

Measuring English Pragmatics Skills through Game-Based Performance

G. Tanner Jackson, Patricia Inglese, Jennifer Wain, Veronika Timpe-Laughlin, Lindsay Grace

Abstract—The current work describes efforts related to the evaluation of a game-based assessment called “Awkward Annie” which targets English language pragmatics by having players intentionally select the most awkward responses. Results are explored from 328 native English speaking adults collected through Amazon Mechanical Turk. This article explores a conceptual model of potential relations between individual differences, user experience, game performance, and English pragmatic skills. Specifically, results suggest that a foundational understanding of how to be inappropriate has significant implications for user experience and game performance, both of which are related to overall English pragmatics skills.

Index Terms— Educational Games, English Pragmatics, User Experience

I. INTRODUCTION

Successful communication requires more than the ability to form grammatically accurate sentences; it involves an awareness of the socio-cultural environment, the interlocutors, and the local context in any given communicative interaction. While traditional language learning instruction provides a foundation of what to say, pragmatics focuses on the more nuanced understanding required to know how to say what to whom and when [1]. Sociocultural linguistic skills can be learned through direct instruction and/or conversational (immersed) experience. However, these face-to-face educational approaches can be costly and/or impractical to implement (e.g., traveling to another country) and are frequently absent from traditional classroom environments [2]. The growing feasibility and popularity of educational games, provides a technological platform for pragmatics instruction and provides the advantages of being experiential, engaging, and fun [3]. Additionally, an educational game offers a cost effective platform for distribution, in a low stakes environment affording many opportunities to engage, experiment, learn, and fail. Thus, Game-Based Assessment (GBA) for measuring English language pragmatics is a promising area of research due to its critical role in communicative language use.

G.T. Jackson is with Educational Testing Services (ETS), Princeton, NJ 08541 USA. (609-252-8491; e-mail:gtjackson@ets.org).

P. Inglese is with ETS, Princeton, NJ 08541 USA. (609-252-8437; e-mail: pinglese@ets.org).

J. Wain is with ETS, Princeton, NJ 08541 USA. (609-734-5536; e-mail: jwain@ets.org).

L. Grace is with American University, Washington DC, 20016 (212-885-1000 e-mail: grace@american.edu).

V. Timpe-Laughlin is with ETS, Princeton, NJ 08541 USA. (609-252-8489; e-mail: vlaughlin@ets.org).

II. ENGLISH LANGUAGE PRAGMATICS

Pragmatic ability refers to the appropriate use of language in a particular socio-cultural context. Pragmatic ability helps people deliver and interpret a communicative message as intended and is therefore considered to have a major role in successful communication [4], while a lack thereof has been known to cause severe communication breakdown in the workplace [5] and create a negative impression of the speaker [6]. Despite its recognized importance, materials specifically targeting pragmatic instruction and assessment are limited.

Because pragmatic ability relies on an understanding of social and cultural norms, an assessment of the construct requires the presentation of rich contextual and sociocultural input, making an adaptive online virtual environment more feasible and appropriate than traditional text-based instructional materials [7]. The field of pragmatics currently lags behind other fields in terms of educational games for learning [8], and most pragmatics game research has been conducted by only one researcher, Julie Sykes, and her colleagues who focus exclusively on the acquisition of Spanish pragmatics [9]. To advance the field into the learning and assessment of English pragmatics, the GBA being developed for this project aims to represent some of the complexity of sociocultural interactions by providing the necessary contextual information (e.g., character backgrounds, power relationships, sociocultural environment) to assess English pragmatic awareness within an engaging and interactive game design. Next, we will discuss the specific designs of that GBA followed by details of an empirical evaluation.

III. GAME-BASED ASSESSMENT DESIGN

Interactive GBAs have the distinct advantage of providing a more engaging instructional format than traditional methods, showing consistently higher associations with motivation and persistence, which are two critical components of learning [8]. GBAs are well-suited platforms for pragmatics learning and assessment given that contextual information and feedback, critical pragmatics components, can be saliently presented. Additionally, the interactive nature of simulation games mimic the real world, activating processes from social constructionist and experiential learning theories [11], [12].

The design goal for the current GBA was to engage adults and measure aspects of English language pragmatic dependencies across three social dimensions (social familiarity, power differences, and imposition size). These social dimensions, identified in prior work [13], are considered a foundational means to elicit and identify pragmatic ability.

In addition to assessing pragmatic ability, the current work also aimed to investigate the impact of game design choices on

user experience. An Evidence Centered game Design approach (ECgD)[14] was used to iterate over game concepts which attempted to balance game and assessment design constraints. Ultimately this process resulted in a decision to use a series of selected-response conversations which evolved into a game called Awkward Annie (see Figure 1). Annie is an employee starting a new job set within an experientially rich environment including conversations between people with varying degrees of familiarity and a diverse hierarchical environment (i.e., strangers and friends as colleagues and bosses). Sampling these features across the social dimensions facilitates the assessment (and learning) of pragmatic abilities [15].

Awkward Annie's design takes an unorthodox approach to conversation selection by having the players intentionally select the most inappropriate and awkward response. It was hypothesized that reversing the goal (being inappropriate rather than appropriate), would improve the design from a game perspective by allowing users to play with social norms and potentially escape from reality by interacting in ways not feasible or socially acceptable during day-to-day interactions. This approach also allows for the inclusion of more playful and amusing conversations with exaggerated Non-Player Character (NPC) reactions [3], [16]. A potential limitation from the perspective of pragmatic literacy assessment, this strategy depends on the assumption that identifying the least appropriate response indicates knowledge of the appropriate response (by the absence of its selection).



Figure 1. Awkward Annie conversation selection (top panel), environment (bottom-left), and characters (bottom-right).

Thus, Awkward Annie is a GBA where users play the role of “Awkward Annie” who is an odd, but endearing character who humorously fails at social interactions (top-right character in Figure 1). Players navigate through an office environment to engage in the conversations with NPCs (bottom-left of Figure 1). In each conversation the player is presented with three dialogue options (appropriate, mildly inappropriate, and most inappropriate) and asked to select the most inappropriate statement, which is then displayed as Annie’s conversational

turn. Scores were awarded based on the correctness of the selection and a minimum score was required to complete a given conversation (see procedure below for specific values). A conversation can be replayed until the minimum score is reached for players to progress to subsequent conversations with a particular character; however, players may interact with any of the currently available characters in the order of their choosing. When players select awkward statements they can watch as the NPCs find this behavior strange and rightly react to her ostentatious shenanigans (bottom-right images of Figure 1), but (amazingly) continue to tolerate her.

IV. STUDY DESIGN

The current study aims to explore several factors related to GBA design of English language pragmatics. Specific research questions include: RQ1) how does the evidence from Awkward Annie relate to a conventional measure of pragmatics, RQ2) how do users perceive their interactions with Awkward Annie, and RQ3) how do aspects of user experience relate to various measures of performance (game and external survey)? Specifically, this work explores relations among pragmatic performance, game performance, user experience, and individual differences. A previous study with predominantly non-native English speaking adults revealed significant trends among these critical variables [17]. The current work expands on that initial study and aims to establish a baseline of relations through a sample of 328 native English speakers. Participants were recruited through Amazon Mechanical Turk (MTurk), received \$10 for participating, and the sample self-reported as 43% female, 79% as Caucasian (8% African American, 7% Asian American, 8% Hispanic or Puerto Rican, and 2% Native American/Alaskan), and 84% as being between 21 to 45 years old.

A. Procedure

Participating adults worked for one hour and completed three self-paced phases: presurvey, gameplay, and postsurvey.

Presurvey. Participants used their MTurk ID to log in and begin the presurvey, which took approximately 15 minutes. The presurvey was used to collect demographic information as well as participants’ education, languages spoken, self-reported use and application of the English language, and use of and comfort with technology. The end of the presurvey presented introductory information for Awkward Annie.

Awkward Annie. After completing the presurvey, participants played Awkward Annie (v1.39) for approximately 15-20 minutes. As previously described, Awkward Annie focused on the adult workplace and included 18 total conversations with the NPCs, with each conversation consisting of three user turns. In v1.39 individual response scores for appropriate, inappropriate, and very inappropriate corresponded to 0, 3, and 5 points respectively, with a minimum of 10 points needed to complete a conversation.

Game performance was represented by two variables: game success efficiency and average conversation score. Game success efficiency (% Success) is the number of successful

conversations divided by the total number of conversation attempts, then multiplied by 100 to produce a percent. This variable represents players' effectiveness at navigating the conversations with a higher value indicating more successful game play. Each player's average conversation score (Avg Score) serves as an indicator of how well they discriminated between the three conversational options (appropriate, mildly inappropriate, and most inappropriate). A higher average score per conversation indicates more successful discrimination.

Postsurvey. After playing the game, participants completed the postsurvey including items related to their user experience (UX) and English language pragmatic ability. The UX questions consisted of 26 likert scale items (1-strongly disagree, 6-strongly agree) focusing on positive aspects (e.g., "I had fun while working on the game"), negative emotions (e.g., "I was frustrated while playing the game"), basic gameplay (e.g., "I understood how to play the game"), and general perceptions (e.g., "I felt that the characters responded appropriately to my choices", "I selected statements to the best of my ability"). Pragmatic skills were assessed through 27 conventional selected-response items requiring identification of appropriate language use within particular contexts.

V. RESULTS

Analyses focused on identifying basic relations between English language pragmatics skills, game performance, and user experience. Initial analyses on the proportion of correct responses on the postsurvey revealed a range of pragmatic ability for these native English speakers ($M=.77$, $SD=.08$, $range=.57$). Subsequent correlations yielded a significant positive relation between the postsurvey pragmatics performance and both game performance variables (% Success: $r=.132$, $p=.017$; Avg Score: $r=.158$, $p=.004$). These initial results provide evidence that Awkward Annie gameplay captures aspects related to players' pragmatic ability (RQ1).

An interesting caveat is that players' performance within the game is significantly correlated with how well a player "understood how to play the game" through the underlying mechanic of selecting the most inappropriate conversation option (% Success: $r=.183$, $p=.001$; Avg Score: $r=.164$, $p=.003$). However, understanding how to play the game was not significantly related to pragmatic skills ($r=.056$, $p=.304$). Thus, pragmatic ability did not appear to impact players' ability to understand that the intent was to be awkward.

After examining the initial performance relations, subsequent analyses focused on how the unique mechanic of being awkward impacted players experience and perceptions (RQ2). Specifically, Table 1 displays a representative subset of correlations exploring the relations between individual items that represent the aforementioned categories of user experience variables (positive, negative, gameplay, and perceptions). For example, players who had fun while playing Awkward Annie (#1 in Table 1) were also less frustrated (#2), had a better understanding of how to play (#3), liked selecting awkward statements (#4), felt like the characters reacted appropriately to

the selections they made (#5), felt that they performed to the best of their ability (#6), and (#7) felt that the interaction was more playful (i.e., less like homework).

Table 1. User Experience Correlations (n=328, all $p<.001$).

	(2)	(3)	(4)	(5)	(6)	(7)
(1) Had fun while playing	-.552	.336	.636	.369	.222	-.659
(2) Was frustrated	1	-.281	-.468	-.309	-.188	.451
(3) Understood how to play		1	.292	.350	.548	-.352
(4) Liked being awkward			1	.321	.258	-.608
(5) Appropriate reactions				1	.248	-.442
(6) Performed at my best					1	-.332
(7) Felt like homework						1

Interestingly, there were no significant relations among technology usage (frequency of using computers, comfort level with technology, frequency of playing video games) and players' understanding of how to play the game, game performance, or even pragmatic skills (all $p>.05$). Although the game mechanic itself (being inappropriate) may be somewhat out of the ordinary, the game design did not require significant technical expertise to engage.

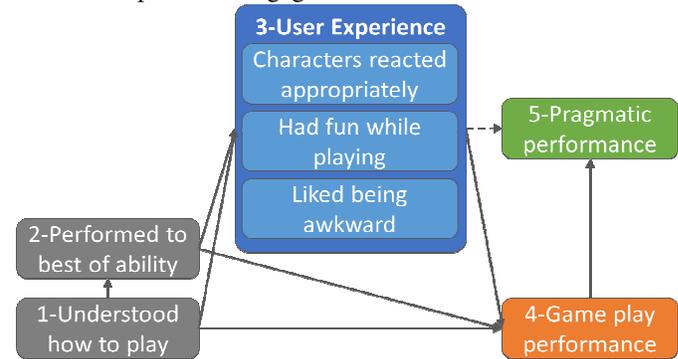


Figure 2. Proposed conceptual model of study relations.

Synthesizing these and additional correlations, a tentative conceptual model is proposed in Figure 2. The findings suggest that a player's reported level of understanding how to play the game (#1 in Figure 2) relates to their overall user experience (#3) as well as both their perceived (#2) and actual (#4) performance levels in the game. Thus, people who understood the intent to be awkward found the experience to be fun, playful (less like homework), and knew how to succeed in the game (RQ3). Conversely, understanding the need to be inappropriate may pose a difficulty for certain users, and those users seem to have difficulty playing and enjoying the game. Additionally, a player's perception of performing to the best of their ability (#2) has a significant impact on their experience (#3) and game performance (#4), with the latter variables (#3, #4) being positively related to pragmatic ability (#5)(RQ3).

VI. CONCLUSIONS AND DISCUSSION

Overall, the Awkward Annie efforts provided an operational conceptualization of (a portion of) English pragmatic ability aligned with existing literature [15], developed a GBA which addresses these skills, and conducted an investigation of relevant factors for game and pragmatics performance.

Although the results from the set of performance relations may draw a subtle distinction, it is an important one for the design of the GBA. Namely, for people who understood that the goal of the game was to play as an awkward character (i.e., not how they likely behave in everyday life), their game performance was significantly related to their pragmatic performance (RQ1). Additionally, understanding this novel intent of the game was not reliant on already having a specific level of pragmatic ability. This suggests that the Awkward Annie environment is successfully measuring aspects of English pragmatic ability, and that particular pragmatic skill levels may not be a prerequisite to understand the game. These results are also consistent with an initial study using predominantly non-native English speakers [17]. These results are informative for assessment purposes as the unique approach to confirming the negative response (being inappropriate) does appear to be providing evidence of the overall competency (RQ1). The findings are also relevant for education and game-based research communities due to the novel game design, associated successes, and shortcomings.

In addition to the performance variables, the current work investigated a variety of factors that contribute to the resulting experience with the environment and impact perceptions of performance (RQ2). Specifically, people who enjoyed the experience really enjoyed the playful nature of being awkward (as hypothesized during design). Additionally, those people who understood how to play (i.e., intent was to be awkward), also reported that they played the game to the best of their ability (and, in fact, they did have higher game performance scores). It appears that once people understood the game was designed as an escape from reality where they should drop social norms to explore inappropriate interactions, it afforded much more positive and productive experiences (RQ3).

In terms of the Awkward Annie GBA, these results suggest revisions to certain key areas targeting the clarity of game navigation and pragmatically appropriate content will need to be addressed. Based on the findings above (understanding gameplay and user experience), future efforts will empirically address how choosing the inappropriate (vs. appropriate) response impacts aspects of engagement and the measurement of pragmatic ability. Additionally, the novelty, enjoyment, and impact of selecting (in)appropriate options may be dependent upon the timescale of interaction (short or long) [18].

Another aspect of the game that will merit further refinement is the conversation options. Although the Awkward Annie pragmatics content was developed based on prior research [19], qualitative, offline feedback suggested that some users found it difficult to choose between the mildly and most inappropriate responses. These users noted that both options were often

awkward enough to be considered the “correct” choice. While this difficulty may have affected the average score per conversation, it should not have significantly impacted overall conversation success rates. Additional efforts are now underway to refine and norm the game content.

Despite these potential limitations, and the tentative nature of the conceptual model, the current work with Awkward Annie has illustrated some of the potential for the environment and provided initial support for this unique approach. This work has identified several key factors and constructs that may have significant impact on game design, game-based assessment practices, and measures of pragmatic ability.

REFERENCES

- [1] Bardovi-Harlig, K. (2013). Developing L2 pragmatics. *Language Learning*, 63(1), 68-86.
- [2] Limberg, H. (2015). Principles for pragmatics teaching: Apologies in the EFL classroom. *ELT Journal*, 69(3), 275-285.
- [3] Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE)*, 1(1), 20-20.
- [4] Riddiford, N., & Joe, A. (2010). Tracking the development of sociopragmatic skills. *TESOL Quarterly*, 44(1), 195-205.
- [5] Clyne, M. (1994). *Inter-cultural communication at work: Cultural values in discourse*. Cambridge, UK: Cambridge University Press.
- [6] Timpe, V. (2013a). *Assessing intercultural language learning*. Frankfurt a.M.: Peter Lang.
- [7] Belz, J. A. (2007). The role of computer mediation in the instruction and development of L2 pragmatic competence. *Annual Review of Applied Linguistics*, 27, 45.
- [8] Taguchi, N. & Roeber, C. (2017). *Second language pragmatics*. Oxford: Oxford University Press.
- [9] Sykes, J. & Reinhard, J. (2012). *Language at Play: Digital Games in Second and Foreign Language Teaching and Learning*. Series on Theory and Practice in Second Language Classroom Instruction, J. Liskin-Gasparro & M. Lacorte, series eds. Pearson-Prentice Hall.
- [10] Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & gaming*, 33(4), 441-467.
- [11] Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.
- [12] Woo, Y., & Reeves, T. C. (2007). Meaningful interaction in web-based learning: A social constructivist interpretation. *The Internet and higher education*, 10(1), 15-25.
- [13] Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge university press.
- [14] Hoffman, E., John, M. & Makany, T. (2014). How do game design frameworks align with learning and assessment design frameworks? *Paper presented at the annual meeting for the National Council for Measurement in Education*, Philadelphia, PA.
- [15] Timpe-Laughlin, V., Wain, J., & Schmidgall, J. (2015). *Defining and operationalizing the construct of pragmatic competence: Review and recommendations*. ETS Research Report Series. Princeton, NJ: Educational Testing Service.
- [16] Cole, H., & Griffiths, M. D. (2007). Social interactions in massively multi-player online role-playing gamers. *CyberPsychology & Behavior*, 10(4), 575-583.
- [17] Jackson, G.T., Grace, L., Inglese, P., Wain, J. (submitted). Awkward Annie: Game-based assessment of English pragmatic skills. *Submitted to 2017 CHI Play conference*.
- [18] Jackson, G.T., & McNamara, D.S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology*, 105, 1036-1049.
- [19] Timpe, V. (2013b). The difficulty with difficulty: The issue of determining task difficulty in TBLA. *Journal of Linguistics and Language Teaching*, 4(1), 13-28.