequent Sect. 5, we report the results and further empirical findings. Section 6 concludes.

2 Theoretical background and empirical evidence

2.1 Readability

Prior studies in the field of accounting research investigated the influence of different levels of readability² primarily from investors’ perspective. It was demonstrated that disclosure formats can affect the acquisition, evaluation, and weighting of disclosed information (Asay et al. 2017). Following the assumption that lower levels of readability impede investors’ understanding of financial reports and their willingness to extract information from such disclosures (Bloomfield 2002), some accounting studies concentrated on stock market behavior and found that less readable and lengthy 10-K filings are accompanied by noticeable market underreactions (You and Zhang 2009) and reduced trading volumes, especially for small (non-professional) investors (Miller 2010). There is also evidence that disclosure readability has an impact on professionals

² Exact definitions of readability are imprecise in prior literature. However, the general agreement is that readability refers to the ease with which a reader can process and comprehend written texts, see Bonsall IV et al. (2017).

as well. For instance, less readable and longer disclosures are associated with higher exertion of effort (Lehavy et al. 2011), and the forecast dispersion increases when analysts deal with long 10-K filings (Loughran and McDonald 2014).

Even if a less readable financial report does not hamper individual's ability to extract and understand the information, readability may still have an effect on how judgements and decisions are shaped. Rennekamp (2012) posits that more readable disclosures influence investors through processing fluency. She demonstrated that when readability is high and good news are conveyed in the disclosure, investors' valuation judgements are more positive, and the opposite is true when readability is high and bad news are conveyed. Besides, language sentiment in disclosures proved to impact investors' judgements more strongly for lower levels of readability (Tan et al. 2014). In particular, positive language paired with low readability leads to higher earnings judgments, and this effect is especially pronounced for less sophisticated investors. Tan et al. (2014) explain their findings with the help of the dual-process model. They argue that less readable reports increase information processing difficulty and, as a result, especially less sophisticated investors are more likely to rely on heuristic cues such as language sentiment.

The effect of readability manipulations has been intensively investigated in the field of educational psychology. Studies in this area of research examined the influence of different readability levels of learning materials on learning outcomes. On the one hand, it was demonstrated that less legible texts can serve as 'desirable difficulty' in improving long-term learning and retention performance through the so-called *disfluency effect* (Seufert et al. 2017). Increasing perceived difficulty (e.g., via a less readable text) associated with a cognitive task activates the more analytic-elaborative thinking mode of System 2 (Alter et al. 2007). On the other hand, from a certain level of illegibility onwards *extraneous cognitive load*³ ties up scarce cognitive resources, needed for an effective operation of System 2, and the learning performance decreases (Seufert et al. 2017).

Further evidence that readability interacts with System 2 can be gathered from the fact that System 2 is working memory dependent (Whitney et al. 2008). For example, Lehmann et al. (2016) could only find a positive effect of disfluency for learners with higher working memory capacities, indicating that individuals need sufficient resources to successfully employ System 2. As less readable texts burden limited working memory capacities by imposing extraneous cognitive load, too low readability levels may interfere with the deliberative thinking mode of System 2.

2.2 Foreign language effect

On the individual employee-level, the introduction of a foreign language as a common corporate language is mainly seen from the perspective of cognitive load, and, in this regard, foreign language use is expected to entail negative consequences for a decision-maker. Volk et al. (2014) cover this position in their "brain drain"-model by suggesting

³ Extraneous cognitive load comprises all processes which are unrelated to the learning task at hand such as navigating or searching within the learning material, see Lehmann et al. (2016).

that cognition in a foreign language ties up scarce cognitive resources which leads to biased judgements and reduced self-regulation.

However, recent evidence from studies in the field of cognitive-psychology suggest that the “brain drain”-model is not exhaustive and has to be complemented by potentially positive effects of foreign language use which are subsumed under the term ‘foreign language effect’. For example, it has been demonstrated that foreign language use prompts less severe moral evaluations which are supposed to be guided by affective mental processes. Such lenient moral judgments were measured towards actions such as siblings having safe and consensual sex or sacrificing one person’s life to save five other persons (Cipolletti et al. 2016; Geipel et al. 2015; Hayakawa et al. 2017b). Another positive effect of foreign language use was found in the domain of self-regulation. People tend to order healthier deserts in a restaurant (Klesse et al. 2015) and are less willing to lie in a foreign language (Bereby-Meyer et al. 2017).

The most important aspect of the foreign language effect for our study concerns risk-taking and risk perception in general. In the experimental study of Hadjichristidis et al. (2015), participants received materials either in their native or a foreign language and were asked to indicate their attitudes towards innovative technologies such as biotechnology or nanotechnology. Subjects in the foreign language condition rated the technologies as less risky and more beneficial. These findings offer support for the hypothesis that people are less guided by negative feelings towards risks and appreciate potential benefits more when they use a foreign language. Moreover, the use of a foreign language was shown to encourage the willingness to undertake ‘smart’ risks (Hadjichristidis et al. 2017b).

However, another stream of research (partially) failed to detect any foreign language effect. For example, Oganian et al. (2016) scrutinized the effect of foreign language use on the framing effect. The only difference between the experimental groups in terms of risk choices was due to language switching and not as a result of foreign language use per se. Hayakawa et al. (2017a) set out to investigate whether using a foreign language increases the willingness to take risks in general or whether it promotes a more strategic view on risks. The main result was that the effect of foreign language use is not particularly robust across different contexts and populations. Finally, in the study of Winskel et al. (2016) foreign language use leveled out the framing effect only in Study 1 (Asian disease problem), but this effect was completely absent in Study 2 (financial crisis problem).

There are three main explanations for the foreign language effect and they are all based on the dual-process model. The leading account is that the foreign language effect operates through *reduced emotionality* (Hayakawa et al. 2016). For example, reading emotionally laden texts in a foreign language proved to activate brain areas associated with emotional processes to a lesser extent than reading the same texts in one’s native tongue (Hsu et al. 2015). Following this notion, reduced emotionality in the course of foreign language use will enhance the weight of deliberative System 2 relative to affective System 1 in the decision-making process. Another explanation for the foreign language effect refers to an increased *psychological distance* which means that people think in a foreign language on a more abstract level of construal. In other words, using a foreign language could lead people to take a ‘bird’s eye view’ (Hayakawa et al. 2016) wherein the focus on ends over means is increased (Fujita

et al. 2006). This would contribute to the more utilitarian and less risk-averse behavior associated with System 2. Lastly, processing the information in a foreign language relative to one's native tongue is more costly and less fluent. According to the *disfluency effect*, processing difficulty may serve as a metacognitive cue that signals the need to slow down and to engage in a more deliberative thinking mode of System 2 (Alter et al. 2007).

All of the above mentioned explanations have to be seen in combination with language proficiency. As both native-like and lower levels of language proficiency are expected to minimize or prevent the foreign language effect, the positive influence of foreign language use on promoting the use of System 2 is anticipated to be present only for a medium–high language proficiency level (Costa et al. 2017).

3 Development of hypotheses

3.1 Hypothesis 1

In our hypothesis 1 we address the effect of different levels of readability on decision-makers' willingness to accept 'smart' risks. As discussed above, the positive effects of disfluency flip from a certain level of illegibility onwards, causing the negative consequences of the cognitive load effect. Too much extraneous load caused by lower levels of readability proved to deteriorate cognitive performance due to the fact that extraneous tasks distract mental capacities from the actual task at hand (Seufert et al. 2017).

Less readable texts burden working memory by imposing extraneous cognitive load. As working memory capacities are limited and working memory processes are crucial for a successful operation of System 2 (Whitney et al. 2008), lower readability levels may negatively impact the deliberative thinking mode. For example, increased information processing difficulty due to low readability limits the effects of investor sophistication. This is because investors are limited in effectively processing financial documents and, consequently, rely more on heuristic System 1 instead of analytic System 2 (Tan et al. 2014).

In the context of risk-taking, there is strong evidence that burdening working memory induces increased risk aversion. For example, Deck and Jahedi (2015) demonstrated that burdening subjects' working memory capacities via a working-memory task significantly increased their risk aversion. The same result can be expected, if working memory capacities are inhibited by extraneous cognitive load of a low readability level which precludes the possibility to decide in the deliberative thinking mode of System 2. According to this argument, our first hypothesis is:

H1 Decision-makers take less 'smart' risks if they process a management accounting report with a low as opposed to a high readability level.

3.2 Hypothesis 2

In our second hypothesis we address the effect of presenting a management accounting report either in the decision-maker's native tongue or a foreign language on her

willingness to accept ‘smart’ risks. Next, we will discuss how language choice may have an impact on both System 1 and System 2 and argue why the foreign language effect is conditional on high readability in the context of beneficial risk-taking.

There is strong empirical evidence that people are less emotionally aroused if they process information in a foreign language as compared to their native tongue (Hsu et al. 2015). According to the “risk as feeling” hypothesis (Loewenstein et al. 2001) situations involving risks induce strong negative feelings and, hence, fall into the sphere of influence of the foreign language effect. Furthermore, risks may involve some level of emotionality even in the absence of losses. This notion is corroborated by Costa et al. (2014) who demonstrated by means of the Holt–Laury test—where all lottery pairs have positive expected values of outcomes—that individuals who performed the test in a foreign language were less risk-averse than their counterparts in the native language condition. The authors conclude that: “[...] the poor choices, in terms of expected value, prompted by risk aversion stem from the emotional reaction to risk itself” (p. 245). Taken together, the use of a foreign language is supposed to strengthen deliberative System 2 relative to affective System 1, leading to a higher willingness to undertake beneficial risks.

In line with the psychological distance mechanism, the perception of risks is modulated depending on whether individuals think on a concrete or more abstract level of construal (Sagrignano et al. 2002). Using a foreign language as compared to one’s native tongue could lead a decision-maker to take a more abstract level of construal whereby the focus on ends over means is increased (Hayakawa et al. 2016). Through this, an abstract level of construal should promote utilitarian behavior associated with System 2 and exercise a positive effect on the willingness to accept beneficial risks.

Finally, according to the disfluency effect, information processing difficulty serves as a metacognitive cue that signals the need to slow down and activate the more deliberative thinking mode (Alter et al. 2007). Because information processing in a foreign language is more cognitively demanding than in one’s native tongue, the use of a foreign language could activate System 2, thus, leading to a higher willingness to accept beneficial risks. However, the disfluency effect necessitates linguistically enriched texts which will impose some (non-critical) cognitive load. As a typical management accounting report comprises diagrams, tables, as well as textual components (Ohlert and Weißenberger 2015), it can be supposed that the foreign language effect will emerge in a management accounting report as well.

It has been demonstrated that linguistic processing in a foreign language and thinking performed concurrently interfere with each other. Because linguistic processing is a prerequisite for any appropriate response, thinking is sacrificed first if there are insufficient cognitive resources to operate the both tasks in parallel (Takano and Noda 1993). With a low readability level cognitive resources needed for an effective operation of System 2 will be solely occupied by the extraneous cognitive load through linguistic processing. By contrast, in a high readability condition it is possible for the foreign language effect to evolve through enhanced contribution of System 2 in the decision-making process.

Empirical evidence that the foreign language effect is conditional on a high readability level can be gathered from previous psychological studies. Tasks in these studies are formulated either only in a few words (e.g., Holt–Laury test with lottery choices)

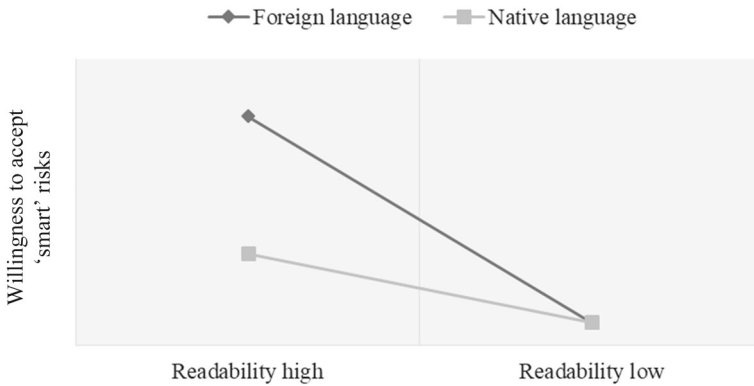


Fig. 1 Visual representation of hypotheses H1 and H2

or in limited simple sentences. Mostly, the Asian disease problem or similar framing scenarios are employed to scrutinize the foreign language effect (Costa et al. 2014; Keysar et al. 2012). The original version of the Asian disease problem in the gain condition (loss condition is comparable) consists of only six sentences with 88 words and a readability value, as measured by the Gunning Fog Index (hereafter, Fog Index), of 9.5—meaning that the text can be understood even by a child (Li 2008). All in all, we hypothesize (Fig. 1):

H2 Only with a high readability level, decision-makers take more ‘smart’ risks if they process a management accounting report in a foreign language as opposed to their native language.

3.3 Hypothesis 3

The impact of language use on risk-taking behavior may be conditional on language proficiency (Hayakawa et al. 2016). On the one hand, it is reasonable to predict that native-like language proficiency should only have a minimal effect on risk-taking behavior because there will be, if any, only a marginal difference in the nature of information processing compared to a native speaker. On the other hand, lower levels of language proficiency impair cognitive resources due to information processing difficulties and this will negatively impact the relative contribution of System 2 in the decision-making process. Thus, the foreign language effect is most likely to be observed only with a medium–high language proficiency level (Costa et al. 2017). Following this line of argumentations we predict:

H3 The foreign language effect, as stated in H2, persist only for a medium–high level of language proficiency.

4 Experimental design and participants

4.1 Experimental design

4.1.1 Procedure and task

Our experiment was set up as a 2×2 between-subject design with randomly assigned participants. The experiment was conducted in four separate rooms simultaneously in the summer term of 2018. Upon arrival in each room, subjects received short written and verbal instructions which were held in German or English, according to the respective language condition. Moreover, to prevent the language switching effect (Oganian et al. 2016), participants in the foreign language condition were invited to exclusively communicate in English.

In part one, participants were provided with a paper-based case study and a questionnaire. After part one had been finished by all participants, it was collected and part two was handed out. Part two consisted of checks of subjects' understanding of the case study, demographic questions, and a language history questionnaire adapted from Li et al. (2006). Finally, participants in the foreign language groups had to complete a short English test (see Sect. 5.2.2 and Appendix 3) with the time limit of 15 min.

With respect to the experimental task, subjects were provided with a case study which comprised a management accounting report of a fictitious company in which two potential purchase offers for the company's production were presented. One purchase offer constituted the safe option with a fixed revenue for the firm, and the other one represented the risky option with a higher expected value of revenues as compared to the safe option.⁴

Participants had to take the role of a company manager and to choose from the two purchase offers. They were instructed that the decision would impact the company's financial performance and that their assumed compensation would partly depend on this performance. We chose this setting in order to put the decision into a context that resembles the real-life context of corporate decision-making. Managers as the recipients and users of management accounting reports usually act on the behalf of others, i.e. the owners of the firm, and their remuneration contains variable components based on the company's performance in order to mitigate agency problems (Hoskisson et al. 2016). As performance is usually measured outcome-based and on an aggregate level, e.g. share price increase or corporate earnings, the link between any single decision and the variable payment received is far less direct, however, as for instance for a sales person with a bonus tied to sales. In order not to overemphasize the link between the decision and its consequences for personal pay we refrained from actually remunerating the participants.

⁴ The risky purchase offer implied two possible levels of revenue, \$200,000 and \$100,000, with equal probabilities of outcomes. The full production costs were given by \$100,000 and, thus, the expected value of net revenues of the risky option amounted to \$50,000. In order to set up an appropriate level of revenue for the safe purchase offer, we consulted prospect theory and used the value and the probability weighting functions as provided by Tversky and Kahneman (1992) for our calculations. This resulted in a revenue of \$138,000 (rounded up) for the safe purchase offer. To rule out any uncertainty, participants were informed that companies offering the purchase prices are long-term clients who pay their bills always on time.

The decision not to remunerate participants and how to frame the context of the decision is important as the degree of accountability for a decision and its outcomes has been found to influence decision-makers willingness to bear risks (Losecaat Vermeer et al. 2019). Depending on context, a high degree of accountability may reinforce risk-averse behavior as observed in a setting in which decision-makers make decisions for themselves (Pollmann et al. 2014). It can be expected that the overall level of risk-aversion would be higher if participants would actually be remunerated, thereby emphasizing the consequences for themselves (as in a lottery game with small payments), and lower if remuneration would not be addressed at all, thereby shifting attention more to the accountability to the company's owners. With our design choice of addressing remuneration hypothetically but not actually remunerating participants we aim at framing the use of the management accounting report in a way that resembles real-life governance structures. The influence of the degree of accountability on the effects of readability and language choice, however, was not investigated in our study.

4.1.2 Main independent and control variables

Our first independent variable is language, native vs. foreign. In the native language condition experimental instructions and the case study were written in German, and in the foreign language condition these materials were provided in English. In order to guarantee that the meaning conveyed in both language conditions is the same, materials, originally written in German, were translated into English and then back-translated into German by two independent bilingual speakers (Brislin 1970). A comparison of the original materials with the back-translated materials showed no substantial differences.

Readability, high vs. low, serves as our second independent variable. In the high readability condition, information contained in the case study should impose only little extraneous cognitive load, and the opposite was pursued in the low readability condition. Several accounting studies used the Fog Index as a readability proxy (e.g., Leheavy et al. 2011; Li 2008; Miller 2010). The problem with the Fog Index is that some multisyllabic words identified as "complex" (e.g., company) would be understood effortlessly even by least sophisticated investors (Loughran and McDonald 2014). Bonsall et al. (2017) argue that traditional readability indices like the Fog Index do not capture all relevant readability aspects of financial disclosures and appeal for the use of a more comprehensive readability measure. In order to create a clear and concise financial disclosures, the SEC provides some very specific recommendations which are formulated in the Plain English Handbook (SEC 1998).⁵ Because readability manipulations with reference to these recommendations proved to be successful in triggering altered cognitive processes through fluency perception and cognitive load (Rennekamp 2012; Tan et al. 2014), we decided to employ some of the SEC's plain English recommendations in our study. As suggested by Rennekamp (2012), we used only the following linguistic features in order to keep the information content

⁵ Generally, the SEC's plain English principles comprise: (1) short sentences; (2) definite, concrete, everyday words; (3) active voice; (4) tabular presentation or bullet lists for complex material, whenever possible; (5) no legal jargon or highly technical business terms; and (6) no multiple negatives.

unchanged: short sentences; active voice; no hidden verbs; no superfluous words; language written in the positive; simple synonyms; personal pronouns; and sentences that keep subject, verb, and object close together. With regard to formatting features, we employed clear headings, bullet points, and appropriate layout of tables. We manipulated the high readability version of our management accounting report by linguistic and formatting choices in accordance with the above mentioned principles, and, by contrast, these principles were violated in the low readability condition. Appendix 2 provides examples of our readability manipulations.

With regard to risk-taking, some control variables have to be taken into account. For example, gender differences in risk-taking are well documented with males predominantly taking more risks than females in the vast majority of contexts (Byrnes et al. 1999). Another important factor is age. Figner et al. (2009) argue that risk-taking peaks in adolescence and decreases again during adulthood. Additionally, Figner and Weber (2011) posit that risk-taking is domain specific which means that some individuals may enjoy risks in one domain (e.g., skydiving), but are more risk-averse in other domains (e.g., investing). In our study, we used the subscale ‘gambling and investing’ from Blais and Weber’s (2006) Domain-Specific Risk-Taking Scale to capture general risk attitudes in this particular domain. Finally, as risk aversion is associated with System 1 rather than with System 2, there is a possibility that cognitive style as a personal trait may influence risk-taking as well. The tendency to rely more on gut feelings and intuition associated with System 1 can be captured by the construct ‘faith in intuition’. By contrast, people who enjoy more logical thinking, which is characteristic of System 2, are expected to score higher on the construct ‘need for cognition’. The (shortened) version of the scale ‘faith in intuition’ was taken from Epstein et al. (1996), and we employed the scale from Beißert et al. (2015) to measure the construct ‘need for cognition’.

4.1.3 Dependent variables

We measured our construct of interest, ‘smart’ risk-taking behavior, in two steps. First, subjects had to make a discrete choice between the two purchase offers in the case study. Hereafter, we asked the participants to provide their individual certainty equivalent (i.e., the required revenue of the safe option with which the subject becomes indifferent between the safe and the risky option). There were no predefined answers for the certainty equivalents, but participants had to provide their answers in form of a free entry. We used this design choice to assure that the subjects took the experimental task seriously as acceptable answers could not be provided by simply checking an arbitrary box. For this reason, we base our statistical analyses on individual certainty equivalents as our main dependent variable.⁶

⁶ Besides individual certainty equivalents, we asked the participants to indicate their preference ratings for the chosen purchase offer. Similar to the discrete choices, subject provided their answers by checking a box. Hence, compared to individual certainty equivalents, preference ratings have no added value in terms of design choice and we skip this variable from our statistical analyses. This being said, the only difference in the results between the two metric variables is that the foreign language effect, which is only marginal for individual certainty equivalents anyway, was not present with the preference ratings as the dependent variable.

Table 1 Descriptive statistics: distribution of participant characteristics between groups

	Readability				Total
	High		Low		
	Language		Language		
	Foreign	Native	Foreign	Native	
Gender					
Female	9 (32.1%)	11 (37.9%)	6 (20.7%)	7 (20.0%)	33 (27.3%)
Male	19 (67.9%)	18 (62.1%)	23 (79.3%)	28 (80.0%)	88 (72.7%)
Program ^a					
Bachelor	13 (50%)	14 (48.3%)	17 (58.6%)	13 (37.1%)	57 (47.9%)
Master	13 (50%)	15 (51.7%)	12 (41.4%)	22 (62.9%)	62 (52.1%)
Age (<i>M</i>)	23.7	23.5	22.6	23.8	23.4
Risk attitude ^b (<i>M</i>)	2.9	3.1	3.1	3.2	3.1
Faith in intuition ^c (<i>M</i>)	4.5	4.7	4.8	4.8	4.7
Need for cognition ^c (<i>M</i>)	5.1	5	4.9	4.7	4.9

^aTwo participants in group 1 didn't provide information on whether they are graduate or undergraduate students

^b*Risk attitude* stands for participant's general attitude towards risks in the domain of 'gambling and investing' according to Blais and Weber (2006)

^c*Faith in intuition* and *need for cognition* are constructs for capturing individual's tendency to rely more on intuition associated with System 1 or logical thinking associated with System 2, accordingly. The (shortened) version of the scale *faith in intuition* was taken from Epstein et al. (1996), and we employed the scale from Beißert et al. (2015) to measure the construct *need for cognition*

4.2 Participants

The majority of participants were business students⁷ from a middle-size university in Germany. All participated on a voluntary basis and anonymity was guaranteed to the subjects. Descriptive statistics on the participants and their distribution between the four experimental groups are provided in Table 1. Subjects were on average 23.4 years old, 27.3% were female, undergraduate and graduate students were nearly equally represented. With respect to our control variables (gender, age, domain-specific risk attitude, need for cognition, and faith in intuition) there were no substantial differences across the experimental groups.

Some of the participants had to be excluded from our main statistical analyses for several reasons. First of all, we employed multiple-choice questions to ensure that participants comprehended that the risky option in the case study had a higher expected value of revenues relative to the safe option and to make sure that everyone understood that the risky option entails no potential loss situation because domains of gains and losses exhibit different patterns of risk-taking (Tversky and Kahneman 1981).

⁷ There were also a small number of students enrolled in mathematical and engineering subjects. For these students it was ensured that they attended relevant business classes necessary for understanding of the case study.

In addition to the objective manipulation checks in form of multiple-choice questions, we asked the participants to indicate their self-reported understanding of the case study. Subjects who failed to provide correct answers to the multiple-choice questions and/or self-reported an understanding of less than 50% were excluded.⁸

Additionally, subjects were skipped from the main analyses if they missed to provide or provided implausible certainty equivalents (i.e., risky option choice in combination with a certainty equivalent below the default safe revenue of \$138,000, and vice versa). Finally, in line with studies on the foreign language effect (e.g., Costa et al. 2014; Keysar et al. 2012), we used specific screening and exclusion criteria with respect to participants' language background. For example, subjects were excluded if they reported not being a native (or comparable) German speaker in the native language condition.

Descriptive statistics of the above mentioned exclusion criteria and their distribution between the experimental groups can be taken from Table 5 in Appendix 1. Our screening procedure led to a relatively steep decline in group sizes, which was especially pronounced in the first group. We will return to this concern in Sect. 5.2.1.

5 Results

5.1 Descriptive statistics

Table 2 reports descriptive statistics of our two dependent variables which capture 'smart' risk-taking behavior of the participants in the experiment: (1) *discrete choice* between the safe and the risky option, i.e., the purchase offers in the case study; and (2) individual *certainty equivalent* for the risky option. As discrete choice is a categorical variable, only absolute numbers and percentages are depicted for each experimental group in Panel A. The certainty equivalents constituting our main dependent variable are metric and allow calculations of descriptive statistics. For this variable, the mean and the standard deviation are reported for each experimental group in Panel B.

Descriptive results suggest that the low readability level induces less willingness to accept 'smart' risks. Taken together, participants in the high readability condition prefer the risky over the safe option more often (80.7%) than their counterparts in the low readability condition (57.8%). Similarly, the average means of certainty equivalents (high readability: 145,901 vs. low readability: 135,158) corroborate H1.

With respect to our hypothesis H2, it is stated that only with a high readability level a management accounting report which is processed in a foreign language as opposed to the native language induces more willingness to accept 'smart' risks. The descriptive results provide first evidence with respect to this hypothesis. With the high readability level, there are more choices in favor of the risky over the safe option in the foreign language groups (89.3%) as compared to the native language groups (72.4%). The

⁸ The same threshold for self-reported understanding is usually employed in studies on the foreign language effect, e.g., Costa et al. (2014), and we argue that a self-reported understanding of at least 50% together with correctly answered multiple-choice questions are sufficient to assure sound responses in the experiment. However, as a robustness check, we will report the overall results with thresholds of 60% and 70% for self-reported understanding in footnotes as well.

Table 2 Descriptive statistics: dependent variables

	Readability				Total
	High		Low		
	Language		Language		
	Foreign	Native	Foreign	Native	
Panel A: Discrete choice ^a (absolute numbers [percentages])					
Risky option choices	25 [89.3%]	21 [72.4%]	15 [51.7%]	22 [62.9%]	83 [68.6%]
Safe option choices	3 [10.7%]	8 [27.6%]	14 [48.3%]	13 [37.1%]	38 [31.4%]
Panel B: Certainty equivalent ^b (mean [standard deviation])					
Certainty equivalents	149,768 [18,847]	142,034 [16,879]	132,673 [17,668]	137,643 [21,447]	140,310 [19,680]

^a*Discrete choice* is a dummy variable conditional on whether the safe or the risky option, i.e., which purchase offer in the case study, was chosen

^b*Certainty equivalent* is our main dependent variable and stands for the individually required revenue of the safe option with which the subject becomes indifferent between the safe and the risky option

same is true for the certainty equivalents within the high readability treatment (foreign language: 149,768 vs. native language: 142,034). By contrast, if the low readability level is taken into consideration, the difference in risk-taking behavior between the foreign and the native language groups is reduced. All in all, **H2** is supported on the descriptive level.

Finally, our third hypothesis **H3** says that the foreign language effect, as stated in **H2**, persists only for a medium–high level of language proficiency.⁹ There is some evidence in favor of this hypothesis on the descriptive level (untabulated). Within the high readability condition, all of the participants with a medium–high level of language proficiency chose the risky option, participants with a low level showed the second highest rate for the risky option choices (82.4%), and the lowest rate was exhibited by the native speakers (72.4%). The same pattern of results can be observed for the certainty equivalents. Within the high readability condition, subjects with a medium–high level of language proficiency indicated on average the highest certainty equivalents, followed by the participants with a low level and the native speakers (161,273; 142,324; and 142,034, respectively).¹⁰

⁹ We will describe the method for measurement of language proficiency and how we categorized the different language proficiency levels in more detail in Sect. 5.2.2.

¹⁰ Within the low readability condition, the difference between the participants with a medium–high language proficiency level and the other participants is less pronounced. Preferences for the risky over the safe option for each language proficiency level are distributed as follows: 66.7% (low); 35.7% (medium–high); and 62.9% (native). And for the certainty equivalents: 135,223 (low); 129,940 (medium–high); and 137,643 (native).

Table 3 Results of ANCOVA and planned contrasts analysis with language

Panel A: Main and interaction effects resulting from ANCOVA				
Predictor variable (certainty equivalent ^a)	df	Sum of squares	<i>F</i> -value	<i>p</i> -value
Language ^b	1	6.23×10^7	0.18	0.674
Readability ^c	1	3.26×10^9	9.33	0.003***
Risk attitude ^d	1	1.44×10^9	4.11	0.045**
Language \times readability	1	1.21×10^9	3.47	0.065*
Panel B: Result of planned contrasts with ANCOVA				
Contrast (certainty equivalent)	df	Contrast estimate	<i>t</i> -value	<i>p</i> -value
Readability high: language: foreign vs. native	116	8,303	1.67	0.097*
Readability low: language: foreign vs. native	116	- 4,397	- 0.94	0.352

All *p*-values are two-tailed with *, **, and *** indicating significance at the 0.10, 0.05, and 0.01 levels

^a*Certainty equivalent* stands for the individually required revenue of the safe option with which the subject becomes indifferent between the safe and the risky option

^b*Language* is a dummy variable with the two categories: native language (German) and foreign language (English)

^c*Readability* is a dummy variable with the two categories: high readability and low readability. Readability was manipulated with the help of the SEC's Plain English Handbook (SEC 1998) by means of linguistic and formatting choices

^d*Risk attitude* stands for participant's general attitude towards risks in the domain of 'gambling and investing' according to Blais and Weber (2006)

5.2 Tests of hypotheses

5.2.1 Hypotheses H1 and H2

To formally test our hypotheses H1 and H2, we conducted an analysis of covariance (ANCOVA) and, as a follow-up test for group differences, a planned contrasts analysis (Sedlmeier and Renkewitz 2008) with control variables as covariates and the following predictor variables: language, readability, and their interaction term. The individual certainty equivalents serve as the dependent variable.¹¹

The results of ANCOVA together with the planned contrasts analysis are displayed in Table 3. With respect to readability, our results are at the highest significance level ($F = 9.33$, $p = 0.003$, Panel A), which fully corroborates our hypothesis H1.¹²

¹¹ Because we had to skip a relatively high number of participants from our final sample and exclusion rates are not equally distributed between the experimental groups, we, additionally, include into our preliminary analysis a dummy variable to account for the exclusions. The output of our full ANCOVA model with the raw sample (untabulated) reveals that from all of our control variables only risk attitude is significant ($F = 3.85$, $p = 0.051$). More importantly, the dummy variable accounting for the exclusions is not significant at all ($F = 0.26$, $p = 0.612$) suggesting that the fact whether a participant was excluded or not does not alter the choice for the individual certainty equivalent. Hence, we base our main statistical analyses on a sample after the exclusions and include in our later ANCOVA models only risk attitude as a control variable.

¹² Highly significant results for readability are also gained if we use models with the raw sample and if we include only participants who reported understanding 60%, 70% of the case study (untabulated).

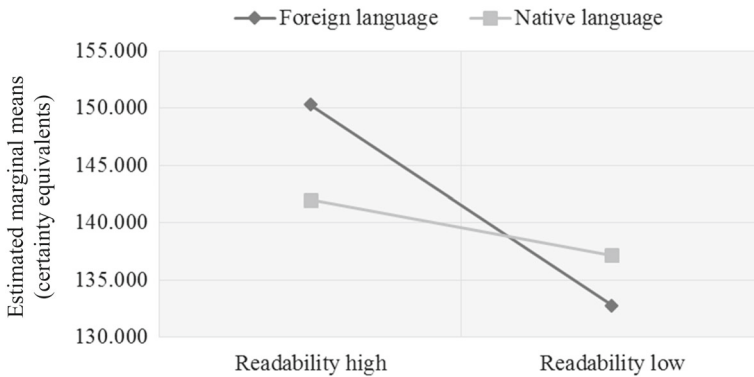


Fig. 2 Interaction between language and readability

The significant interaction term ($F = 3.47$, $p = 0.065$, Panel A) allows us to further scrutinize the foreign language effect as stated in H2. Figure 2 together with Table 3, Panel B depicts the interaction between language and readability. As can be seen, within the high readability condition, there is a positive and slightly significant difference between the foreign and native language groups ($t = 1.67$, $p = 0.097$). Contrary to this, the difference between the foreign and native language groups becomes insignificant if the low readability level is observed ($t = -0.94$, $p = 0.352$). This pattern of results reflects our predictions in hypothesis H2. However, we draw our conclusions with caution as the difference between the foreign and native language groups within the high readability level is only marginally significant and not fully robust against changes in the sample composition due to the application of different exclusion criteria.¹³

In summary, we are able to fully corroborate our hypothesis H1 supporting the notion of the cognitive load effect induced by a low readability level. In other words, participants in this treatment condition took significantly less ‘smart’ risks relative to their counterparts in the high readability condition. In addition, our results provide some tentative evidence in favor of H2. According to this, the use of a foreign language as compared to one’s native tongue may enhance ‘smart’ risk-taking behavior, but only if a high readability level is guaranteed.

5.2.2 Hypothesis H3

The aim of this section is to disentangle the foreign language effect by considering language proficiency levels of the participants in our experiment. In hypothesis H3, we posit that the foreign language effect, as stated in H2, persists only for a medium–high language proficiency level. Before presenting the results, we first briefly explain the elicitation procedure for language proficiency in our experiment.

¹³ With the raw sample the interaction term is not significant at all. If we consider participants who reported understanding 60%, 70% of the case study, results are dispersed around the 10% significance level (untabulated).

We employed the so-called C-Test (Raatz and Klein-Braley 2002), a quick and reliable language test (for an example see Appendix 3).¹⁴ In the context of a C-Test it is important to note that individual scores have to be regarded relative to the performance of the whole test group, and the interpretation of absolute language proficiency scores alone is not meaningful (Grotjahn 2002). Following this notion, we subdivided our participants into three groups of language proficiency levels: (1) all subjects in the mother tongue groups indicating themselves as native German speakers or comparable were classified as native speakers; (2) 25% of the best performing subjects in the C-Test were assigned to the medium–high proficiency level; and (3) all other participants were grouped into the low proficiency level. Subjects who were grouped into the low proficiency level scored on average 56%, and those in the medium–high proficiency level 78% in the C-Test. The difference between the both groups is highly significant ($p < 0.010$, untabulated). Moreover, as our three texts were taken from English textbooks designed for a medium–high language proficiency level on average, we conclude that our categorizations reflect the designated language proficiency levels on average.

In order to test our hypothesis H3—similar to the procedure of Oganian et al. (2016) in the context of the foreign language effect—we disentangle the language variable in our previous ANCOVA model by the language proficiency variable (Table 4, Panel A).

There is a highly significant main effect of readability ($F = 10.16$, $p = 0.002$) and a significant interaction term ($F = 4.57$, $p = 0.012$). Moreover, the follow-up contrasts analysis (Table 4, Panel B) reveals that within the high readability condition, there is a significant positive difference between both the medium–high vs. low language proficiency groups ($t = 2.61$, $p = 0.010$) and the medium–high vs. native speakers groups ($t = 3.01$, $p = 0.003$). On the other hand, the difference between the low vs. native speakers groups is not significant ($t = 0.18$, $p = 0.860$), as well as all comparisons within the low readability conditions (all p values > 0.269).¹⁵ These results are in line with the prediction in H3 and provide—to our best knowledge—the first statistical evidence that the foreign language effect, indeed, only persist for a medium–high level of language proficiency. This effect is also clearly represented by an inverted U-shaped pattern of estimated marginal means of certainty equivalents with language proficiency for the high readability condition (Table 4, Panel C).

¹⁴ In order to evaluate the appropriateness of our C-Test, we conducted an item analysis according to Grotjahn (2002). For all three texts the item difficulty and item discrimination are sufficiently high with values above 0.45 and 0.70, respectively. The reliability coefficient is on a very high level with a Cronbach's α of 0.86. Finally, the explanatory factor analysis confirmed the existence of only one factor suggesting that the underlying texts measure the same unobserved latent variable, namely, language proficiency.

¹⁵ Inferentially identical results are gained if we include only participants who reported understanding 60%, 70% of the case study (untabulated). Notwithstanding this, the interaction term becomes insignificant if we use the raw sample (untabulated).

Table 4 Results of analyses with language proficiency

Panel A: main and interaction effects resulting from ANCOVA				
Predictor variable (certainty equivalent ^a)	df	Sum of squares	<i>F</i> value	<i>p</i> value
Language proficiency ^b	1	3.77×10^8	0.56	0.571
Readability ^c	1	3.40×10^9	10.16	0.002***
Risk attitude ^d	1	1.52×10^9	4.57	0.035**
Language proficiency \times readability	1	3.06×10^9	4.57	0.012**
Panel B: result of planned contrasts with ANCOVA				
Contrast (Certainty equivalent)	df	Contrast estimate	<i>t</i> value	<i>p</i> value
Readability high: language proficiency				
Medium–high vs. low	114	18,507	2.61	0.010**
Medium–high vs. native	114	19,499	3.01	0.003**
Low vs. native	114	992	0.18	0.86
Readability low: language proficiency				
Medium–high vs. low	114	– 3910	– 0.57	0.568
Medium–high vs. native	114	– 6460	– 1.11	0.269
Low vs. native	114	– 2550	– 0.45	0.652
Panel C: estimated marginal means of certainty equivalents				
	Language proficiency			
	Native	Medium–high	Low	
Readability high	141,982	161,481	142,974	
Readability low	137,217	130,757	134,667	

All *p* values are two-tailed with *, **, and *** indicating significance at the 0.10, 0.05, and 0.01 levels

^a*Certainty equivalent* stands for the individually required revenue of the safe option with which the subject becomes indifferent between the safe and the risky option

^b*Language proficiency* was measured by means of a C-Test (see Sect. 5.2.2 and Appendix 3). The participants were subdivided into three language proficiency groups: low; medium–high; and native

^c*Readability* is a dummy variable with the two categories: high readability and low readability. Readability was manipulated with the help of the SEC's Plain English Handbook (SEC 1998) by means of linguistic and formatting choices

^d*Risk attitude* stands for participant's general attitude towards risks in the domain of 'gambling and investing' according to Blais and Weber (2006)

5.3 Supplemental analysis

5.3.1 Manipulation checks and alternative explanations

In this section we check the validity of our readability manipulations and address potential drivers which could serve as alternative explanations for our results.

First of all, we provide evidence that our readability manipulations were successful and the results conform to H1. Based on rating scales of Elliott et al. (2015), we measured perception of readability by asking participants to indicate on a seven-point

Likert scale ranging from 1 (= very easy) to 7 (= very difficult) how ‘difficult-to-read’, ‘difficult-to-understand’, and ‘difficult-to-process’ they perceived the information presented in the case study to be. As all of our readability manipulation questions are highly correlated (all values for Pearson’s $r > 0.61$), we computed a single average value which was then used as our readability measure. Results of ANOVA (untabulated) with the readability measure as the dependent variable and language, readability, and their interaction term as the independent variables showed that language and the interaction term are insignificant ($F = 1.55$ and $F = 2.03$, $p = 0.215$ and $p = 0.157$, two-tailed, respectively). By contrast, readability is highly significant ($F = 7.91$, $p < 0.006$, two-tailed) with the average mean value of 2.22 for the high readability condition and 2.77 for the low readability condition. It can be concluded that our readability manipulations were successful and conform to H1.

Another potential key driver behind the results could be diminished understanding of some participants in our experiment due to low readability and/or foreign language use. In order to rule out this potential explanatory factor, we computed an ANOVA (untabulated) with participants’ self-reported understanding of the case study in dependence of language, readability, and their interaction term. According to the results of ANOVA none of the dependent variables is significant (all p -values > 0.500). Furthermore, subjects who indicated an understanding of less than 50% accounted only for 8.6% of all participants in the experiment. Taken together, we conclude that the lack of understanding cannot serve as an explanatory factor for our results.

Finally, although we ensured that all participants of our experiment attended relevant business classes necessary for understanding of the case study, the most exclusions were due to incorrectly answered manipulation checks, one of which encompassed basic economic understanding (i.e., expected value concept). With respect to the experimental design we put great emphasis on the clarity of the relevant aspects in the case study—particularly, revenues and probabilities were printed in bold—and a short definition of the expected value concept was presented in parentheses along with the manipulation check questions. Furthermore, there were no special issues during the experimental procedure. We can only speculate about the reason for the high dropout rate. As subjects were intentionally not triggered to rely on formal calculations (i.e., expected value concept), some individuals could have based their decisions on another individual criteria without paying much attention to the specific probabilities and revenues in the case study.

5.3.2 Psychological mechanisms behind the foreign language effect

The aim of this section is to investigate if we can attribute the foreign language effect to any of the discussed psychological mechanisms potentially underlying this effect. In particular, we tested the following psychological mechanisms: reduced emotionality, psychological distance, and the disfluency effect.

The first mechanism posits that reduced emotionality through the use of a foreign language leads individuals to become less prone to affective behavior, and, consequently, decision-makers rely more on the deliberative thinking mode of System 2 which promotes rational decisions—in our case, a higher willingness to accept

'smart' risks. However, with lower levels of language proficiency and/or a low level of readability, System 2 is occupied by information processing difficulties so much that decision-makers are supposed to be still guided by System 1. So, it can be expected that the positive effect of reduced emotionality persist only for a medium–high level of language proficiency and under the condition of a high readability level.

To estimate to what extent participants' decisions in the case study were triggered by any emotional component, subjects indicated on a scale ranging from 1 (= neither happy nor unhappy) to 7 (= very happy) their *positive* feelings towards the safe revenue of the riskless purchase offer in the case study. In addition, to measure the extent of *negative* feelings towards the potential low revenue resulting from the risky purchase offer, we asked the participants to indicate on a scale ranging from 1 (= not sad at all) to 7 (= very sad) how they would feel in the event of the low revenue.

In the first step, we tested the reduced emotionality account by employing an ANOVA model (untabulated) with language, readability and their interaction term in dependence of positive/negative feelings. The results are significant neither for positive nor for negative feelings. In the second step, we used a more accurate measure for the foreign language effect in form of language proficiency levels. The interaction term then becomes significant for positive feelings ($F = 3.30, p = 0.040$). For the high readability level, participants from our language proficiency groups indicated on average the following positive feelings: 3.00 (low); 1.91 (medium–high); and 2.76 (native). For the low readability level, these values are as follows: 2.60 (low); 2.93 (medium–high); and 2.57 (native). Follow-up tests by means of a contrasts analysis (untabulated) showed that subjects with a medium–high language proficiency level and within the high readability condition were significantly less affected by positive feelings than the other language proficiency groups.

Eventually, we used a causal model to see whether the effect of a medium–high language proficiency level on enhanced willingness to accept 'smart' risks, as measured by the individual certainty equivalents, is mediated by reduced positive feelings towards the safe revenue within the high readability condition (see Fig. 3). First, we regressed individual certainty equivalents on language proficiency and risk attitude as a control variable. Consistent with our expectations, compared to other levels of language proficiency, subjects with a medium–high language proficiency level provided substantially higher certainty equivalents ($\beta = 18,466, p = 0.006$). Likewise, there is a highly significant negative link between the medium–high language proficiency level and positive feelings ($\beta = -1.09, p = 0.009$). Lastly, if we employ a regression model with certainty equivalents as the dependent variable and language proficiency, risk attitude, and positive feelings (mediator) as the independent variables, the association between the medium–high language proficiency level and certainty equivalents becomes far less significant ($\beta = 12,924, p = 0.053$). To test whether the extent of observed mediation is statistically significant, we used a bootstrapping-based mediation analysis (Preacher and Hayes 2004) with 5,000 simulations and find that the mediation effect is statistically significant ($p = 0.005$). Summing up, our results support the reduced emotionality account, which is the leading explanation for the foreign language effect. In particular, participants in our experiment with a medium–high language proficiency level were significantly less attracted by the safe but less beneficial

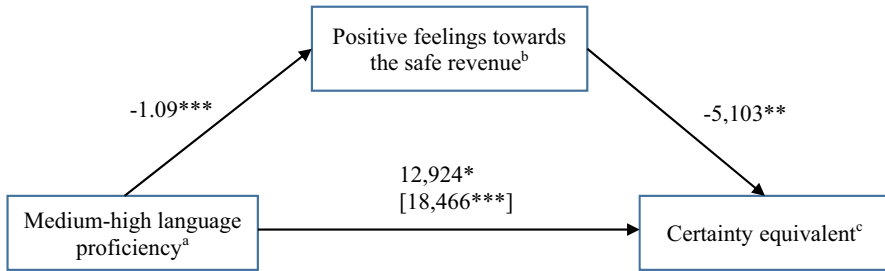


Fig. 3 Test of causal model (high readability condition). All p-values are two-tailed with *, **, and *** indicating significance at the 0.10, 0.05, and 0.01 levels. ^a *Language proficiency* was measured by means of a C-Test (see Sect. 5.2.2 and Appendix 3). The participants were subdivided into three language proficiency groups: low; medium–high; and native. In the regression model, the low level of language proficiency serves as the reference category. ^b *Positive feelings toward the safe revenue* indicate the extent of attraction of the safe purchase offer in the case study. ^c *Certainty equivalent* stands for the individually required revenue of the safe option with which the subject becomes indifferent between the safe and the risky option

financial result and, hence, exhibited more willingness to undertake ‘smart’ risks within the high readability condition.¹⁶

Another explanation for the foreign language effect is based on the psychological distance mechanism. Applied to risk-taking behavior, this explanation would mean that foreign language use promotes more sensitivity to desirability considerations (e.g., the possible high revenue) through an abstract high-level construal. Whereas, in the native language, sensitivity to feasibility considerations (e.g., the probabilities of the revenues) is more pronounced due to a concrete low-level construal (Sagrignano et al. 2002). Following this logic, participants were asked to indicate how much their decision was influenced by the: (1) highest potential revenue; and (2) probabilities for the different revenues. For these dependent variables none of our ANOVA models with language (proficiency), readability and their interaction term as the independent variables produce significant results (untabulated). Taken together, the psychological distance account can be rejected with our results.

Finally, according to the disfluency effect, the use of a foreign language may promote the deliberative thinking mode of System 2 and, hence, more willingness to ‘accept’ smart risks, by generating a metacognitive cue of difficulty. We tested this account by using the same questions we employed for the readability manipulation checks. As already described in the previous section, only the readability variable is significant for the composite value of these questions (same applies to language proficiency). Thus, our results do not support the disfluency effect as an explanation for the foreign language effect.

6 Discussion and conclusion

Our study was designed to examine whether and how certain linguistic characteristics of management accounting reports influence ‘smart’ risk-taking behavior. In partic-

¹⁶ Inferentially identical results are gained if we include only participants who reported understanding 60%, 70% of the case study (untabulated).

ular, we scrutinized the impact of readability and foreign language use. Besides, we further investigated the foreign language effect by incorporating the role of language proficiency and by testing the potential psychological mechanisms.

To test our hypotheses, we conducted a between-subject experiment. Participants, who took the role of a manager, were provided with a fictitious management accounting report. Based on the information in the accounting report, subjects had to select between a safe and a risky but more beneficial option. Hereafter, subjects provided their individual certainty equivalents, which constituted our main dependent variable for measuring the willingness to accept 'smart' risks. We manipulated the high readability level by adhering to certain linguistic and formatting principles provided by the SEC (1998). In contrast, in the low readability condition these principles were violated. With respect to language, one part of the subjects received the management accounting report in their native language and the other part in a foreign language. Language proficiency was measured by a standardized C-Test.

The key finding is that the willingness to undertake 'smart' risks is significantly reduced when readability is low. This phenomenon can be explained by the cognitive load effect which posits that a very difficult to read document induces cognitive load preventing a decision-maker from reaching a decision in the deliberative thinking mode of System 2 according to the dual-process model. Instead, a decision-maker is guided by the affective thinking mode of System 1.

According to the foreign language effect, information processing in a foreign language enhances the contribution of System 2 relative to System 1 in a decision-making process and, as a result, the willingness to accept 'smart' risks due to reduced emotionality, psychological distance, and/or the disfluency effect. However, as a low readability level occupies scarce cognitive resources, needed for an effective operation of System 2, the foreign language effect is expected to occur only within a high readability condition. The foreign language effect can be further disentangled if language proficiency is considered. Language proficiency approaching native-like levels will alter decision outcomes only marginally. On the other hand, lower levels of language proficiency are anticipated to induce cognitive load by itself through information processing difficulties which, in turn, inhibit effective operation of System 2. Thus, the effect of foreign language use is expected to only persist for a medium-high level of language proficiency.

We were able to provide only marginally significant and not very robust evidence in favor of the classical foreign language effect for the aggregated language variable. However, after disaggregating language by language proficiency, the results become more significant and reveal that the foreign language effect persists only for a medium-high language proficiency level within the high readability condition.

Finally, our supplemental analyses with respect to the potential underlying psychological mechanisms behind the foreign language effect showed that subjects with a medium-high language proficiency level and under the condition of a high readability level were significantly less affected by positive feelings towards the safe revenue. This pattern of results corroborates the reduced emotionality account. We could not find support for the other psychological mechanisms with our data.

One possible reason for less significant results with respect to the foreign language effect can originate in the design choice for our experiment in which the manager

makes a risk decision on behalf of others. According to the *empathy gap* hypothesis (Loewenstein 1996), individuals make inaccurate inferences about others' risk preferences and base their decisions on that. In particular, when an agent makes a risk decision on behalf of a principal, who is assumed to be risk-neutral, this decision is partly based on his own risk preferences, and partly on risk neutrality (Desmoulin-Lebeault and Meunier 2018), which should lead to higher risk tolerance in sum.¹⁷ In the context of our experiment, if risk acceptance is heightened in general, the potential for finding a strong foreign language effect is minimized. In the experiment, participants were instructed to assume that their remuneration as a manager responsible for the decision would depend on the financial outcome. It remains unclear to which extent this description of the context helps to counter the previously described effect. Future accounting studies could build on our findings and extend it to incorporate different levels of accountability ranging from a setting with a strong focus on accountability and without any mention of remuneration and personal financial consequences at all to a setting with strong financial incentives provided by actual payments to participants and without reference to typical corporate governance structures involving agency conflicts. Introducing real payments could possibly reinforce the foreign language effect as they can be expected to trigger a higher degree of emotional involvement. Indeed, subjects in our experiment provided on average only 2.66 for positive and 2.79 for negative feelings on a seven-point scale. However, psychological studies suggest that these emotionality levels are sufficient to elicit foreign language effect. For example, Hadjichristidis et al. (2017a) scrutinized superstitious beliefs and indicated a strong foreign language effect for participants in bad-luck scenarios, who provided values below 3 on a 9-point scale for their feelings (Study 2 and 3). Moreover, the foreign language effect emerged in the context of risk-taking without real payments as well (e.g., Study 2 of Keysar et al. 2012).

There remain further fruitful areas for future research. First of all, we reduced the decision rule in our case study to a simple calculation of expected values. By contrast, real-world business situations are usually characterized by complex optimization strategies. For example, Schedlinsky et al. (2016) examined risk-taking under tournament compensation schemes which require consideration of the behavior of other contestants in accordance with game theory. The authors found that in such a sophisticated decision situation individuals were likely to be misled by simplified decision rules, which, in turn, induced excessive risk-taking. Now, it is an open question whether the use of a foreign language will lead to more or less risk-taking in contexts with sophisticated decision rules and undesirable risks.

Summing up, our study sheds light on the behavioral aspects of readability and language choice on 'smart' risk-taking in a management accounting context. Management accountants who prepare information and provide it to a decision-maker should take into account that, even if the information was properly understood by the recipient, some linguistic characteristics may still modulate the decision-making process. Our study empirically shows that a poorly written management accounting report is potentially detrimental to the willingness to undertake 'smart' risks. Moreover, in the

¹⁷ Furthermore, in the domain of positive outcomes, as in our experiment, gains for others might not be integrated in the decision-maker's own mental account which should lead to lower risk aversion (Losecaat Vermeer et al. 2019).

light of the fact that English is increasingly used as the corporate language all over the world, we provide first evidence that the use of a foreign language in management reporting may have a positive effect on the willingness to accept ‘smart’ risks. However, the necessary prerequisite for this effect are: a medium–high language proficiency level of the decision-maker, and a highly readable management accounting report.

Acknowledgements Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix 1: Exclusions

See Table 5.

Table 5 Exclusion criteria and their distribution between groups

	Readability				Total
	High		Low		
	Language		Language		
	Foreign	Native	Foreign	Native	
Sample (raw)	69 (100%)	49 (100%)	52 (100%)	52 (100%)	222 (100%)
Manipulation checks ^a	36 (52.2%)	15 (30.6%)	18 (34.6%)	16 (30.8%)	85 (38.3%)
Understanding ^b	1 (1.5%)	2 (4.1%)	1 (1.9%)	0 (0%)	4 (1.8%)
Other ^c	4 (5.8%)	3 (6.1%)	4 (7.7%)	1 (1.9%)	12 (5.4%)
Sample (final)	28 (40.6%)	29 (59.2%)	29 (55.8%)	35 (67.3%)	121 (54.5%)

^a*Manipulation checks* consisted of two multiple-choice questions. The first question dealt with the expected value of the revenues of the risky option. The second question was intended to rule out the assumption of a potential loss situation. If one of the both questions was answered incorrectly, subjects were rejected

^b*Understanding* represented our second exclusion criterion after the manipulation checks. Remaining participants after the manipulation checks questions who reported understanding of less than 50% for the case study were excluded. Following participants provided incorrect answers to the manipulation checks and reported understanding of less than 50% at the same time: 8 in group 1; 1 in group 2; 3 in group 3; and 3 in group 4

^c*Other* exclusions comprised missing or incorrectly provided individual certainty equivalents (i.e., risky option choice in combination with a certainty equivalent below the default safe revenue of \$138,000, and vice versa) and language specific issues (e.g., subjects were excluded if they reported not being a native (or comparable) German speaker in the native language condition). We employed this exclusions to the remaining subjects after the first two screening stages. Following participants failed at all three exclusion categories at the same time: 2 in group 1; 0 in group 2; 1 in group 3; and 1 in group 4

Appendix 2: Readability manipulations

See Table 6.

Table 6 Examples of linguistic characteristics included in the experiment

Linguistic characteristic ^a	Low readability level	High readability level
Short sentences	Since demand forecasts also showed no fluctuating increases, the company last year made a decision not to entertain a time lag with regard to extensive restructuring work	Demand forecasts at that time showed stable increases. That's why we decided to carry out extensive renovations
Active voice	A base price of \$100,000 is offered to us	They offer us a base price of \$100,000
No hidden verbs	The overtaking of Micro-PC AG's strongest competitor was made possible with the help of a net income of \$10.25 Mio. in the last fiscal cycle	With a net income of \$10.25 Mio. in the previous year, we were able to overtake our strongest competitor
No superfluous words	Now I am pleased to declare that the restructuring work has been successfully completed [...]	The renovations have now been completed [...]
Language written in the positive	Now, this leadership position ought not to be lost	Now, we need to maintain our leadership position in the market
Personal pronouns	Both companies are Micro-PC AG's long-term clients [...]	Both companies are our long-term clients [...]
Sentences that keep subject, verb, and object close together	The cost overview of the additional production of 1,000 computers, which was summarized by the management accounting department of Micro-PC AG, can be found on the next page	Our management accounting department has summarized the cost overview for the additional production of 1,000 computers on the next page

^aLinguistic characteristics are taken from the SEC's (1998) Plain English Handbook

Appendix 3: C-Test

A C-Test is a reliable and quick procedure to elicit general language proficiency. It consists of several short texts and is based on the concept of reduced redundancy. In these texts the first and the last sentence remain unchanged. In the other sentences the “rule of two” is applied: beginning in sentence two the second half of every second word is deleted (Raatz and Klein-Braley 2002). Numbers and proper names are not manipulated. The deletion procedure is continued until approximately 25 blanks have been produced.

The development and evaluation of our version of the C-Test was based on the instructions provided by Grotjahn (2002). To attain a heterogeneous sample of texts, our three texts originated from English textbooks with the language proficiency levels: B1, B2, and C1. For the whole test the time restriction was set to 15 min. The following passage is an excerpt from our first text and should give an impression of how a C-Test looks like:

Many people think that we laugh because we see or hear something funny, but most of the time this isn't true.

In a st_____, a prof_____ of psych_____ and h_____ students list_____ to a_____ made no_____ on hund_____ of conver_____ in pub_____ places.[...].

In order to evaluate the appropriateness of our C-Test, we conducted a pretest with 107 university students several months before our actual experiment. The item analysis revealed inferentially identical results to those as described in Sect. 5.2.2.

References

- Alter AL, Oppenheimer DM, Epley N, Eyre RN (2007) Overcoming intuition: metacognitive difficulty activates analytic reasoning. *J Exp Psychol Gen* 136:569–576
- Asay HS, Elliott WB, Rennekamp K (2017) Disclosure readability and the sensitivity of investors' valuation judgments to outside information. *Account Rev* 92:1–25
- Beißert H, Köhler M, Rempel M, Beierlein C (2015) Kurzsкала Need for Cognition NFC-K. Zusammenstellung sozialwissenschaftlicher Items und Skalen. <https://doi.org/10.6102/zis230>
- Bereby-Meyer Y, Hayakawa S, Shalvi S, Corey JD, Costa A, Keysar B (2017) Honesty speaks a second language. *Top Cognit Sci*. <https://doi.org/10.1111/tops.12360>
- Blais A-R, Weber EU (2006) A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgm Decis Mak* 1:33–47
- Bloomfield RJ (2002) The “incomplete revelation hypothesis” and financial reporting. *Account Horiz* 16:233–243
- Bonsall SB IV, Leone AJ, Miller BP, Rennekamp K (2017) A plain English measure of financial reporting readability. *J Account Econ* 63:329–357
- Borzykowski B (2017) The international companies using only English: a growing number of global firms are using English as their main language—even if they are based in Japan or France. BBC. <https://www.bbc.com/capital/story/20170317-the-international-companies-using-only-english>. Accessed 19 October 2018
- Brislin RW (1970) Back-translation for cross-cultural research. *J Cross Cult Psychol* 1:185–216
- Byrnes JP, Miller DC, Schafer WD (1999) Gender differences in risk taking: a meta-analysis. *Psychol Bull* 125:367–383
- Cipolletti H, McFarlane S, Weissglass C (2016) The moral foreign-language effect. *Philos Psychol* 29:23–40
- Costa A, Foucart A, Arnon I, Aparici M, Apesteguia J (2014) “Piensa” twice: on the foreign language effect in decision making. *Cognition* 130:236–254

- Costa A, Vives M-L, Corey JD (2017) On language processing shaping decision making. *Curr Dir Psychol Sci* 26:146–151
- Deck C, Jahedi S (2015) The effect of cognitive load on economic decision making: a survey and new experiments. *Eur Econ Rev* 78:97–119
- Desmoullins-Lebeault F, Meunier L (2018) Moment risks: investment for self and for a firm. *Decis Anal* 15:242–266
- Elliott WB, Rennekamp KM, White BJ (2015) Does concrete language in disclosures increase willingness to invest? *Rev Acc Stud* 20:839–865
- Epstein S, Pacini R, Denes-Raj V, Heier H (1996) Individual differences in intuitive–experiential and analytical–rational thinking styles. *J Pers Soc Psychol* 71:390–405
- Evans JSB (2008) Dual-processing accounts of reasoning, judgment, and social cognition. *Annu Rev Psychol* 59:255–278
- Figner B, Weber EU (2011) Who takes risks when and why? Determinants of risk taking. *Curr Direct Psychol Sci* 20:211–216
- Figner B, Mackinlay RJ, Wilkening F, Weber EU (2009) Affective and deliberative processes in risky choice: age differences in risk taking in the Columbia Card Task. *J Exp Psychol Learn Mem Cogn* 35:709–730
- Frankish K (2010) Dual-process and dual-system theories of reasoning. *Philosophy Compass* 5:914–926
- Fujita K, Henderson MD, Eng J, Trope Y, Liberman N (2006) Spatial distance and mental construal of social events. *Psychol Sci* 17:278–282
- Geipel J, Hadjichristidis C, Surian L (2015) How foreign language shapes moral judgment. *J Exp Soc Psychol* 59:8–17
- Grotjahn R (2002) Konstruktion und Einsatz von C-Tests: Ein Leitfaden für die Praxis. In: Grotjahn R (ed) *Der C-test: Theoretische Grundlagen und praktische Anwendungen*. AKS-Verlag, Bochum, pp 211–225
- Hadjichristidis C, Geipel J, Savadori L (2015) The effect of foreign language in judgments of risk and benefit: the role of affect. *J Exp Psychol Appl* 21:117–129
- Hadjichristidis C, Geipel J, Surian L (2017a) Breaking magic: foreign language suppresses superstition. *Q J Exp Psychol* 72:18–28
- Hadjichristidis C, Geipel J, Surian L (2017b) How foreign language affects decisions: rethinking the brain-drain model. *J Int Bus Stud* 48:645–651
- Hayakawa S, Costa A, Foucart A, Keysar B (2016) Using a foreign language changes our choices. *Trends Cognit Sci* 20:791–793
- Hayakawa S, Lau BKY, Holtzmann S, Costa A, Keysar B (2017a) On the reliability of the foreign language effect on risk-taking. *Q J Exp Psychol* 72:29–40
- Hayakawa S, Tannenbaum D, Costa A, Corey JD, Keysar B (2017b) Thinking more or feeling less? Explaining the foreign-language effect on moral judgment. *Psychol Sci* 28:1387–1397
- Hoskisson RE, Chirico F, Zyung J, Gambeta E (2016) Managerial risk taking: a multitheoretical review and future research agenda. *J Manag* 43:137–169
- Hsu C-T, Jacobs AM, Conrad M (2015) Can Harry Potter still put a spell on us in a second language? An fMRI study on reading emotion-laden literature in late bilinguals. *Cortex* 63:282–295
- Kahneman D (2011) *Thinking, fast and slow*. Farrar, Straus and Giroux, New York
- Keysar B, Hayakawa SL, An SG (2012) The foreign-language effect: thinking in a foreign tongue reduces decision biases. *Psychol Sci* 23:661–668
- Klesse A-K, Levav J, Goukens C (2015) The effect of preference expression modality on self-control. *J Consum Res* 42:535–550
- Lehavy R, Li F, Merkley K (2011) The effect of annual report readability on analyst following and the properties of their earnings forecasts. *Account Rev* 86:1087–1115
- Lehmann J, Goussios C, Seufert T (2016) Working memory capacity and disfluency effect: an aptitude-treatment-interaction study. *Metacognition Learn* 11:89–105
- Li F (2008) Annual report readability, current earnings, and earnings persistence. *J Account Econ* 45:221–247
- Li P, Sepanski S, Zhao X (2006) Language history questionnaire: a web-based interface for bilingual research. *Behav Res Methods* 38:202–210
- Loewenstein G (1996) Out of control: visceral influences on behavior. *Organ Behav Hum Decis Process* 65:272–292
- Loewenstein GF, Weber EU, Hsee CK, Welch N (2001) Risk as feelings. *Psychol Bull* 127:267–286

- Losecaat Vermeer AB, Boksem MAS, Sanfey AG (2019) Third-party decision-making under risk as a function of prior gains and losses. *J Econ Psychol*. <https://doi.org/10.1016/j.joep.2019.102206>
- Loughran T, McDonald B (2014) Measuring readability in financial disclosures. *J Financ* 69:1643–1671
- Miller BP (2010) The effects of reporting complexity on small and large investor trading. *Account Rev* 85:2107–2143
- Oganian Y, Korn CW, Heekeren HR (2016) Language switching—but not foreign language use per se—reduces the framing effect. *J Exp Psychol Learn Mem Cogn* 42:140–148
- Ohlert CR, Weißenberger BE (2015) Beating the base-rate fallacy: an experimental approach on the effectiveness of different information presentation formats. *J Manag Control* 26:51–80
- Pavlenko A (2012) Affective processing in bilingual speakers: disembodied cognition? *Int J Psychol* 47:405–428
- Pollmann MMH, Potters J, Trautmann ST (2014) Risk taking by agents: the role of ex-ante and ex-post accountability. *Econ Lett* 123:387–390
- Preacher KJ, Hayes AF (2004) SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav Res Methods Instrum Comput* 36:717–731
- Raatz U, Klein-Braley C (2002) Introduction to language testing and to C-test. In: Coleman JA, Grotjahn R, Raatz U (eds) *University language testing and the C-test*. AKS-Verlag, Bochum, pp 75–91
- Rennekamp K (2012) Processing fluency and investors' reactions to disclosure readability. *J Account Res* 50:1319–1354
- Sagrignano MD, Trope Y, Liberman N (2002) Time-dependent gambling: odds now, money later. *J Exp Psychol Gen* 131:364–376
- Schedlinsky I, Sommer F, Wöhrmann A (2016) Risk-taking in tournaments: an experimental analysis. *J Bus Econ* 86:837–866
- Securities and Exchange Commission (SEC) (1998) A plain English handbook: how to create clear SEC disclosure documents. SEC Office of Investor Education and Assistance. <https://www.sec.gov/pdf/handbook.pdf>. Accessed 25 Oct 2018
- Sedlmeier P, Renkewitz F (2008) *Forschungsmethoden und Statistik in der Psychologie*. Pearson Studium, München
- Seufert T, Wagner F, Westphal J (2017) The effects of different levels of disfluency on learning outcomes and cognitive load. *Instr Sci* 45:221–238
- Slooman SA (1996) The empirical case for two systems of reasoning. *Psychol Bull* 119:3–22
- Stanovich KE, West RF (2000) Individual differences in reasoning: implications for the rationality debate? *Behav Brain Sci* 23:645–665
- Stanovich KE, West RF (2008) On the relative independence of thinking biases and cognitive ability. *J Pers Soc Psychol* 94:672–695
- Takano Y, Noda A (1993) A temporary decline of thinking ability during foreign language processing. *J Cross Cult Psychol* 24:445–462
- Tan H-T, Wang E, Zhou B (2014) When the use of positive language backfires: the joint effect of tone, readability, and investor sophistication on earnings judgments. *J Account Res* 52:273–302
- Tietze S (2008) *International management and language*. Routledge, London
- Tversky A, Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* 211:453–458
- Tversky A, Kahneman D (1992) Advances in prospect theory: cumulative representation of uncertainty. *J Risk Uncertain* 5:297–323
- Volk S, Köhler T, Pudelko M (2014) Brain drain: the cognitive neuroscience of foreign language processing in multinational corporations. *J Int Bus Stud* 45:862–885
- Whitney P, Rinehart CA, Hinson JM (2008) Framing effects under cognitive load: the role of working memory in risky decisions. *Psychon Bull Rev* 15:1179–1184
- Winkel H, Ratitamkul T, Brambley V, Nagarachinda T, Tiencharoen S (2016) Decision-making and the framing effect in a foreign and native language. *J Cognit Psychol* 28:427–436
- You H, Zhang X-j (2009) Financial reporting complexity and investor underreaction to 10-K information. *Rev Acc Stud* 14:559–586