The Development of An Online Test to Measure the Interpretation of Implied Meanings as A Major Constituent of Pragmatic...

Research · December 2017

DOI: 10.13140/RG.2.2.33994.06084

CITATIONS
0

READS
65

3 authors, including:

Ugur Recep Cetinavci
Uludag University
9 PUBLICATIONS **11** CITATIONS
SEE PROFILE

Ismet Öztürk
Uludag University
3 PUBLICATIONS **148** CITATIONS
SEE PROFILE

The Development of An Online Test to Measure the Interpretation of Implied Meanings as A Major Constituent of Pragmatic Competence*

Uğur Recep ÇETİNAVCI
English Language Teaching Department, Faculty of Education
Uludag University, Turkey
cetinavci@uludag.edu.tr

İsmet ÖZTÜRK
English Language Teaching Department, Faculty of Education
Uludag University, Turkey
iozturk@uludag.edu.tr

ABSTRACT
Pragmatic competence is among the explicitly acknowledged sub-competences that make the communicative competence in any language (Bachman & Palmer, 1996; Council of Europe, 2001). Within the notion of pragmatic competence itself, "implicature (implied meanings)" comes to the fore as one of the five main areas there (Levinson, 1983). In this regard, the present study developed an online multiple-choice discourse completion test (MDCT) to specifically measure the interpretation of formulaic, thus teachable implied meanings in English. After the initial version produced with the help of a native speaker colleague, the test was piloted rigorously with different comparison groups. At the end of the process, 112 EFL Teacher Trainees, 33 native speakers of English, 37 EFL learners at a university School of Foreign Languages and 11 high school EFL students had taken the test as the comparison subgroups with varying proficiency levels. Moreover, seven ELT professional native speakers had been interviewed about each test item to have their direct feedback on wording and all the revision/improvement alternatives. Consequently, significant performance differences were detected between particular pairs of subgroups, which is a strength of the test as it did reflect the performance variability between the participants from different proficiency levels. The outcome was a MDCT that the native speaker takers reached a good compromise on and is usable both in computerized and pen-and-paper format, which can be used in any ELT or ELT teacher training program concerned to diagnose the students' performance in a major area of pragmatic competence to later take informed instructional decisions.

Keywords: English Language Teaching, Pragmatic Competence, Implicature (Implied Meanings), Testing, Teacher Training

INTRODUCTION

"Pragmatic competence (the ability to process and use language in context)" is an essential constituent of having an overall communicative competence. On the other hand, it is a construct relatively hard to develop in EFL contexts like in Turkey, where teaching practices could be grammar-oriented and chances of processing sufficient authentic input are minimal. This all is worthy of more note as interlocutors have been reported to tend to evaluate pragmatic flaws more severely than grammatical ones (Thomas, 1983; Bardovi-Harlig, 1998; Economidou-Kogetsidis; 2015).

When the focus is shifted onto the research agenda within pragmatics, one can see the reports that the domains other than "speech acts" have been studied to a lesser extent. They cover also "implicatures" though they are likely to prove troublesome for learners to interpret even after prolonged exposure to the target language.

In light of the abovementioned considerations, the aim of this study is to develop an online multiple-choice discourse completion test (MDCT) to investigate specifically the interpretation of implicatures (implied meanings) in English, which add up to an essential constituent of pragmatics (Levinson, 1983). In this regard, the following research questions were formulated to guide the study:

* This article is based on the first author's PhD dissertation titled "The effects of explicit film-based instruction on English as a foreign language teacher trainees' interpretation of implied meanings".

1) Do NSs of English reach a good compromise with their interpretations of the test items?

2) Are the performance differences between particular pairs of test-taker groups attributable to their proficiency differences?

LITERATURE REVIEW

In terms of the historical continuum of competences that an efficient language user would need to have (Chomsky, 1965; Hymes, 1972; Canale & Swain, 1981; Canale, 1983; Bachman & Palmer, 1996; Council of Europe, 2001), years saw the inclusion of pragmatic competence as one of the basic components of overall language ability (Bachman & Palmer, 1996; Council of Europe, 2001, p. 33). In this context, pragmatic assessment is important. The argument is supported in reference to EFL environments like the one in Turkey, about which the related literature reports grammar-oriented language teaching practices, materials and assessment (Özmen, 2012; Uztosun, 2013; Erkmen, 2014). The concern here is that even if such a learning context works in the best way possible, it is still open to question whether pragmatic competence can develop jointly with grammatical competence (Bardovi-Harlig, 1996; Jianda, 2006). In light of the reports like Yu's (2006), which suggests that that language learners may need to better understand pragmatic aspects of the target culture so that they can interpret appropriately what they hear and interact effectively with members of that culture, pragmatic assessment proves to be certainly worth considering.

Pragmatic flaws and communication

In view of the abovementioned points, one could voice doubts about learners' probable communication problems in encounters especially with native speakers (NSs). Like Thomas (1983), Bardovi-Harlig and Dörnyei (1998) and Crandall and Basturkmen (2004), Economidou-Kogetsidis (2015, p. 1-2) appreciates the probable gravity of such problems by noting that a NS could attribute a grammatically competent and fluent non-native speaker's pragmatic failure to impoliteness. Accordingly, she states that when EFL learners move to the target language community, sophisticated pragmatic competence becomes essential since pragmatically inappropriate language can cause pragmatic failure by unintentionally violating social appropriateness.

Pragmatics and Implied Meanings

Within the framework set above, "implied meanings (implicature)" were addressed as the focal point of this study. Besides the significance of pragmatic competence in overall language ability, the principal reason for this choice is the fact that, among its five main areas including "implicature" (Levinson, 1983), the study of pragmatics has focused on speech acts (Aijmer, 2011; Eslami & Mirzaei, 2012; Roever, 2013; Bardovi-Harlig & Shin, 2014; Bella, 2014) and to a lesser extent on implicatures (Roever, 2006; Bardovi-Harlig & Shin, 2014). This would acquire an additional dimension when we consider the fact that implicature is an "unremarkable and ordinary" conversational strategy (Green, 1989, p. 92) used frequently and extensively in daily conversation (Pichastor, 1998; Matsuda, 1999; McTear, 2004). What is more, in terms of pragmatic testing, implicature was made a major constituent of the first pragmatic competence test developed in the field of Applied Linguistics (Roever, 2005).

In this regard, it is considered worthwhile here to touch on the blanket term "implicature" and several related points.

Implicature

It was Grice (1975) who introduced the notion of "implicature" to denote cases in which what is meant is distinct from what is uttered (Davis, 2007). He also categorized the concept into "conventional" and "conversational" implicatures, the latter of which originates from his well-known Principle of Cooperation (Grice, 1975, 1981) and Maxims (quality, quantity, relevance, manner). When an interlocutor deliberately disregards one or more of these maxims, that could well be to go beyond the extent of what s/he utters with the intention of expressing his/her aims indirectly but more effectively that way though. This played a historically important role in pragmatics (Hadi, 2013). It has led to new developments in people's understanding of conversation, in the light of which Bouton (1988, 1994, 1999) was the first researcher to study implicatures in specific relation to pragmatic assessment of comprehension. In conformity with the scholars postulating that it is often difficult for L2 learners to notice how people in a given culture express meaning indirectly (Wolfson, 1989) as they show an inclination for taking utterances at face value (Kasper, 1997), Bouton discovered that the ability of nonnative speakers (NNSs) to interpret implicatures is highly questionable. This later kept being confirmed in studies like Kubota (1995), Lee (2002), Taguchi (2005), Roever (2005) and Rızaoğlu and Yavuz (2017).

Using his research findings as base, Bouton (1994) divided implicatures into two sets: idiosyncratic and formulaic. While the comprehension of the former relies mainly on a shared perception of the context, the latter is based on a formula of some sort, which would be structural, semantic, or pragmatic and crucial to a person's effective

interpretation. Bouton's key finding on formulaic implicatures was that they might prove considerably difficult for NNSs and they are less susceptible to even prolonged exposure effects. Nevertheless, they were very much teachable, which provided the direct inspiration for the implied meanings to be included in this study. To make it a basis for further "interventionist studies" where the effect of a particular instructional treatment on students' acquisition of the targeted pragmatic feature is examined (Kasper, 1999), the idea was to develop a test that could investigate as a pretest, posttest and/or delayed posttest the interpretation of such teachable implied meanings by EFL/ESL learners and teacher trainees, who are supposed to help future EFL/ESL learners to have pragmatic competence besides other language ability areas. In this regard, the following section provides the introduction and discussion of each implied meaning type included in the test.

Implied meanings covered in the present study
To start with, the list below gives the implied meanings included in the instructional phase of this study:

* Pope Questions
* Indirect Criticism
* (Verbal) Irony
* Indirect Refusals
* Topic Change
* Disclosures
* Indirect Requests (Requestive Hints)
* Indirect Advice

It should be mentioned that Pope Questions, Indirect Criticism, Irony, Topic Change, Disclosures and Indirect Refusals had already been included in several other studies (Bouton, 1994, Roever, 2005; Taguchi, 2005). Indirect requests and indirect advice have not been bunched together with the abovementioned implied meanings in any test before. They were included in this study on the basis of a consideration like Verschueren's (2009, p. 9), who observes that Grice's (1975) account of implicatures and Searle's (1975) definition of indirect speech acts are very similar, or Birner's (2013, p. 195), who posits that indirect speech acts are a subtype of conversational implicature.

Considering the aforementioned fact that formulaic implicatures are highly convenient to test and teach, it is also worth emphasizing here that some of the implied meanings included in this study are ones that were reported earlier as formulaic in the related literature. For the rest, which have not been overtly declared as formulaic, the researchers' claim is that some of their variations can be deemed formulaic, or tentatively formulaic at least, thus worth being included in the instructional program and tested in terms of teachability. This was both a novelty and a risk for the present study, but one that is worth taking as the intention was to respond to Bouton's (1994, p. 106) call that we should be alert to implicature types of which we are not fully aware *with an eye to including them in instruction programs*.

What follows is the discussion of the abovementioned points with a focus on each implied meaning covered in the present study:

Pope Questions
Pope Questions are reported to be clearly formulaic (Bouton, 1994; Bouton, 1999) as they always tend to work according to the following pattern:

* One is asked a question.
* S/he thinks that the answer is an obvious "Yes" or "No".
* To answer that question with an indirect but emphasized "Yes" or "No", s/he asks a new question to which the answer is a clear "Yes" or "No".

The following is an example:

> *A mother and her daughter Jenny have been discussing the upcoming weekend. Jenny's parents are leaving town and this is the first time Jenny has been left at home alone.*
>
> **Mother:** Are you sure you can take care of yourself this weekend?
> **Jenny:** Can a duck swim, Mother? (Bouton, 1988, p. 193)

As the answer to Jenny's question is an obvious "Yes", she is telling her mother indirectly that she will of course be able to take care of herself okay (Bouton, 1988, p. 193).

Indirect criticism

Like Pope Questions, utterances that contain Indirect Criticism (also called "Understated Negative Evaluation" or "Damning with Faint Praise") are reported to be formulaic (Bouton, 1994, 1999). It happens when we are asked what we think of something or someone that we, in fact, do not like - but we don't want to say so explicitly. Instead, we reply indirectly, commenting about features of the thing that are not central to its evaluation in any way (Bouton, 1988, p. 193). See the following example:

> *George and Sheila are looking for a house to buy. Sheila just went to look at a house in their price range and is reporting back to George.*
>
> **George:** So, what did you think of the house?
> **Sheila:** Well, it had a nice mailbox. (Broersma, 1994, p. 3)

As Sheila responds with a praising comment on just a subsidiary feature of the house (the mailbox), she could be interpreted to imply that some more important aspects of the house merit considerable criticism. By praising the house in such a weak way, she makes it obvious that she does not really admire the features that are central to the evaluation she has been asked to make. In other words, she criticizes the whole through a slight compliment to a part.

(Verbal) Irony

Ironic utterances are a type of implied meanings that are considered in the same category with formulaic implicatures (Bouton, 1994, p. 105), thus eligible to be the focus of pragmatic instruction. Especially verbal irony, which is of interest to the present study, can be deemed as based on a particular semantic pattern. That is, an ironic statement must be contrary to the true state of affairs to be interpreted correctly. There must be some discrepancy between the reality and the utterance, and the listener must recognize this discrepancy in order to interpret the utterance (Kreuz & Roberts, 1995, p. 22). To put it differently, the speaker uses words that mean the opposite of what s/he really thinks. This is exemplified below:

> *Joan and Anne are classmates. Joan has some problems reading his paper and he is asking Anne for help.*
>
> **Joan:** Hi, Anne.
> **Anne:** Hi Joan. What's up?
> **Joan:** I was wondering if I could ask a small favor of you. Would you read my Linguistics 441 paper?
> **Anne:** Gosh, John, I wish I could, but I promised Jack I'd go bowling with him tonight.
> **Joan:** Yeah. Well, thanks for the help. (Bouton, 1994, p. 101)

We see that after being refused by Anne, Joan's latest remark suggests that he is thankful for the response. However, on second thought if necessary, one could see that the statement is contrary to the true state of affairs and there is a discrepancy between the reality and the utterance. That is, Joan feels dissatisfied with Anne's response and he means to express it with a sarcastic remark.

Indirect refusals

Indirect refusals can be viewed as another type of formulaic implied meanings in the light of the pertinent literature. They are defined as routinized expressions reflecting relatively fixed patterns of discourse exchange (e.g., giving an excuse when refusing) (Taguchi, 2007, p. 329). What is more, they are cited as notably appropriate for classroom instruction of pragmatic comprehension with their abovementioned conventional features (Taguchi, 2007, p. 331). See the following example:

> **Mary:** Hey, John, what're you doing?
> **John:** I'm working on my paper for the English class.
> **Mary:** You've been working on that paper for a week. Why don't you take a break? Let's go to the movies tonight.
> **John:** I have to finish my paper by eight in the morning. (Taguchi, 2007, p. 322)

It is seen that John does not refuse Mary's offer with explicit linguistic markers of refusals such as "I can't", "No", or "I don't want to", which were identified as direct refusal expressions by Beebe, Takahashi, and Uliss-Weltz (1990) (as cited in Taguchi, 2007, p. 321). Instead, he provides his reason/excuse for not accepting the offer.

Topic change

Topic Change (Change Subject) is another type of implied meanings which has been reported to be formulaic (Roever, 2011, p. 466). In terms of the relevant body of research (Bouton, 1988, p. 190; Roever, 2005), it happens when a person feels that a current line of discussion is really inappropriate and leaps into another topic. In other words, it occurs when one comes up with an irrelevant, unexpected utterance as s/he does not like what has just been said or asked. The purpose can be considered to be making the inappropriateness perceptible to the interlocutor(s). See the following example from "The Prince of Tides (1991)", a romantic drama film based on the 1986 novel of the same name by Pat Conroy:

> *Susan, a psychiatrist in New York, is questioning Tom, a football coach from South Carolina burdened with many details of his dysfunctional family's secrets. Susan decides to discuss the topic of his sister Savannah's last suicide attempt after their brother Luke's death. This is one of the initial meetings between Susan and Tom. Therefore, Tom is reluctant to disclose some certain facts about his family.*
>
> **Susan:** Savannah's last suicide attempt was right after his death, correct?
> **Tom:** Yeah, she had a few bad days over it.
> **Susan:** Were there other times?
> **Tom:** I don't know. There might have been another time when we were young, but I'm not sure . . . How are you getting paid?
> **Susan:** Why change the subject?

As is seen, Tom does not seem to like the turn that the conversation takes. Instead of satisfactorily answering Susan's query, he chooses to ask an irrelevant question at that moment of the talk. We also see that Susan does not fail to understand Tom's attempt to change the subject, which is uncomfortable from his own viewpoint.

Disclosures

Another type of implied meanings covered in the study is Disclosures, which are defined as indirect replies used to avoid disclosing embarrassing information (Taguchi, 2002, p. 157). To the best of the researcher's knowledge, the implied meaning type of Disclosures has not yet been openly reported as formulaic, routinized, homogeneous or predictable. Nevertheless, the researcher still thought that they could be teachable, thus suitable for being included in the study. The rationale was that the definitions and examples in the relevant studies (Taguchi, 2002; Taguchi, 2005) can be considered to contain some semantic clues that point to a tentatively identifiable pattern: When one is questioned about the reality of something and when the answer would urge him/her to give embarrassing or disturbing information from his/her own viewpoint, s/he might not make a full confession. Instead, s/he might just give the reason(s) why the consequence (to be mentioned in a direct answer of confession) really arose or not. Doing that, s/he could produce an indirect answer of revelation, confirmation or negation about the reality that is being questioned. See the following example:

> **Jim:** Hi Mom, I'm home.
> **Mom:** Hi Jim. Didn't you get the report card today? How were your grades this semester?
> **Jim:** You know mom, I don't think the teacher grades fairly. (Taguchi, 2002, p. 171)

We see that Jim does not respond to his mother's question with a direct answer of confession that his grades were poor, which appears to be an item of too embarrassing information for Jim to disclose directly. Instead, he just gives the reason in his opinion (the fact that the teacher does not grade fairly) why the grades in the report card were low. In that way, he indirectly makes the revelation that the reality is his poor grades. It can be thought that the answer is intended to function also as a call for understanding and empathy.

Another example of Disclosures provided below is from the American sitcom "Friends (1994)":

> *Monica, the mother hen in her group of friends and a chef known for her perfectionist, bossy and competitive nature (Retrieved on July 21, 2016 from http://en.wikipedia.org/wiki/Friends#Characters), is trying to organize a big special dinner. She is in search of a waitress for it. Rachel, Monica's best friend from childhood and a waitress herself, infers from Monica's telephone conversation that she has arranged for another waitress to serve in the organization.*
> [Phone rings. Monica answers it.]

Monica: [on phone] Hello? Oh, hi Wendy! Yeah, eight o'clock. What did we say? Ten dollars an hour?... OK, great. All right, I'll see you then. Bye. [hangs up]
Rachel: Ten dollars an hour for what?
Monica: Oh, I asked one of the waitresses at work if she'd help me out.
Rachel: [hurt] Waitressing?
Joey: Uh-oh.
Monica: Well... of course I thought of you! But... but...
Rachel: But, but?
Monica: But, you see, it's just... this night has to go just perfect, you know? And, well, Wendy's more of a... professional waitress.
Rachel: Oh! I see…

As we see, Rachel questions Monica so that she states the obvious fact for Rachel that she was not hired as the waitress to help Monica out. Instead of a direct response in the affirmative or negative, Monica just gives the reason (the fact that the hired waitress is more professional than Rachel is) why she did not pick Rachel. In that way, Monica indirectly makes the confirmation that she did choose another waitress, which seems to be an item of embarrassing information hard for Monica to disclose directly. It can be thought that her reply is also an attempt to justify her decision and a call for understanding.

Indirect requests (Requestive hints)
Another type of implied meanings included in this study is Indirect Requests, which were labeled as Requestive Hints (Rinnert & Kobayashi, 1999; Weizman, 1985, 1989, 1993) in the pertinent literature.

Considering the pursuit in this study of some formulaic implied meanings with several clues that point to a particular pattern, the fact must be acknowledged here that Requestive Hints have never been reported to be formulaic, routinized or homogeneous. They represent a heterogeneous category which includes various sub-strategies (Weizman, 2007, p. 144). According to the model that Weizman (1985, 1989, 1993) posits, requestive hints should be considered in two dimensions, which are the propositional and illocutionary meaning of the request.

The first dimension, 'propositional content' of the request, contains 3 categories: (1) *zero* (no reference to the hearer, the act or any of its components, e.g., 'There's a problem'), (2) *component* (reference to some component of the requested act, e.g. 'Are there any batteries?'), and (3) *act* (reference to the requested act, including some or all of its components, e.g. 'The sign to change the master [for the duplicating machine] came on but ...'.

The second dimension, 'illocutionary device', contains 4 categories: (1) *zero* (no statement of illocutionary intent, e.g., 'Here's the mail' as a request to take the mail to the mailroom); (2) *stating potential grounder* (giving a reason why the request is necessary, e.g., 'The printer is running out of ink'); (3) *questioning feasibility* (asking about some prerequisite for the request to be granted, e.g., 'Do you have any chalk? '); and (4) *other* (illocutionary device not falling into one of the three preceding categories, e.g., 'I'm going to borrow this pen'). (Rinnert & Kobayashi, 1999, p. 1188)

In the heterogeneity of implied requests as itemized above, it was out of question for this study to be aimed at measuring the comprehension of and teaching about all the reported categories. However, considering the facts that speech acts have commanded a good deal of attention in pragmatics research (Cohen, 2012a, p. 33) and requesting is one of the especially popular speech acts (Cohen, 2012b, p. 280) in terms of instructed pragmatics as well (Taguchi, 2015, p. 5), the researcher had the intention to include the requests in the study since the very beginning.

In that regard, the decision to be made was which category of the aforementioned indirect requests would be integrated into the data collection instrument and instruction program of this study. The choice was the ones that are based on "stating potential grounder (giving a reason why the request is necessary)". The primary basis for that decision was the fact that they were found to be the most frequent English hints on the illocutionary scale (47.2%) in the naturally occurring data in Rinnert and Kobayashi (1999, p. 1189). What is more, when the analyses of the occurrences on the propositional and illocutionary scales were combined, English speakers' most frequent strategy was "potential grounder" added "component" (30.6%) (Rinnert & Kobayashi, 1999, p. 1189).

Besides the research findings on their frequency as mentioned above, the other basis for the inclusion of "requestive hints by stating potential grounder" was the fact that the way they are reported to occur sounds fairly clear: giving a reason why the request is necessary (Rinnert & Kobayashi, 1999, p. 1188). This was thought to be compatible with the present study's principle of including formulaic, routinized or predictable, thus teachable

implied meanings. For an example of the requestive hints mentioned so far and integrated into the study, see the situation and dialogue below taken from the sitcom "Friends (1994)":

> *Monica, a chef, is trying to finish the job of preparing enough food for a special meeting. She has figured out that it will not be possible for her to complete the preparation in time as she did not schedule things properly. While cooking, she is talking about the situation to her housemate Rachel.*
>
> **Monica:** Anyway, see, I planned everything really well. I planned and I planned and I planned. It just turns out, I don't think I planned enough time to actually do it…
> **Rachel:** Hey, Mon, you want some help?
> **Monica:** If you want.

As is seen, Monica does not use a direct statement of request like "(Could/Can you) please help me finish cooking. (?)" Instead, while cooking hastily at the same time, she just indicates the problem (Taguchi, 2005, p. 549), which is the reason why a request for help is necessary. In that way, she makes the requestive hint that her housemate Rachel cooks with her so that she can finish the job in time. Considering Rachel's offer of help that follows Monica's words, we also see that it does not take long at all for Rachel to get the hint. Monica jumping at the offer confirms the fact that her initial words were meant to function as an implied request (requestive hint) for Rachel's help.

Indirect advice

The last type of implied meanings included in this study is Indirect Advice (Matsumura, 2001; 2007), which is explained as "indirect comments with no advice" (Matsumura, 2001, p. 646) where the speaker's intentions are not made explicit (Brown & Levinson, 1987; Levinson, 1983). As is the case with Disclosures and Indirect Requests, to the best of the researcher's knowledge, indirect pieces of advice have not yet been openly reported as formulaic, routinized, homogeneous or predictable. Nevertheless, examining the advice-giving scenarios and options in Matsumura's (2001, p. 676; 2007, p. 187) multiple-choice questionnaire adapted from Hinkel (1997), the researcher believed that at least a certain way of indirect advice-giving could be teachable, thus suitable for being included in this study. The rationale was that the advice-giving way in question can be considered to contain some semantic clues that point to a tentatively identifiable pattern: Without using forms like "should (not), had better (not) etc", one offers the advice indirectly by just giving a/the reason why the hearer should or should not do the thing which is the subject of the advice. See the following example modified and adapted from Matsumura (2001, p. 679; 2007, p. 190), which was not used in the data collection instrument or instruction program of this study as it would require knowledge about the Canadian cities Banff and Vancouver, especially the distance between them:

> *You have just heard from your supervisor that s/he is considering a trip to Banff from Vancouver in a car which breaks down frequently. You think it would be appropriate to say:*
> "Taking such a long trip in this car may be risky."

As shown here, without employing some well-known advice-giving patterns like "should (not), had better (not) etc", the speaker offers his/her advice by just giving the reason why the hearer should not take the car for such a long trip and that reason is the fact that doing it might be risky. The logic here could be likened to what is typically done by people who work as financial advisors. As is known, they often set forth a good number of reasons to buy or sell some particular financial instruments. However, with the concern that their statements could be interpreted as sound advice likely to burden them with responsibility for any loss of addressees, they use warning notices like the following:

> This document is for information and illustrative purposes only and does not purport to show actual results. It is not, and should not be regarded as investment advice or as a recommendation regarding any particular security or course of action (Retrieved on July 7, 2016 from http://www.nisa.com/psrx-disclaimer/).

The strategy described above can be claimed to be one of the advice-giving options in most of the scenarios in Matsumura's (2001, p. 676; 2007, p. 187) multiple-choice questionnaire. The only exception can be considered to be the scenario about a broken vending machine from which people cannot get a pop or the money back (Matsumura, 2001, p. 677; 2007, p. 188).

Besides viewing it as formulaic and teachable because it contains some semantic clues that point to a particular pattern, there were two more reasons why the above-discussed advice giving way was included in this study as modified from Matsumura's (2001, 2007) scenarios.

First, Matsumura (2001) reported that offering indirect advice was a strategy favored to a considerable extent by his native speaker participants. When we exclude that "broken vending machine" item from his four-option multiple-choice questionnaire, for the three scenarios where advice is to be offered to a higher status person, "Indirect" was the native speakers' most frequent choice in one of the scenarios and the second most in the other two. For the three scenarios where advice is to be offered to an equal status person, "Indirect" was the native speakers' most frequent choice in again one of the scenarios and the second and third most in the other two. For the three scenarios where advice is to be offered to a lower status person, "Indirect" was the native speakers' most frequent choice in two of the scenarios and the third most in the other one.

It is worth noting here that the lower status hearers were 1st-year university students addressed by higher-year university students, about which Matsumura (2001, p. 645; 2007, p. 172) asserts that it is a part of an existing Japanese hierarchical system where 2nd- and 3rd-year students are considered to be *senpai*, that is, to be in a higher status than 1st-year students, and according to this hierarchy, 1st-year students normally use polite expressions when talking to *senpai*. Considering the status relationships from the viewpoint of the 1st- and higher-year university students in Turkey, where it is impossible under normal conditions to talk about such a hierarchy and titles like "senpai", the higher status speaker-lower status hearer interactions in Matsumura (2001, 2007) were adapted to this study as scenarios where individuals of relatively equal statuses interact.

The second reason why indirect advice was included in this study as modified from Matsumura (2001, 2007) concerns the consideration given in the ending of the preceding paragraph. For the data collection instrument of this study, all the test items were based on either conversations that take place between people of relatively equal statuses or utterances that speakers make to themselves. Similar to what Roever (2005) did for the speech acts section of his pragmatic assessment battery, the aim was to keep the social distance and power differential relatively low so that the participants' comprehension performance of implied meanings would be measured under as controlled contextual parameters as possible. In this regard, the higher status speaker-lower status hearer interactions in Matsumura (2001, 2007) were included in the data collection instrument of this study as scenarios where people of relatively equal statuses interact.

Fillers

Like in Taguchi (2005), in addition to the item types described so far, a certain number of filler items that tested literal comprehension were included in the test. They dealt with basic, direct interpretation. They were excluded from the analyses.

The filler items were meant to deflect the participants' attention from the true purpose of the test, which is to investigate how test takers comprehend the nonliteral meanings. If the test had been composed of only implied meaning items, participants who discover it after answering some initial questions could stop examining the rest and continue by just searching for the response options that give indirect interpretation.

DEVELOPMENT AND DESIGN OF THE TEST

To the best of the researchers' knowledge, no study has so far attempted to measure specifically the interpretation of formulaic and thus teachable implied meanings. For this purpose, the previous studies on pragmatic interpretation (Bouton, 1988, 1992, 1999; Roever, 2005; Taguchi, 2002, 2005; Rızaoğlu & Yavuz, 2017) were all examined. On their basis, an online MDCT was decided to be developed following the steps detailed below.

Theoretical background to the data collection instrument

First of all, a certain number of scenarios (situations) that contain the target implied meanings had to be determined. Those scenarios were supposed to provide the basis on which the test's questions and response options would be built. To that end, all of the related studies providing appropriate scenarios were examined to be adapted to this study. The table below shows the final numbers of the test items in each group of target implied meanings and the studies from which they were adapted.

Table 1: *The Numbers of the Test Items in Each Group of Implied Meanings and their Sources*

Implied Meaning	Number of Test Items	Source
Pope Questions	5	(Bouton, 1994)
Indirect Criticism	4	(Bouton, 1994; Kubota, 1995)
Indirect Advice	4	(Matsumura, 2001; 2007)
Topic Change	4	(Roever, 2005)
Disclosure	3	(Taguchi, 2005)
Irony	3	(Colston & O'Brien, 2000)
Indirect Refusals	3	(Taguchi, 2005)
Indirect Requests	2	(Rinnert and Kobayashi, 1999)
Filler Items	5	(Taguchi, 2005)

Modification of the language in the test items

After deciding on the initial versions of the scenarios and ensuing talks in the test items, a (British) native speaker of English, who is a colleague with 25 years' experience in foreign language teaching and EFL teacher training, was asked to proofread them all. This step was seen strictly necessary as the researcher had tried to shorten and/or simplify the language of all the scenarios and utterances. The aim was to minimize the effect of language proficiency nuances between participants so that the validity of the test could be enhanced to primarily measure the construct of implied meaning comprehension.

This simplification and modification procedure was inspired by the way Roever (2005, p. 46) standardized, shortened and simplified the items he had adapted from Bouton's (1988, 1994, 1999) test. The procedure employed in this study will be described later in detail. Taguchi (2005, p. 550) as well tried to reduce the effect of some construct-irrelevant factors in the implied meaning comprehension measurement to be done by her test. In order to minimize the variance from her learners' difference in vocabulary knowledge for example, all vocabulary in her items was drawn from Longman's 2,000-word defining vocabulary list (Longman, 1995). The 2,000 words in question are identified as common and basic English words, which makes them the ones chosen to write all the word definitions in the Longman dictionary. Accordingly, the 2,000-word-level vocabulary items were considered to be relatively attainable by her L2 participants. Besides that, Taguchi (2005, p. 550) took great care to keep the lengths of all of the conversations in her test approximately the same. Her aim was to control the burden on short-term memory. Moreover, she used equal number of words in her question and option sentences across item categories so that the effect of some irrelevant variables like reading time could be lessened.

After the abovementioned colleague proofread the initial versions of this study's test items, a meeting was held with him to discuss the alteration and revision suggestions that he had come up with. In that meeting, almost all the items were refined to varying extents in terms of grammar and some word choices. With the refined versions of the test items at hand, the next step was writing response options for each item so that the instrument could serve as a multiple-choice test.

Writing the response options for the test items

Appointing the correct answers in the multiple-choice test developed for this study was fairly easy. The favored responses in the studies that the items had been adapted from were already self-evident.

Selecting some of the incorrect responses was made through adoption with no or minor change(s) from the studies that the items had been borrowed from. For the rest, the present study drew on a synthesis of three methods employed in the related literature to write response options for multiple-choice tests designed to investigate implied meanings comprehension. As mentioned above, there were a certain number of test items adapted already with some ready-made response choices. In this regard, the synthesis of the methods in question served also as a step to converting the other borrowed items (originally with no response options) into multiple-choice test items.

Bouton's (1988) method was the first to be manipulated for that synthesis. It called for having nonnative speakers of English respond to the item stems and then using their responses different from the favored ones as distractors. To that end, the first step was dividing the total number of the test items into three even groups. After that, they were printed on three separate handout forms and administered to three different EFL teacher trainee groups of 60 people. They were students who had enrolled in the summer school courses of Uludag University ELT Department, and they did not participate in any further phase of the study. They were asked to respond to each item, which

consisted of a brief description of the situation, the utterance(s) and an open ended question that reads: *"What does (the last speaker's name) probably mean?"* Below is an example:

> *Maria and Frank are working on a class project together but they won't be able to finish it by the deadline.*
> ***Maria:*** *"Do you think Dr. Gibson is going to lower our grade if we hand it in late?"*
> ***Frank:*** *"Do fish swim?"*
> *What does Frank probably mean?*

As explicated before in the section devoted to "Pope Questions" in the Literature Review, the favored interpretation for the item above would be something like "he (Dr. Gibson) will of course lower our grade if we do that." Accordingly, the teacher trainees' responses that differed from such an interpretation were all recorded as the distractor alternatives for the item. An example to the erroneous interpretations was interestingly in reference to a well-known saying in Turkish where the central figure is a fish: "Battı balık yan gider." Within the context of the item above, it can be interpreted to mean something like "As we do not seem to have any other chance, let's just take the risk and hand in the project late to face the consequences." This procedure was followed for each item, and the first group of distractor alternatives was thereby obtained. Like in many other ways possible to write good multiple-choice items, this was also a response to the calls like Haladyna and Downing's (1989) for common errors of students to be incorporated in distractors.

For the second group of distractor alternatives, Taguchi's (2005, p. 550) principles for distractor writing were considered. They are as follows:

> *** Principle 1:** The option contains a meaning that is the opposite of the implied meaning.
> *** Principle 2:** The option contains words taken from the last part of the dialogue.
> *** Principle 3:** The option is related to the overall conversation.

For each item, the researcher tried to apply all the above-mentioned principles. However, just as the impossibility that Taguchi (2005, p. 560) encountered herself, it was not possible to follow all the three distractor principles for all the items. An example reason is the fact that, when the last utterance in a dialogue was extremely short containing only a few words, it was difficult to write a distractor following the second principle, "taking words from the last utterance" (Taguchi, 2005, p. 560). Nonetheless, the procedure did contribute to the pool of distractor alternatives for almost all the items in the present study. For example, two of the distractors in the item below were provided by this procedure:

> *Roger is thinking of taking his car to a repair shop in the city centre. His friend Melanie knows that the shop is known for doing careless work.*
>
> **Melanie:** "I don't usually take my car there. It has a really bad reputation."
> What does Melanie probably mean?
> - Roger should take his car there for only small repairs.
> - She advises Roger not to take his car to that repair shop. (the favored response)
> - The reputation of a place is important. (the one based on principle 2)
> - Roger can take his car there. (the one based on principle 1)

As mentioned before, this study drew on a synthesis of three methods to develop the response options to be counted as the distractors. Accordingly, for the third group of distractor alternatives, Roever's (2005) viewpoint was employed. Despite finding it intuitively appealing, Roever viewed Bouton's aforementioned procedure for item design as questionable. His postulation was that incorrect response choices produced by nonnative speakers do not guarantee unambiguous, good distractors (Roever, 2005, p. 46), which is reported by Hudson, Detmer and Brown (1995) as well. In this regard, exercising his own judgment, Roever built new distractors wherever he found the ones borrowed from Bouton ambiguous. This procedure was followed in this study too when the distractors developed with the two aforementioned methods were considered quantitatively or qualitatively inadequate. Below is an item to exemplify how it was done. In it, all the three distractors were written with the researcher's own judgment as the ones produced with the two other procedures had not been considered unambiguous or challenging enough:

Susan and Tom, friends, are talking about what is going on in their lives. Susan knows Tom had a job interview recently.

Susan: 'So how was your interview? Did you get the job you applied for?'
Tom: 'Um . . I think I need to improve my interview skills.'
What does Tom probably mean?

- He did not get the job. *(the favored response)*
- He wants help from Susan to improve his interview skills.
- He will have the interview when he feels his interview skills are good enough.
- They gave him the job with the advice that he should improve his interview skills.

Conversion of the data collection instrument into a web-based test
After developing the initial version of the test items with the principles and procedures described above, the next step was to create a web-based MDCT out of it. The following two sections give some fundamental aspects of that web-based MDCT.

Technical aspects of the test
First of all, a professional computer programmer was paid to cooperate. Keeping in close touch with the researcher before and during the development of the system, he designed the test as a web-based one that should run on any common web browser. He wrote the codes in a way that the system would control item delivery, scoring, data storage and all other functionality.

Content aspects of the test
Each test item had the same format and elicited what a character in the item stem probably means with his or her utterance. What a test taker would say or mean in the situation was not elicited in any way, which is common in pragmatics research instruments (Roever, 2005, p. 45). The idea was to investigate test takers' interpretation of the included implied meanings rather than their favor or disfavor of particular conversational strategies.

In a similar way to Roever's (2005) test on implicatures, all the items were standardized with the details as explained below:

1. The response choices counted as "correct" for the items included in the final analyses were designed to occur as equally frequent as possible in all response option positions. The aim was to ensure that systematic guessing by test takers would lead to only chance-level correctness.

2. All the characters in the items have names and all are introduced in the item stems. This is intended to be an improvement to Bouton's (1988, 1994, 1999) items with generic descriptions like "two friends", "two teachers" etc. and to some scenarios adapted from studies like Matsumura's (2001, 2007) where the characters are given false names like "P.D", "C.J", "X.L".

3. Except for two of them, every item is based on a conversational situation where a male interlocutor addresses a female one or vice versa. This is for the sake of gender balancing and aimed to be an improvement to male-male or female-female items.

One of those two exceptional items includes an ironic utterance that the male speaker makes to himself like muttering. Its original version in Colston and O'Brien (2000) was already that way. Besides that, the researcher and the assisting native speaker colleagues could not figure out a way to add a female interlocutor in the situation without making the item sound unnatural. Below are the original version of the item stem and the related ironic utterance (Colston & O'Brien, 2000, p. 1581):

Henri was an avid cyclist and was eagerly awaiting a new, very expensive, high tech bicycle he had ordered from this new company. When it finally arrived, it turned out to be really heavy and poorly constructed. When Henri saw that he was cheated by the bike company, he said,

"This company is incredibly honest."

The other exceptional item includes a scenario with indirect criticism (damning with faint praise) and its original version takes place between two female characters as given below (Bouton, 1988, p. 194):

Brenda and Sally have lunch every Tuesday. As they meet on this particular day, Brenda stops, twirls like a fashion model, and the following dialogue occurs:
Brenda: I just got a new dress. How do you like it?
Sally: Well, there certainly are a lot of women wearing it this year. When did you get it?
How does Sally like Brenda's new dress?

For the initial version of the test used in the first pilot study, it was one of the items that were included in the attempt to achieve gender balancing. It was modified so that it took place between one male and one female speaker. The result is provided below:

Brenda is waiting for her boyfriend Jim at a cafe for lunch. When he comes to the table, Brenda stands up, and twirls like a fashion model, smiling.
Brenda: I just got a new dress. How do you like it?
Jim: Well . . . there certainly are a lot of women wearing it this year. When did you get it?
What does Jim probably mean?

In the period between the first pilot study and the main study, six native speakers were interviewed about each item. The consensus emerged between them on the fact that the item would sound much more natural if the dialogue happened between two female characters, which is the situation in its original version anyway. One native speaker even objected that Jim sounded homosexual in the way the item was modified as shown above. Therefore, in the versions of the test used after the first pilot study, the item was re-modified so that the conversation occurred between two female characters and with a rephrased answer to the first character's question.

4. The question before each set of response options is always in the same format, which was adopted from Roever (2005): "What does NAME of the SPEAKER probably mean?" This is intended to be an improvement to Bouton's (1988, 1994, 1999) items using different questions for different items like "Which of the following best says what Bill meant?", "Which of the following is the closest to what the friend meant by this remark?"

This standardization served also as another step to the conversion into multiple-choice test items of some adapted scenarios originally with no question and/or response options.

Vocabulary explanations in the test
In order to minimize the effects of vocabulary knowledge differences between the participants, all the salient vocabulary items were displayed as underlined on the computer screen. Whenever a test taker positioned his/her cursor on any of them, the related definition from Cambridge Learner's Online Dictionary (reference) automatically appeared. Most of the words underlined for this functionality were determined as early as when the test items were administered to teacher trainee groups in the open-ended format previously explained. Before, during and after responding to the items, the teacher trainees were systematically encouraged to ask about any lexical units that posed a problem for them. Every query of theirs was noted down so that the decision could later be made on the vocabulary that required the incorporation of explanations from Cambridge Learner's Dictionary. Besides that, the researcher included some other vocabulary items that he considered salient even though they had not been queried by the teacher trainees.

Having developed the initial version of the web-based MDCT with the procedures and aspects described above, the following step for the researchers was to conduct the pilot study to refine the test.

PILOT STUDY
Following Roever (2005) the pilot study in this research was carried out with different groups at different times.

The first group consisted of two subgroups: 69 first year EFL teacher trainees at Uludag University and 13 Turkish citizens (all over the age of 30) who had been schooled and lived in an English-speaking country for between 9 and 36 years. The first subgroup was meant to work like Roever's EFL group of target proficiency while the second one was intended to function as his English as a Second Language (ESL) group.

The second group was comprised of 23 EFL learners at the School of Foreign Languages at Uludag University. They had all been ranked at beginner/elementary level a year earlier by the university's official placement test. They participated in this study after a year's intensive EFL instruction given to put them at a level relatively higher than intermediate. They were meant to function as Roever's EFL group below target proficiency.

The third group was 12 native speakers of English (5 American, 4 British, 1 Canadian, 1 Australian and 1 South African). They were intended to work like Roever's group above target proficiency. Besides that, as seven of them (5 American, 1 British and 1 Australian) were later interviewed one by one about each test item, they functioned also like Roever's native speaker participants that produced verbal protocols.

The aim of collecting pilot study data from such distinctly different groups was to cross-validate the decisions to refine and improve the test. In this regard, the results provided by especially the EFL teacher trainees, ESL group members and native speakers were examined to determine the test items with malfunctioning response options. The distractors that had not been chosen by any of the EFL teacher trainees and ESL group members were identified as in need of alteration. Additionally, the common items with the lowest item-total correlations for all the three groups were categorized as in need of revision or complete replacement.

The scores of the EFL teacher trainees were separately considered to identify the general suitability of the test for the target proficiency group. The test proved relatively suitable, with test takers scoring on average 52.95%. The scores of the EFL learners at the School of Foreign Languages were used in a comparison with those of the ESL group members. The objective was to have more data on the evaluation of general suitability and item revision. The expected great variability between the groups did arise in the test scores: EFL learners scored 29.19% while ESL group members scored 73.90%.

The results seemed promising in that the test proved generally suitable for the EFL teacher trainees, who could also be viewed as relatively high-level language learners to whom implied meanings can be taught. In addition, the test reflected the variability between the relatively low and higher proficiency groups. Nonetheless, as they would serve also as the referent group for the favored responses in the test, the average performances of the ESL group and native speakers were relatively unsatisfactory with 73.90% and 72.91% respectively. Besides, there were items with some particular response options chosen saliently less or more frequently than expected. The doubts arising were resolved when seven native speakers were interviewed about the test items. Their comments that overlapped with each other led to a considerable number of rightful changes in terms of the points laid below:

* The wording of the situations in some item stems was revised. This added clarification to the contexts in which the implied meanings occur.
* The distractors were altered or replaced when any of them was interpreted as not clearly enough correct or incorrect. This decreased the number of the ambiguous items where two or more response options were likely to be picked as the favored option.
* Several revisions were made in the grammar and/or word choices of some items. This was to help them sound more native speaker-like.

The most important result of the debriefing sessions with the seven native speakers was that a second pilot study was decided to be conducted. This was not planned at all at the beginning of the larger research project where this study is the initial step as the test development phase for further data collection. From this point of view, while the second pilot study did serve as the second piloting stage of the larger research project in question, it was the final main phase of the present study aimed at developing an online MDCT to investigate formulaic implicatures comprehension.

MAIN STUDY

Before carrying out this phase, four (three American, one British) of the native speakers who had contributed in the previous debriefing sessions were interviewed again one by one. Their common point was that they were all trained and experienced in the field of language teaching (in Turkey as well). Before the talks, a considerable number of revisions and alterations were already done according to the data gathered in the previous phase. Moreover, thanks to the help of a friend of one of the researcher's, a new native speaker group of 14 people at the physics department of an American University had taken the revised version of the test, with five of them providing also their direct feedback on wording and some alternative distractors.

Eventually, in the printouts prepared for each one of the four abovementioned ELT professionals, beneath the revised version of every test item, there were also the alternative revision ideas inspired by the debriefing sessions of the first pilot study and the contributions of the additional 14 American test-takers. In this way, the ELT professional native speakers, who were assisting the researcher face-to-face, were provided the favored revisions together with their alternatives so that they were able to discuss the most appropriate changes by taking account of all the options that had accumulated.

The new version of the test was developed in consideration of these four native speakers' paralleling views on the revision alternatives. The new ideas that came up during the talk with any one of them were later shared with the others via emails, and compromise was sought. Some information is provided below to exemplify how a considerable number of items evolved to varying extents through the painstaking stages of the test development procedure explained so far. First, the sample item is given in the way it was in its source (Colston & O'Brien, 2000, p. 1581):

> *Henri was an avid cyclist and was eagerly awaiting a new, very expensive, high tech bicycle he had ordered from this new company. When it finally arrived, it turned out to be really heavy and poorly constructed. When Henri saw that he was cheated by the bike company, he said,*
>
> This company is a tiny bit sneaky. (UNDERSTATEMENT)
> This company totally stole my money. (LITERAL)
> This company is incredibly honest. (VERBAL IRONY)
>
> What follows is its final version used in this study:
>
> *Henry loves cycling. He orders a new, very expensive bicycle from a new bicycle company. When it arrives, he sees that it is really heavy and does not look well-made at all.*
>
> **Henry:** "Wow, this company's really honest."
> *What does Henry probably mean?*
>
> - The company is dishonest.
> - The company is a bit sneaky.
> - The company is really honest.
> - It is normal as the company is new.

As illustrated above, apart from the abridgement and simplification work, the test items sometimes needed to be added characters, a question and proper response options. In addition, as Taguchi (2005, p. 549) did for the dialogues in her study, linguistic units that characterize the interactive nature of spoken English, such as discourse markers (e.g., *well, you know*), interjections (e.g., *oh*), or hesitation markers (e.g., *um*; see Biber, Johansson, Leech, Conrad, & Finegan, 1999) were included in as many utterances as possible with the help of the assisting native speakers.

Participants

Consequently, the new test with the finally decided changes were administered online to

* 43 EFL Teacher Trainees at Uludag University (10 to 11 students from 1st, 2nd, 3rd and 4th graders each),
* 21 native speakers of English (13 American, 3 British, 2 Australian, 2 Canadian, 1 New Zealander),
* 14 EFL learners at the School of Foreign Languages at Uludag University, who were ranked at pre-intermediate level four months earlier by the university's official placement test and participated in this study after a three and a half months' intensive EFL instruction,
* 11 high school students, who were grouped with regard to their previous achievements in EFL and were getting a language intensive education to enroll for such university programs as ELT, English Language and Literature, Translation and Interpreting Studies.

RESULTS AND DISCUSSION

The data were analyzed with SPSS 22. The Cronbach Alpha's Reliability Coefficient for the EFL teacher trainees (both the target and biggest group) was calculated as ".777", which can be considered acceptably high.

To see if there were any significant differences between the four participant groups, i.e. EFL teacher trainees (henceforth EFLTTs), EFL Learners at the School of Foreign Languages (EFLLs), native speakers of English (NSs) and High School Students (HSSs), one-way ANOVA was performed. As the homogeneity of the variances of groups (Levene's test) was not satisfied ($p<0.01$), non-parametric tests (Kruskal Wallis) were conducted. The tests showed significant differences among the groups investigated: $\chi^2 = 54.589$, $p<0.01$. The results are given in Tables 2-7.

In order to see if there were significant differences between the specific pairs of participant groups, Mann-Whitney pair-wise comparisons were performed. Table 2 shows the results comparing EFLTTs with EFLLs.

Table 2: *Mann-Whitney Pair-wise Comparisons between the Teacher Trainees and School of Foreign Languages Students*

Group	N	Mean Rank	Sum of Ranks	U	P
EFLTT*	43	34.64	1498.50	58.50	0.000
EFLL**	14	11.68	163.50		
EFLTT*: EFL Teacher Trainees					
EFLL:** EFL Learners at the School of Foreign Languages					

As displayed in Table 2, a significant difference ($p<0.01$) was found between the two groups in favor of the EFL teacher trainees, which would be expectable considering the differences in terms of the length and content of their work with English. The results concerned with the comparison of EFLTTs with NSs are given in Table 3.

Table 3: *Mann-Whitney Pair-wise Comparisons between the Teacher Trainees and Native Speakers of English*

Group	N	Mean Rank	Sum of Ranks	U	P
EFLTT*	43	22.78	979.50	33.500	0.000
NAT**	21	52.40	1100.50		
EFLTT*: EFL Teacher Trainees					
NAT:** Native Speakers of English					

According to the results in Table 3, there is a significant difference ($p<0.01$) between the two groups in favor of the native speakers, which is expectable considering the fact that English is their mother tongue while it is still a foreign language for the other group's members though they were at a relatively advanced level. The results of the comparison between EFLTTs and HSSs are provided in Table 4.

Table 4: *Mann-Whitney Pair-wise Comparisons between the Teacher Trainees and High School Students*

Group	N	Mean Rank	Sum of Ranks	U	P
EFLTT*	43	27.23	1171.00	225.000	.804
HSS**	11	28.55	314.00		
EFLTT*: EFL Teacher Trainees					
HSS:** High School Students					

The results given in Table 4 show that there is no significant difference ($p>0.05$) between the two groups. This could be considered predictable as students like those in the high school group function as the primary source of undergraduates for university programs such as English Language Teaching. Therefore, it is possible to postulate that the teacher trainee participants had the position of the high school students a couple of years ago while some of the latter would probably be the 1st year students of different ELT departments a couple of months later. Table 5 shows the results comparing EFLLs with NSs.

Table 5: *Mann-Whitney Pair-wise Comparisons between the School of Foreign Languages Students and Native Speakers of English*

Group	N	Mean Rank	Sum of Ranks	U	P
EFLL *	14	7.50	105.00	.000	.000
NAT**	21	25.00	525.00		
EFLL*: EFL Learners at the School of Foreign Languages					
NAT:** Native Speakers of English					

As put in Table 5, a significant difference ($p<0.01$) was found between the two groups in favor of the native speakers. This is expectable considering the fact that English is their mother tongue while the students of the School of Foreign Languages were officially diagnosed as "false beginners" for English nearly a year earlier.

As for the comparison of EFLLs with HSSs, according to the results provided in Table 6, a significant difference ($p<0.05$) was found between the two groups in favor of the High School students. This would be predictable considering the fact that they were a group that was formed with regard to their former achievements in EFL and they were getting a language intensive education to enroll for university programs based on EFL study. On the other hand, as mentioned before, the students of the School of Foreign Languages were officially diagnosed as false beginners nearly a year earlier.

Table 6: *Mann-Whitney Pair-wise Comparisons between the School of Foreign Languages Students and High School Students*

Group	N	Mean Rank	Sum of Ranks	U	P
EFLL*	14	8.89	124.50	19.500	.002
HSS**	11	18.23	200.50		

EFLL*: EFL Learners at the School of Foreign Languages
HSS:** High School Students

Finally, the results comparing NSs with HSSs are given in Table 7.

Table 7: *Mann-Whitney Pair-wise Comparisons between the Native Speakers of English and High School Students*

Group	N	Mean Rank	Sum of Ranks	U	P
NAT*	21	21.29	447.00	15.000	.000
HSS**	11	7.36	81.00		

NAT*: Native Speakers of English
HSS:** High School Students

The results in Table 7 show that there is a significant difference ($p<0.01$) between the two groups in favor of the native speakers, which would be expectable since English is their mother tongue while a foreign language for the other group's members although they were at a relatively advanced proficiency level.

The results presented in Tables 2-7 above suggest that the expectations and predictions were more satisfactorily fulfilled than the pilot study.

The first research question of the present study was aimed at exploring whether native speakers of English could reach a compromise in their interpretation of the test items. The results show that the performance of the native speakers was remarkably high, which points to a good compromise with their interpretations of the test items. At this point it would be worth mentioning that, in our day characterized by globalized communication in multiculturalism, using NS norms as a benchmark for pragmatic behavior may not be so crucial in a foreign language situation (Wyner & Cohen, 2015, p. 547). Nevertheless, in the strenuous attempt to develop a valid and recent MDCT to measure pragmatic comprehension about formulaic and teachable implied meanings, this study had the compelling need for norms to count as the "favored interpretations of the items". In that regard, no other appropriate way to have them could be conceived than taking the response options on which the native speaker participants reached a satisfactory compromise in the measurements. Apart from that, as Wyner and Cohen (2015, p. 547) put it with a comprehensive look, NS norms as a benchmark can be valuable for learners to have familiarity with what these norms are and help them figure out not only what went wrong in experienced pragmatic failures but also ways in which they could be avoided in future interactions.

The second research question was aimed at investigating whether the performance differences between particular pairs of test-taker groups could be attributable to their proficiency differences only. The findings reveal that there are statistically significant differences between the NS' performance and those of all the other participant groups. This shows that the study was able to address a problem worth pragmatic assessment and instruction. Apart from the comparison between the NSs and the other three groups, it is also seen that the performance differences between particular pairs of groups seem to be attributable to the proficiency differences. This could be argued to be a strength of the test as it seems to reflect the performance variability between participants from different proficiency levels. This is important since educational assessment is supposed to discriminate among those who are assessed, and a good test should produce scores that vary between high and low performers (Biggs, 1996). Besides, in the subsequent stage of the larger research project where this study is a part, all the above-mentioned results were confirmed (Çetinavcı, 2016). On the one hand, the mean performance of the 127 new native speakers was 26.61 out of 28 items, which verified once again the research hypothesis that the online MDCT would prove to be one on which NSs of English reach a good compromise with their interpretations of the items. On the other hand, in comparison with a new group of 144 first-year EFL teacher trainees, there was a statistically significant ($p < 0.01$) NS superiority at opting for the favored interpretations of the implied meanings. This verified once again the research hypothesis that the test will prove to be one where the performance differences between particular pairs of test-taker groups could be attributable to their proficiency levels.

Several studies (Kehoe, 1995; Hughes, 2003, p. 228; Quaigrain & Arhin, 2017) have indicated that distractors with no or very low choice frequency would point to their ineffective functioning. This was a concern in Roever's (2005) study as well, which reports on the first pragmatic competence test developed in the field of Applied

Linguistics. In this regard, another positive feature of the online MDCT developed in this study is the fact that all the distractors of every item were chosen by the target group members (EFL teacher trainees) in varying frequencies. This means that none of the distractors was just an ineffective space filler, which could give the possibility to claim that they functioned in the way they had been supposed to.

Like Roever's (2005, p. 52) piloting experience pursuant to Hudson et al. (1995); the process explained above took nearly a year, but it was felt that a well-designed test was essential to obtaining meaningful results. From the administration of the initial pilot test till that of its final version, after receiving a thank-you note from the researchers, some native speaker participants e-mailed their comments about their experience even though none of them had been asked or encouraged to in any way. The change between the beginning and end of the process could be viewed as quite dramatic, which justifies the work during the long development period. Below are given some comments to illustrate the point that has just been made. While the first three are from the beginning, the fourth one is from the midst and the others are from the end of the process:

1) "Hope it works out. I'm sure you know what you're doing but some of the phrases weren't actually idioms?"
2) "Some parts do not sound like native English at all."
3) "It was still possible in most cases to see what the intent was but it just sounded weird if that makes any sense. Anyways, best of luck!"
4) "It was an interesting test, although I do admit, I think some of the questions had 'wrong' answers."
5) "The test was very well written, and one can see a lot of thought went into it."
6) "Good evening, I wanted to let you know that I have completed your exam, it looks great."
7) "Hi, I have completed the test. It was kind of fun. I enjoyed it. Glad i was able to help."
8) "Thought this to be very interesting. Went quickly. The discussions seemed pretty clear cut to me."

CONCLUSION

To conclude, it would be worthwhile to state that the present study took almost a year for the development of a valid, well-designed MDCT that can be used as a data collection instrument to have comparable sets of data for further descriptive and/or instructional studies. Special care was taken in the test to include some particular subsections of items in response to the calls in the literature for integration of different implied meaning types to add to our understanding of pragmatic comprehension in a target language. What is more, the overall focus of the test was on "implicature (implied meanings)" so that the study could keep out of the reported weight of "speech acts" in pragmatics research and provide a new perspective upon another important but lesser-studied component of pragmatics.

Within this framework, we could emphasize the fact that the present study gave a tangible product: A data collection instrument on the comprehension of eight particular implied meanings in English, all of which have been scholarly conceptualized. Being a test which is usable both in computerized and pen-and-paper format, it proved to be one on which a relatively big, heterogeneous group of NSs of English reached a good compromise with their interpretations of the test items. When considered together with the other phases in the larger research project that covers this study as well, in addition to the total number of 174 NSs (a heterogeneous group of 111 Americans, 39 Britons, 13 Canadians, 10 Australians/New Zealanders and 1 South African), 316 EFL teacher trainees, 37 EFL learners at a university school of foreign languages, 13 Turkish citizens who had been schooled and lived in an English-speaking country for a decade or more and 11 high-school students getting a language-intensive education made quantitative and/or qualitative contribution in the development, refinement and implementation processes of the test. They all add up to 551 people.

In addition, the online test in its final version added another important feature: ability to measure each test taker's response times for every single test item and the whole test. This was triggered mainly by the perspective put by Taguchi (2005, 2007, 2008, 2011a), who noted that not many studies had addressed fluency or processing speed in language learners' pragmatic performance.

As the last point to be made about the significance of the study, it should be restated here that some of the implied meanings included are ones that have already been reported as formulaic. For the rest, as discussed thoroughly earlier in the text, the claim in the present study is that at least some of their variations can be deemed formulaic, thus teachable. To the best of the researcher's knowledge, the test is a pioneering one in pragmatic assessment to bundle together specifically that type of implied meanings. As one could not teach what s/he cannot initially measure, the test can serve as the starting point for any ELT or ELT teacher-training program concerned to help the students be equipped for that specific domain of pragmatic competence as one of the acknowledged requisites for overall communicative competence.

On the grounds of the limitations of this study and the experiences that accumulated throughout its conduct, some recommendations can be made for further research.

First of all, considering the fact that this study measured pragmatic comprehension with a reading instrument (like in many other previous inspiring studies) while people mostly "see and hear" in real-life communication, the procedures in similar future studies could be designed as based on a sufficient number of readymade video extracts or fictionalized dramas to the purpose. Provided that this is achieved with proper validation work in a manner where audiovisual items would not impede but aid the watchers or listeners (Gruba, 2000; Roever, 2005), the measurement of pragmatic interpretation could include such clues as tone of voice, setting, gestures and facial expressions, which all can express so much meaning together with or independently of the words there. Besides these, the ideal to be pursued within this framework would most probably be extracting discourse samples with the target implied meanings via corpora/concordance work and producing scenes out of them with proper use of tone of voice, facial expressions and gestures not open to ambiguity. The fuller the extent to which this is achieved, the more likely it would be to use the products in both testing and instruction procedures, which would give the researchers the chance to base their studies on authentic materials as much as possible.

In the context of discussing the content and scope of studies which are similar to this one, another recommendation for further research could me made about the identification and integration of even more implied meaning types into the designs so that we can add to our understanding of pragmatic comprehension/interpretation and learn which ones of them could be troublesome to EFL/ESL learners and why, which is an attempt made by the present study with the integration of "indirect pieces of advice" and "indirect requests". What is more, the range of L1s and target languages in studies on pragmatic interpretation could be expanded so that investigators and language educators can better assess whether and to what extent findings from studies of a particular L1 or target language may be valid in terms of other language combinations. Moreover, further research could be conducted also on how competent language learners are in terms of "producing" implied meanings. This would provide a new perspective in studies of this kind beyond the focus merely on comprehension/interpretation.

REFERENCES

Aijmer, K. (Ed.). (2011). *Contrastive pragmatics* (Vol. 30). Amsterdam/Philadelphia: John Benjamins Publishing.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press, USA.

Bardovi-Harlig, K. (1996). Pragmatics and language teaching: Bringing pragmatics and pedagogy together. In L.F. Bouton (Ed.) *Pragmatics and Language Learning* (Vol. 7, pp. 21_39). Urbana, IL: University of Illinois at Urbana-Champaign.

Bardovi-Harlig, K. & Dörnyei, Z. (1998). Do Language Learners Recognize Pragmatic Violations? Pragmatic Versus Grammatical Awareness in Instructed L2 Learning. *Tesol Quarterly, 32*(2), 233-259.

Bardovi-Harlig, K., & Shin, S. Y. (2014). Expanding Traditional Testing Measures with Tasks from L2 Pragmatics Research. *Iranian Journal of Language Testing, 4*(1).

Bella, S. (2014). Developing the ability to refuse: A cross-sectional study of Greek FL refusals. *Journal of Pragmatics, 61*, 35-62.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Essex, England: Pearson Education Limited.

Biggs, J. B. (Ed.). (1996). Testing: To educate or to select? Education in Hong Kong at the cross-roads. Hong Kong: Hong Kong Educational Publishing.

Birner, B. J. (2013) *Introduction to Pragmatics*. Chichester: Wiley-Blackwell.

Bouton, L. F. (1988). A cross- cultural study of ability to interpret implicatures in English. *World Englishes, 7*(2), 183-196.

Bouton, L. (1994). Can NNS skill in interpreting implicatures in American English be improved through explicit instruction? A pilot study. In L. Bouton & Y. Kachru (Eds.), *Pragmatics and language learning* (Vol. 5, pp. 88–109). Urbana-Champaign: University of Illinois, Division of English as an International Language.

Bouton, L. (1999). Developing nonnative speaker skills in interpreting conversational implicatures in English: Explicit teaching can ease the process. In E. Hinkel (Ed.), *Culture in second language teaching and learning* (pp. 47-70). Cambridge: Cambridge University Press.

Broersma, D. (1994). Do chickens have lips? Conversational implicature in the classroom. Paper presented at the 8th Annual Meeting of the International Conference on Pragmatics and Language Learning, Urbana, IL.

Brown, P. & Levinson S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Canale, M., & Swain, M. (1981). A theoretical framework for communicative competence. In A. Palmer, P. Groot, & G. Trosper (Eds.), *The construct validation of test of communicative competence: Including proceedings of a colloquium at TESOL '79* (pp. 31-36). Washington, DC: Teachers of English to Speakers of other Languages.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In Richards, J. C., & Schmidt, R. W. (Eds.), *Language and Communication* (pp. 2-27). London: Longman.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

Cohen, A. D. (2012). Teaching pragmatics in the second language classroom. *The European Journal of Applied Linguistics and TEFL, 1* (1), 35-49.

Cohen, A. D. (2012). Research Methods for Describing Variation in Intercultural Pragmatics for Cultures in Contact and Conflict. In J. C. Félix-Brasdefer & D. A. Koike (Eds.), *Pragmatic Variation in First and Second Language Contexts: Methodological Issues* (pp. 271-294). Amsterdam/Philadelphia: John Benjamins.

Colston, H. L., & O'Brien, J. (2000). Contrast and pragmatics in figurative language: Anything understatement can do, irony can do better. Journal of Pragmatics, 32(11), 1557-1583.

Council of Europe. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.

Crandall, E., & Basturkmen, H. (2004). Evaluating pragmatics-focused materials. *ELT Journal, 58*(1), 38-49.

Çetinavcı, U. R. (2016). *The effects of explicit film-based instruction on English as a foreign language teacher trainees' interpretation of implied meanings* (Unpublished doctoral dissertation). Canakkale 18 Mart University, Çanakkale, Turkey.

Davis, W. A. (2007). How normative is implicature. *Journal of pragmatics, 39*(10), 1655-1672.

Economidou-Kogetsidis, M. (2015). Teaching email politeness in the EFL/ESL classroom. *ELT Journal, 69*(4), 415-424.

Erkmen, B. (2014). Novice EFL Teachers' Beliefs about Teaching and Learning, and their Classroom Practices. *Hacettepe University Journal of Education, 29*(1), 99-113.

Eslami, Z. R. & Mirzaei, A. (2012). Assessment of second language pragmatics. In C. Coombe, P. Davidson, B. O'Sullivan & S. Stoynoff (Eds.), *The Cambridge Guide to Second Language Assessment* (pp. 198-208). Cambridge: Cambridge University Press.

Green, G. (1989). *Pragmatics and natural language understanding*. New York: Erlbaum.
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics. 3. Speech Acts* (pp. 41-58). New York: Academic Press.
Grice, H. P. (1981). Presupposition and conversational implicature. In P. Cole (Ed.), *Radical Pragmatics*. (pp. 183-198). New York: Academic Press.
Gruba, P. (2000). The role of digital video media in response to task demands. In *22nd Annual Language Testing Research Colloquium*, Vancouver, BC, Canada.
Hadi, A. (2013). A Critical Appraisal of Grice's Cooperative Principle. *Open Journal of Modern Linguistics, 3*, 69-72.
Haladyna, T. M., & Downing, S. M. (1989). A taxonomy of multiple-choice item-writing rules. Applied measurement in education, 2(1), 37-50.
Hinkel, E. (1997). Appropriateness of Advice: DCT and Multiple Choice Data. *Applied Linguistics, 18*(1), 1-26.
Hudson, T., Detmer, E., & Brown, J. D. (1995). *Developing prototypic measures of cross-cultural pragmatics*. Honolulu: HI: University of Hawaii, Second Language Teaching & Curriculum Center.
Hughes, A. (2003). Testing English for Language Teachers. Cambridge: Cambridge University Press.
Hymes, D. H. (1972). On communicative competence. In J. B. Pride, & J. Holmes (Eds.), *Sociolinguistics (pp. 269-293)*. Baltimore: Penguin Books.
Jianda, L. (2006). Assessing EFL learners' interlanguage pragmatic knowledge: Implications for testers and teachers. *Reflections on English language teaching, 5*(1), 1-22.
Kasper, G. (1997). Can pragmatic competence be taught? *NetWork, 6*, 105-119.
Kasper, G. (1999). Learning Pragmatics in the L2 classroom. Paper presented at the annual international conference on *Pragmatics and Language Learning*. University of Illinois at Urbana-Champaign.
Kehoe, J. (1995). Basic Item Analysis for Multiple-Choice Tests. ERIC/AE Digest.
Kreuz, R. & Roberts, R. (1995). Two cues for verbal irony: Hyperbole and the ironic tone of voice. *Metaphor and Symbol, 10* (1), 21-31.
Kubota, M. (1995). Teachability of conversational implicature to Japanese EFL learners. *Institute for Research in Language Teaching Bulletin, 9*, 35-67.
Lee, J. S. (2002). Interpreting conversational implicatures: A study of Korean learners of English. *The Korea TESOL Journal, 5*(1), 1-26.
Levinson, S. (1983). *Pragmatics*. Cambridge, U.K.: Cambridge University Press.
Longman. (1995). *Longman's dictionary of contemporary English*. New York: Longman.
Matsuda, M. (1999). Interlanguage pragmatics: What can it offer to language teachers ? *The CATESOL Journal, 11*(1), 39-59.
Matsumura, S. (2001). Learning the rules for offering advice: A quantitative approach to second language socialization. *Language Learning, 51*(4), 635-679.
Matsumura, S. (2007). Exploring the aftereffects of study abroad on interlanguage pragmatic development. *Intercultural Pragmatics, 4* (2), 167-192.
McTear, M. (2004). *Spoken dialogue technology: Towards the conversational user interface*. The Netherlands: Springer.
Özmen, K. S. (2012). Exploring student teachers' beliefs about language learning and teaching: A longitudinal study. *Current Issues in Education, 15*(1), 1-15.
Pichastor, M. S. (1998). Politeness and textbooks: how to approach the teaching of communicative competence in a second language. In I. V. Orta & I. G. Galve (Eds.), *Perspectivas Pragmáticas en Lingüística Aplicada* (pp. 299-306). Anubar.
Quaigrain, K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. Cogent Education, 4(1), 1301013.
Rızaoğlu, F. & Yavuz, M. A. (2017). English Language Learners' Comprehension and Production of Implicatures İngilizce Öğrenenlerin Sezdirileri Anlama ve Üretme Becerileri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education)*. doi: 10.16986/HUJE.2017027932
Rinnert, C., & Kobayashi, H. (1999). Requestive hints in Japanese and English. *Journal of Pragmatics, 31* (9), 1173-1201.
Roever, C. (2005). *Testing ESL pragmatics*. Germany: Peter Lang.
Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing, 23*(2), 229-256.
Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing, 28(4)*, 463-481.
Roever, C. (2013). Assessment of pragmatics. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-8). Hoboken, New Jersey: John Wiley & Sons, Inc.
Streisand, B. and Streisand, B./Karsch, A. S. (Director/Producers). (1991). *The Prince of Tides* [Film]. U.S.: Columbia Pictures.
Taguchi, N. (2002). An application of relevance theory to the analysis of L2 interpretation processes: The comprehension of indirect replies. *IRAL, 40*(2), 151-176.

Taguchi, N. (2005). Comprehending implied meaning in English as a foreign language. *The Modern language journal, 89*(4), 543-562.
Taguchi, N. (2007). Development of speed and accuracy in pragmatic comprehension in English as a foreign language. *TESOL Quarterly, 41*, 313–338.
Taguchi, N. (2015). Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *Language Teaching, 48*(1), 1-50.
Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics 4*(2), 91-112.
Uztosun, M. S. (2013). An interpretive study into elementary School English Teachers' Beliefs and Practices in Turkey. *Turkish Online Journal of Qualitative Inquiry, 4*(1), 20-33.
Verschueren, J. (2009). Introduction: The Pragmatic Perspective. In J. Verschueren & J. O. Östman (Eds.), *Key notions for pragmatics* (pp. 1-28). Amsterdam: John Benjamins Publishing.
Weizman, E. (1985). Towards an analysis of opaque utterances: Hints as a request strategy. *Theoretical linguistics, 12*(1), 153-163.
Weizman, E. (1989). Requestive hints. In S. Blum-Kulka, J. House & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 7-95). Norwood, NJ: Ablex Publishing.
Weizman, E. (1993). Interlanguage requestive hints. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage Pragmatics* (pp. 123-137). Oxford: Oxford University Press.
Weizman, E. (2007). Quantity scales: towards culture-specific profiles of discourse norms. In M. Grein & E. Weigand (Eds.), *Dialogue and Culture* (pp. 141-152). Amsterdam/Philadelphia: John Benjamins.
Wolfson, N. (1989). *Perspectives: Sociolinguistics and TESOL*. Boston: Heinle & Heinle.
Wyner, L. & Cohen, A. D. (2015). Second language pragmatic ability: Individual differences according to environment. *Studies in Second Language Learning and Teaching, 5*(4), 519-556.
Yu, M. C. (2006). On the teaching and learning of L2 sociolinguistic competence in classroom settings. *Asian EFL Journal, 8*(2), 111-131.

APPENDIX 1. MDCT Items

Item 1:

Tom is from Atlanta. His friend Sally has recently moved to Atlanta.

Tom: "How do you like Atlanta so far?"

Sally: "I love it!"

What does Sally probably mean?

- She thinks that Atlanta is a dirty city.
- She has not seen much of the city since she moved in.
- She thinks the city needs more great changes.
- She likes Atlanta and enjoys living there.

Item 2:

Henry loves cycling. He orders a new, very expensive bicycle from a new bicycle company. When it arrives, he sees that it is really heavy and does not look well-made at all.

Henry: "Wow, this company's really honest."

What does Henry probably mean?

- The company is dishonest.
- The company is a tiny bit sneaky.
- The company is a really honest one.
- It is normal as the company is new.

Item 3:

Jose and Tanya are professors at a college. They are talking about a student, Derek.

Jose: "How did you like Derek's essay?"

Tanya: "Well . . I thought it was well-typed."

What does Tanya probably mean?

- She did not like Derek's essay.
- She does not really remember Derek's essay.
- She thought the topic Derek had chosen was interesting.
- She liked Derek's essay quite a lot.

Item 4:

Judie and her classmate David are community college freshmen. Judie is considering taking a course but David has heard it is really difficult.

David: "I don't know . . . but people say it's really difficult."

What does David probably mean?

- He thinks the course may not be very difficult.
- He thinks Judie can take that course.
- He recommends not taking that course.
- He thinks Judie should not listen to what people say about the course.

Item 5:

Rob is telling his friend Sheila about a card game he played last night. He lost money and decides not to play with those guys again.

Sheila: "They were good, huh?"

Rob: "Good? Let's say awfully lucky".

Sheila: "Lucky? What's the matter? Don't you trust them?"

Rob: "Is the sky green?"

What does Rob probably mean?

- He thinks they are OK.
- He does not want to talk about the card game anymore.
- He suddenly saw something in the sky.
- He does not trust them at all.

Item 6:

Jack sees his classmate Jane in the faculty hallway.

Jack: "Oh, Jane. I'm so glad I ran into you. I need your help!"

Jane: "What's up?"

Jack: "I have a paper due tomorrow, but I'm working tonight in the cafe. Can you type my paper?"

Jane: "Shoot! I have to study for my finals tonight."

What does Jane probably mean?

- She will type the paper.
- She will think about it.
- She cannot type the paper for tomorrow.
- She can type it when she is done with everything.

Item 7:

Susan and John, friends, are watching a film together.

Susan: “This film is too boring! I can’t watch it anymore.”

John: “Really? I don’t think it’s so bad.”

What does John probably mean?

- He thinks the film is really bad.
- He is doing something else, not watching the film.
- He does not think the film is very bad.
- He is not quite sure.

Item 8:

Bob and Maggie, friends, are talking about school and courses. Bob is taking introductory chemistry this semester.

Maggie: “How are you doing in chemistry?”

Bob: “So . . . did you watch that basketball game yesterday?”

What does Bob probably mean?

- The content of yesterday's lesson was completely irrelevant to chemistry like a basketball game.
- He is doing badly in chemistry.
- Chemistry is like an easy game for him.
- He is doing so well in chemistry that there is no need to talk about it.

Item 9:

Carol, an office secretary at a university, is typing at her desk. Jeff, a teacher, is in Carol's office to make a lot of printouts.

Jeff: "The printer is almost out of ink."

What does Jeff probably mean?

- He wants Carol to refill the ink.
- He uses the printer really very often.
- He does not want to do the printing himself.
- Carol can continue what she is doing.

Item 10:

Mike is trying to find an apartment in New York City. He just looked at a place and is telling his friend Jane about it.

Jane: "So, is the rent high?"

Mike: "Is the Pope Catholic?"

What does Mike probably mean?

- He does not want to talk about the rent.
- The rent is high.
- He did not understand Jane's question.
- The rent is not very high.

Item 11:

Felicity is talking to her co-worker Brian during a coffee break.

Felicity: "So, life must be good for you. I hear you got a nice raise."

Brian: "Um, this coffee is awfully weak. You'd think they'd at least give us decent coffee."

What does Brian probably mean?

- He does not want to talk about how much money he earns.
- He does not like the coffee.
- Reality may not be what you think it is.
- He does not care about money.

Item 12:

Roger is thinking of taking his car to a repair shop in the city centre. His friend Melanie knows that the shop is known for doing careless work.

Melanie: "I don't usually take my car there. It has a really bad reputation."

What does Melanie probably mean?

- Roger should take his car there for only small repairs.
- She advises Roger not to take his car to that repair shop.
- The reputation of a place is important.
- Roger can take his car there.

Item 13:

Paul and Mary, two friends, are having a talk. Paul remembers that he must pay his apartment's rent today but has no money for it now.

Paul: "Oh, the rent is due today, but I don't get paid until Monday. Could I borrow $50? I'll give it back next week."

Mary: "Sure, no problem."

What does Mary probably mean?

- She is not sure about giving money to Paul.
- She will give the money to Paul.
- It is a problem for Paul.
- She will not give the money.

Item 14:

Barbara and Brad, classmates, are talking about what they are going to do during the summer. Barbara's mother wants her to stay home, and entertain the relatives when they come to visit them at the beach.

Brad: "Do you have a lot of relatives?"

Barbara: "Does a dog have fleas?"

What does Barbara probably mean?

- She does not like her relatives and feels like an unlucky dog.
- She does not have very many relatives.
- She has a lot of relatives.
- She wants to learn if a dog usually has fleas.

Item 15:

Toby and Ally are trying a new buffet restaurant in town. Toby is eating something, but Ally cannot decide what to have next.

Ally: "How do you like what you're eating?"

Toby: "Well, let's just say it's . . colorful."

What does Toby probably mean?

- He thinks it is important for food to look good.
- He likes the food.
- He wants Ally to try something colorful.
- He does not like the food much.

Item 16:

Peter promises his friend Mary to help her move to a new apartment. That day, he moves the clock on the wall while Mary moves the heavy boxes.

Mary: "Thanks, you've been terribly helpful."

What does Mary probably mean?

- Peter helped her a lot.
- Moving the clock was really important as it needed special care.
- Peter is weak.
- Peter was not helpful at all.

Item 17:

John's friend Mary asks him about their classmate Sally.

Mary: "You know. I've been curious to know if you went out with Sally."

John: "Um . . Sally's not really my type."

What does John probably mean?

- He is not sure of his feelings.
- He is talking bad about Sally as she refused him.
- Mary is his type.
- They did not go out.

Item 18:

Dale runs into his friend Julia. He knows Julia recently had a job interview.

Dale: 'By the way, did you get that job you applied for?'

Julia: 'Good God, I'm so tired of this cold weather.'

What does Julia probably mean?

- She does not want to talk about the interview.
- She is bored of searching for a job.
- She did not understand Dale's question.
- She could not attend the interview because of cold weather.

Item 19:

Maria and Frank are working on a class project together but they will not be able to finish it by the deadline.

Maria: "Do you think Dr. Gibson is going to lower our grade if we hand it in late?"

Frank: "Do fish swim?"

What does Frank probably mean?

- He thinks they should choose a new project topic on fish.
- He thinks Dr. Gibson will not lower their grade.
- He thinks they will get a lower grade.
- He suggests just giving in the project to see the result.

Item 20:

Hillary sees that her boyfriend Bruce has forgotten to leave a tip while leaving the restaurant they had dinner in.

Hillary: 'You know, leaving a tip is important.'

What does Hillary probably mean?

- She advises him to leave a tip.
- She indirectly asks Bruce if they should leave a tip or not.
- It is OK now but Bruce should not forget the tip next time.
- She wants to leave quickly without tipping.

Item 21:

Jenny is out in the freezing cold after basketball practice. As she often has to do, she has been waiting for her mom to pick her up for an hour. She throws a quick glance at her watch, talking to herself.

Jenny: "She's a bit late huh?"

What does Jenny probably mean?

- Her mom is not very late yet.
- She is anxious about her mom.
- Her mom is really late once again.
- She needs to look at her watch again.

Item 22:

Susan and Ronald, two officemates, are having lunch in a café and discussing their boss.

Ron: 'So, do you think Mr. Davis will give me a raise?'

Susan: 'Do pigs fly?'

What does Susan probably mean?

- She wants to change the topic.
- The boss will not give Ron a raise.
- She has seen outside a pig falling down from a high place.
- Ron will get a raise.

Item 23:

Joan and Dave, classmates, see each other in the school corridor.

Joan: 'Hi Dave.'

Dave: 'Hi Joan. What's up?'

Joan: 'I was going to ask you a favor. Would you read my paper for English 101?'

Dave: 'Oh, Joan, sorry I can't. I have a class in about 10 minutes.'

What does Dave probably mean?

- He will read the paper.
- That is a difficult thing to do for him.
- He will read it after the class.
- He will not read the paper because he is busy.

Item 24:

Susan and Tom, friends, are talking about what is going on in their lives. Susan knows Tom had a job interview recently.

Susan: 'So how was your interview? Did you get the job you applied for?'

Tom: 'Um . . . I think I need to improve my interview skills.'

What does Tom probably mean?

- He did not get the job.
- He wants help from Susan to improve his interview skills.
- He will have the interview when he feels his interview skills are good enough.
- They gave him the job with the advice that he should improve his interview skills.

Item 25:

Ken bought a new car and he showed it to his co-worker, Tina. She drove it around for a couple of times and they are talking at lunchtime the next day.

Ken: 'So what do you think of this new car?'

Tina: 'Well, the color's fine.'

What does Tina probably mean?

- What she liked most about the car is its color.
- She thinks the color of a car is very important.
- She does not know much about cars.
- She did not like the car very much.

Item 26:

Hilda is looking for a new job. She is having lunch with her friend John.

John: "So how's the job search coming along?"

Hilda: "Um, this curry's really good, don't you think?"

What does Hilda probably mean?

- She did not understand John's question.
- She is not looking for a job anymore.
- She wants to talk about nothing but food.
- Her job search is not going very well.

Item 27:

Tom and Mary share the same apartment. Tom finds Mary in the kitchen.

Tom: 'Hey, ah . . . could you clean the house this weekend? I have plans.'

Mary: 'Oh, ah . . I'm going to see my parents this weekend.'

What does Mary probably mean?

- She will clean the house.
- She will try to make some new arrangements.
- She refuses Tom's request.
- She thinks the house does not need cleaning.

Item 28:

Michael is planning not to come to today's class. His housemate Angela knows one absence loses five points in the end.

Angela: 'Well, you know, one absence loses five points from the final marks.'

What does Angela probably mean?

- Michael has already lost 5 points.
- She advises Michael to come to the class.
- She will remind the teacher to take off five points.
- She recommends that he should do as he wishes.

Item 29:

Nina, an office secretary at a university, is working at her desk. Tom, a teacher, is there to make photocopies but the machine is not working.

Tom: 'The copy machine isn't working.'

What does Tom probably mean?

- He asks permission to make the photocopies.
- He indirectly criticizes Nina for not doing her job.
- He wants help from Nina with fixing the machine.
- He wants Nina to continue what she is doing.

Item 30:

Sally and Dennis, old friends, see each other again after a long time. Sally has heard that Dennis got divorced but is not sure.

Sally: 'By the way, is it true you got divorced?'

Dennis: 'You know . . . I think we got married too young.'

What does Dennis probably mean?

- They are still in that unhappy marriage.
- They are not married anymore.
- They are OK, but it would have been better if they had got married older.
- He does not want to answer the question.

Item 31:

Bob and Sarah, two school friends, are halfway to finishing this semester. They are talking about the courses they are taking.

Bob: 'By the way, how are you doing in history?'

Sarah: Um . . . not so well. I got a 'C' on the last test.

What does Sarah probably mean?

- She is doing really well in history.
- She loves history.
- She is not sure about her performance.
- She is not doing so well in history.

Item 32:

Brenda and Sally, friends, have lunch every Tuesday. As they meet on this particular day, Brenda stops and twirls like a fashion model, smiling.

Brenda: 'I just got a new dress. How do you like it?'

Sally: 'Well . . . it's certainly a popular style'

What does Sally probably mean?

- Brenda should have bought it earlier.
- She really likes it.
- Every dress is the same for her.
- She does not like it much.

Item 33:

Mark and Jane work in the same factory. They are both at work.

Mark: 'Hey Jane. Are you busy?'

Jane: 'Ah . . not right now. We just finished that big project.'

Mark: 'Wow, good for you. I know that was a lot of work. By the way, can you work my night shift this Friday Jane? My son is graduating from college.'

Jane: 'Um . . I'm having a party Friday.'

What does Jane probably mean?

- She indirectly invites Mark to the party.
- She will not work Mark's shift.
- She will relieve her tiredness of the night shifts with the party.
- She can work Mark's night shift.

APPENDIX 2. The MDCT Item Specifications

Implied Meaning Types	Item Numbering
1. Fillers (5 items)	1, 7, 13, 23, 31
2. Pope Question (5 items)	5, 10, 14, 19, 22
3. Indirect Criticism (4 items)	3, 15, 25, 32
4. Topic Change (4 items)	8, 11, 18, 26
5. Indirect Advice (4 items)	4, 12, 20, 28
6. (Verbal) Irony (3 items)	2, 16, 21
7. Indirect Refusals (3 items)	6, 27, 33
8. Disclosure (3 items)	17, 24, 30
9. Indirect Requests (2 items)	9, 29

APPENDIX 3. Favored Response Options for the Items

Item Number	Favored Option	Item Number	Favored Option
1.	d*	18.	a
2.	a	19.	c
3.	a	20.	a
4.	c	21.	c
5.	d	22.	b
6.	c	23.	d*
7.	c*	24.	a
8.	b	25.	d
9.	a	26.	d
10.	b	27.	c
11.	a	28.	b
12.	b	29.	c
13.	b*	30.	b
14.	c	31.	d*
15.	d	32.	d
16.	d	33.	b
17.	d		

* excluded from the final analyses for being a filler item