Instructor accents in online education and their effect on learning and attitudes

C. A. Sanchez & S. Khan
Oregon State University, Corvallis, USA

Abstract

Reductions in perceptual fluency have been shown to negatively impact attitudes towards learning material, but not learning itself. The current study extends this work to spoken presentations and examines whether the presence of a foreign accent negatively affects learners’ experience in an online learning environment. Results indicate that the presence of an instructor accent, consistent with prior work on perceptual fluency, does not impact learning, but does cause learners to rate the instructor as less effective. Further, for those who received accented presentations, changes in participants’ attitudes towards both the content area and online instruction were not predicted by learning, but instead their attitude towards the instructor. This suggests that learners in online learning environments with accented narration are potentially miscalibrated, and these biases in judgement could be inappropriately linked to a specific instructor, rather than their success of learning in the field.

Keywords

learning, narration, online instruction, perceptual fluency.

The explosion of online education (OE) opportunities in the past decade is founded on the somewhat fundamental assumption that learning via online instruction provides an instructional experience that is equivalent to more ‘traditional’ classrooms, but deftly sidesteps some of the typical constraints encountered when pursuing higher education. For example, learners are able to receive instruction on their own terms and schedule (often at a much lower or even no cost), that is arguably equivalent to (or better than) what might normally be available to them at their local community college or university.

Meta-analytic research has confirmed that OE options produce at least an equivalent learning experience compared to face-to-face (FTF) classrooms, dubbed the ‘no significant difference’ phenomenon (Bernard et al., 2009). However, a fixation on learning outcomes alone is perhaps short-sighted and potentially produces an overestimation of the true value of such OE. For example, it is also worth considering student attitudes towards learning (both within and outside the content area), when instruction is delivered in these online formats. As a case in point, students have reported issues of isolation or frustration when taking online instruction (Ku & Lohr, 2003; Reilly, Gallagher-Lepak, & Killion, 2012) which may underlie some of the more negative feelings found towards OE versus their FTF counterparts (Johnson, Aragon & Shaik, 2000; Lin, Chen & Fang, 2011). These findings, at the very least, support the notion that something about the online delivery of education negatively impacts learner attitudes at multiple levels and should be considered further.

In regards to learning from OE, there does not appear to be a consistent observable negative consequence for OE. An obvious starting point in this effort to understand potential shortcomings of OE would be to examine the manner in which online content is disseminated, and how basic characteristics of this delivery method can affect student attitudes. Perhaps one of the most common formats of online delivery (in both formal and informal learning settings) involves a narrated video or lecture that learners are able to download and watch at their convenience. Note that there does exist a continuum of
perceived connection to the instructor within this type of OE. At one end, this mode of delivery consists of a videotaped lecture of the instructor teaching in front of a group of students (analogous to a FTF classroom), providing at least some of the traditional classroom ‘feel’ to online learners. On the opposite end (although likely just as common), the instructor merely narrates a slide presentation or demonstration that moves through the material, more or less delivering the content in a semi-anonymous state. In other words, content is the focus in this sub-type of narrated OE, with little to no visual reference to the instructor themselves. In the most extreme sense this might consist of no visual reference to the instructor, or in slightly more moderate cases perhaps a non-dynamic headshot, that although identifies who might be speaking, does not explicitly connect the presenter to the actual presentation at hand. For example, the learner cannot see facial expressions or other social cues in these cases, nor can they necessarily verify that the static picture presented is in fact the actual speaker! Across the entire spectrum however, the primary means with which the instructor communicates with the learner is through (1) auditory feedback in the form of narration, and (2) any visuals or graphics that are included.

While it is certainly possible to draw some parallels between these types of narrated online instruction to more traditional classroom instruction (at least in regards to instructional content), it must also be realized that because of the inherent asynchronous and non-dynamic nature of online instruction in general, potential learners are limited in their exposure to the instructional experience by what is included in the downloaded lesson, and no more. Simply, it becomes imperative to understand how otherwise innocuous characteristics of a lesson in a traditional classroom (i.e., aspects of narration) become more central in OE. For example, it is not unreasonable to expect that in a large number of cases, narration provides the strongest (and sometimes sole) connection students have to the instructor or peers in online learning environments. Thus, there is little opportunity for students in these OE environments to tailor instruction to their own needs, and engage in some of the more social aspects that have been identified as critical for building knowledge (Powell & Kalina, 2009). The question then is: as narration is an often critical component of OE, can simple changes in the perceptual characteristics of the narration otherwise influence learner attitudes, such that it impacts the educational experience at large?

Support for this question can first be drawn from research that has explored changes in simple perceptual characteristics of text and how they affect learner attitudes. In these experiments, perceptual fluency is manipulated by making a text more difficult to read by visually changing characteristics of the typesetting, but not the actual information in the text itself. Importantly, the term ‘fluency’, as used in the literature reported here, does not reference a linguistic or conceptual characteristic of the text, but instead references the ease with which information can be perceptually decoded. By this definition, a text that is less perceptually fluent or ‘disfluent’, should require more effort to perceptual decode the textual information, although all other aspects of the text are still completely legible and intelligible including grammar, syntax, etc. What this means is that although both fluent and disfluent textual presentations can be read and understood (and thus neither are ‘disfluent’ using the more traditional definition of the word), less perceptually fluent texts are slightly more arduous to process, but only at a very low perceptual level. This is perhaps best demonstrated with a simple explicit example.

As mentioned above, there is some evidence that the simple manipulation of typesetting from a more perceptually fluent (i.e., Times New Roman) to a less perceptually fluent font (i.e., Mistral) can have negative impacts on attitudes towards information described in text, even though readers demonstrate equivalent levels of understanding (Alter, Oppenheimer, Epley & Eyre, 2007; Sanchez & Jaeger, 2015; Song & Schwarz, 2008). Thus, although readers do comprehend the texts equally, when asked to report responses to questions like ‘How easy was this text to read?’, those receiving perceptually disfluent presentations respond much more negatively. Critical to the point here is that this negative appraisal of the simple textual decoding process also bleeds over and affects attitudes towards the conceptual information in the text. For example, participants who received a less perceptually fluent text also reported that the procedural task described in the text (making a Sushi roll) would take more time to complete and/or be more difficult to construct, than other participants who received the same exact text, albeit in a more perceptually fluent and easier to read font (Sanchez & Jaeger, 2015; Song & Schwarz, 2008). Further, other research examining perceptual fluency also suggests that perceptually disfluent text presentations also reduce the ability of learners to accurately gauge or predict their learning (Rawson & Dunlosky,
2002; Rhodes & Castel, 2008; Undorf & Erdfelder, 2015). In sum, it appears that changes to the perceptual experience of a text that make perceptual decoding more difficult can have severe implications on not only attitudes towards the information in the text, but also readers’ evaluation of how well they have learned the information. Importantly, this phenomenon has not been limited to text processing alone, and has also been observed in relation to other visual and perceptual characteristics such as visual aesthetics, branding and auditory generation tasks (Sobko & Mautone, 2003). This further suggests that any manipulation that affects perceptual fluency, in any perceptual modality, may produce changes in attitudes or judgements of the to-be-learned information.

Specific to the potential relationship between issues of narrated instruction and perceptual fluency, Mayer, Sobko, and Mautone (2003), investigated how the presence of a foreign accent (Russian) impacted learning of a science lesson on lightning. While individuals who had an accented multimedia lesson learned at the same level as their non-accented counterparts, they did rate the instructor as harder to understand and also significantly lower in other social aspects (e.g., pleasantness, friendliness, etc.). This is also similar to recent work done by Carpenter, Wilford, Kornell, and Mullaney (2013) who manipulated the lecture fluency of a video science lesson by changing instructor behaviour, for example by reducing eye contact or reading in a halting, awkward manner. Again, while students learned equivalently in both fluent and less perceptually fluent video lectures, students given the less fluid presentation were less metacognitively aware of their own knowledge states, and also rated the instructor as being a less effective teacher.

The above results suggest that learners do pay attention to aspects of the instructional narration, and while less perceptually fluent presentations do not reduce learning, they can cause learners to feel worse about both their own understanding and the instructor. Across both studies, however, the change in judgements (whether they be social ratings or rating of instructional effectiveness) was contextualized within the lesson at hand. Importantly, it must be explored how these less perceptually fluent presentations might also impact ratings outside the direct context of the lesson itself, for example by examining more general attitudes towards the task at hand, or perhaps even the method of delivery itself (i.e., OE broadly construed). This next question is important as instructional lessons often exist within a larger context, both within the content domain and pedagogically. If learners are less able to make accurate judgements about the connection between instructor characteristics and learning, and also how well they might learn in the future, this could be problematic for the attainment of educational goals in general.

Further, if attitudes relative to a specific educational experience or instructor resonate through the educational process, this might cause students to harbour attitudes towards the material or field that might be mal-adaptive to their own success. For example, students may be less likely to persist in a given major if negative judgements carry-over to other classes or content areas, thus fundamentally disconnected from their ability to actually learn the material.

To explore these questions of how simple perceptual differences in narration in OE might affect learners, a single study was conducted where students watched a lesson on programming that was narrated by either a fluent Native English speaker, or instead a non-Native English speaker that spoke with an accent. This presentation consisted of a video demonstration (without visible presence of the instructor) of a simple lesson on how to use a programming tool. Participants were evaluated on their learning of the lesson, and their attitudes towards the instructor and the instructional quality. Further, participants were also asked to evaluate their attitudes towards both programming and online instruction in general, in an effort to see whether exposure to this accented narration would predict changes in these more global attitudes.

Method

Participants and design

One-hundred ten (N = 110) undergraduates from a large public university in the western United States were solicited for participation. All participants were Native English speakers and compensated with course credit in an Introductory Psychology Course. However, if participants reported any previous programming experience, they were not retained for later analysis. This strict criterion ensures that prior knowledge levels do not unduly influence performance in the learning or attitude assessment tasks. Nine participants reported programming
experience and were thus omitted, and the remaining 101 participants were randomly assigned to one of two experimental conditions. Forty-nine (69.4% female; average age = 20.27 years) individuals watched an instructional video that contained narration with no accent, and the remaining 52 (69.2% female; average age = 19.65 years) individuals watched the same instructional video, but with accented narration.

Materials

Instructional video
All participants watched an instructional video that was 5 m and 43 s in length and explained how to create a simple animation using Scratch (ver. 1.4; Figure 1). Scratch 1.4 (Lifelong Kindergarten Group at MIT Media Laboratory, 2009) is an educational tool developed in the hopes of giving young novice programmers an effective and positive experience with object-oriented programming, thereby reducing some of the negative stigma associated with programming at large. Based on condition, this video was either narrated by a Native English speaker, or a non-native English speaker with an accent (Mandarin Chinese). Narration was identical across these conditions in terms of both content and timing, with the only difference being the presence of a non-native accent in one condition. Both narrators were also adult males and had no other relationship, instructional or otherwise, with the participants. All visuals were identical across conditions, and the narrator was never presented in any visual capacity to the participants. Participants listened to the narration via headphones, and watched the videos in full-screen with a resolution of 1600×900 on individual computer monitors.

Pre-test
Prior to watching the video, participants completed a pre-test that evaluated their age, gender and whether or not they had programmed before (including Scratch). Participants also provided ratings for two questions designed to assess their attitudes towards (1) online instruction and their estimation of (2) how difficult programming is. Attitudes towards online instruction and programming were evaluated on a scale from 1 to 8 (1 being lowest). For the online instruction measure, 1 represented ‘ Extremely Dislike’ and 8 represented ‘ Extremely Like’, while on the programming difficulty measure, 1 represented ‘ Not Hard’ and 8 represented ‘ Very Hard’.

Figure 1 Screenshot of the Video Used to Instruct the Scratch Programming Interface
Post-test
The post-test consisted of three sub-components. The first two sub-components evaluated how well participants learned the information contained in the instructional video, while the third component again assessed participant attitudes.

Matching. The first component required participants to identify, and then match, information relative to how to use the programming tool. Eight locations were identified on a screen print, and participants were required to match one of the eight locations to one of the six described functions. For example, participants were asked to identify where the ‘Action Window’ was located, and in order to respond correctly would be required to select the corresponding letter that identified the Action Window. This measure was designed to evaluate how familiar participants were with using the interface on a very basic level. In other words, did individuals know where to locate the basic functionality of the system? The maximum score on this measure was 6.

Multiple choice. In this portion of the post-test, participants were asked to complete three multiple choice questions, designed to evaluate how well the participants knew how to use the interface respective to the specific programming task at hand. For example, participants were asked what specific command would be used to ‘animate’ the project as it moves. Participants were required to select the correct response from four potential responses, and were only allowed to select one answer. Thus, while the Matching test evaluated where participants would need to go to find commands, this simple Multiple Choice component attempted to evaluate whether participants knew how and what command to use to programme the current task.

Attitudes. In this final section of the post-test participants were asked to make ratings on a 1–8 (1 being lowest) scale for their attitudes towards the instructor across two questions (i.e., how easy to understand, quality of instructor), and again their attitudes towards online classes and the difficulty of programming. Questions evaluating attitudes towards online classes and programming were identical to the pre-test to evaluate any shift in participant attitude based on the instructional experience.

Procedure
Participants were first given 3 min to complete the pretest measure. After completing the pretest, participants were then asked to view the instructional video. Participants were given 8 min to view the video, and were allowed to rewind and replay the video as they saw fit during that time. All participants were required, however, to watch the entire video, and not skip any portion of it. After the 8 min was up, participants were then given 8 min to complete the Post-test, after which they were debriefed and dismissed. The entire experiment took no longer than 30 min.

Results
To better capture learning and instructor attitudes, composite scores for these metrics were first calculated. Scores on the Matching and Multiple Choice post-test measures (Pearson’s $r = .30$, $p < .01$) were averaged together, and the ratings on the 2 questions assessing attitudes towards the instructor (ease and quality; Pearson’s $r = .63$, $p < .01$) were also averaged together. Prior to averaging, all measures were first standardized on a z-distribution, and the resulting z-scores were used to compute the subsequent composite measures. These composite scores were calculated to minimize any variance unique to the specific measures themselves, and thus hopefully provide a more consistent and valid estimation of participant learning and attitudes towards the instructor. Importantly, as is visible in Table 1, the accent manipulation did have a significant impact on participant ratings of how easy the instructor was to understand, confirming that even the mild accent used here was capable of producing a less perceptually fluent narrative experience. Descriptive statistics for all measures, and pre-post comparisons, are available in Table 1.

Learning and attitudes towards instructor
To evaluate whether narration conditions varied in terms of learning or attitudes towards the instructor, separate simple two-way ANOVAs were conducted comparing the different narration groups across the learning and attitude measures.

In terms of learning, there was found to be no effect of narration condition, and participants in both groups were equally adept and understanding how to use the programming tool ($F(1, 99) = .30$, $p > .05$, $\eta^2_p = .00$). However, while learning was consistent, did the presence of a simple accent change attitudes towards the instructor?
It does appear that there was a significant difference between narration conditions in regards to attitudes towards the instructor ($F(1, 99) = 9.19, p < .01, \eta^2_p = .09$). Participants in the accented video condition viewed the instructor in a significantly more negative light than participants in the non-accented condition.

Taken together, these results suggest that even though participants did not learn differently across the different narration conditions, participants with accented narration felt the instructor was not only more difficult to understand, but also of lower quality (Table 1). Thus, even though participants obviously did not have more difficulty understanding the content of the instructional video in the accented condition (as evidenced by no difference in learning), participants in this condition perceived the instructor as worse, simply because of the accent. These results are captured in both Figure 2 and Table 1.

### Attitudes towards online instruction and programming

Participant attitudes towards online instruction and programming were also evaluated across conditions. A 2 (accent condition) × 2 (pre-post) mixed ANOVA was conducted on both attitude measures. There did appear to be a significant improvement pre-post regarding both online instruction ($F(1, 99) = 28.70, p < .01, \eta^2_p = .23$) and programming ($F(1, 99) = 142.59, p < .01, \eta^2_p = .59$). After viewing the video, all participants were significantly more positive towards both online instruction and programming. However, there was no main effect of narration condition for either of these measures, and also no interaction between narration condition and pre-post evaluation. This suggests that while participant attitudes did change for the better, the magnitude of this change was not different across narration conditions.

However, while these attitudes towards referents that exist outside of the task did change for both groups in an equivalent fashion, of interest was whether these changes in attitudes were based on similar predictors. In other words, it is important to understand what is producing this change in these general attitudes, and whether this change is predicted in the same way for all narration conditions. For example, is learning the key predictor of this change in attitudes, or are instructor

| Table 1. Mean(SD) and Simple Inferential Statistics for the Non-Accented ($n = 49$) and Accented ($n = 52$) Video Conditions for All Reported Measures |
|-------------------------------------------------|-----------------|-----------------|
| Age                                             |    Non-accented |        Accented |      $t(df = 99)$ |
| How hard do you think programming is?            |                |                 |                  |
| Pretest                                          |   20.27(3.41)  |  19.65(1.47)   |   1.18           |
| Posttest                                         |   6.61(1.00)   |  6.38(1.46)    |   0.91           |
| Rate your attitude towards online instruction/classes. |         |                  |                  |
| Pretest                                          |   4.66(1.31)   |  4.66(1.49)    |   0              |
| Posttest                                         |   5.27(1.38)   |  5.46(1.49)    |  -0.69           |
| Matching                                         |   4.10(1.09)   |  4.21(1.02)    |  -0.55           |
| Multiple Choice                                  |   2.65 (.66)   |  2.69(.51)     |   0.33           |
| How easy was the instructor to understand?      |   6.90(1.49)   |  5.90(1.71)    |  3.11**          |
| Please rate the quality of the instructor.       |   6.71(1.38)   |  6.00(1.67)    |  2.33*           |

*p < .05.

**p < .01.
characteristics the basis of this change? Or are both equally relevant?

To evaluate these questions, a difference score was computed by subtracting participant ratings regarding programming and online classes on the post-test from participant answers to the same questions on the pre-test. Both learning and attitudes towards instruction were then used to predict the change in scores in the non-accented and accented condition using a hierarchical linear regression. Learning was entered in the first block, and attitudes towards instructor were entered in the second block.

What predicts changes in attitudes towards programming? In the non-accented narration condition, results from the first block of analysis indicate that participant learning ($\beta = -.35, p < .01$) did significantly predict the positive change in attitudes towards programming ($R^2 = .12, F(1, 48) = 6.45, p < .01$). The addition of attitudes towards instructor in the second block did not predict a significant additional portion of variance ($R^2$ change = .03, $p > .05$), suggesting that attitudes towards instructor did not relate to changes in attitudes towards programming in this group. However, in the accented narration condition, the exact opposite pattern of results was observed. Results from the first block of analysis demonstrated that learning ($\beta = -.08, p > .05$) alone did not predict changes in attitudes towards programming ($R^2 = .01, F(1, 50) < 1, p > .05$), and instead, attitudes towards instructor ($\beta = .36, p < .01$) was a significant predictor in the second block ($R^2$ change = .13, $p < .01$; $R^2 = .13, F(2, 49) = 3.77, p < .05$).

Thus, in the non-accented condition, attitudes towards the general type of task instructed here were predicted by how well participants learned the task, independent of their opinion of the instructor. Whereas, in the accented condition, these participants are instead basing their change in attitude judgements towards programming on their opinion of the instructor, and not their ability to learn the material. This represents a potential miscalibration of learners in this accented condition, as one would expect that general attitudes towards a given task should be predicted most by prior success in learning. This was not the case in the accented condition.

What predicts changes in attitudes towards online courses? Results from the non-accented narration condition indicated that neither learning, nor attitudes towards instructor, were predictive of changes in more general attitudes towards online instruction ($R^2 = .07, F(2, 46) = 1.80, p > .05$). However, in the accented narration condition, results mirrored those of the attitudes towards programming above. Learning again did not predict change in attitudes towards online classes ($R^2 = .03, F(1, 50) = 1.43, p > .05$); however, the addition of attitudes towards instructor ($\beta = .39, p < .01$) in the second block predicted a significant portion of additional variance ($R^2$ change = .15, $p < .01$; $R^2 = .18, F(2, 49) = 5.25, p < .01$).

This pattern of results suggests that learning did not predict more general changes in attitude towards online instruction in either condition. This is perhaps not surprising as OE is a fairly separate issue from the direct content instruction provided here. In other words, while OE is the method of delivery, it was not emphasized as a primary focal point in the current lesson. However, despite the conceptual separation of OE from the content here, those participants that experienced accented narration relied on their attitudes towards the instructor as a basis for this change in attitudes, rather than their success in the current context.

Discussion

Results from the current study suggest that the presence of a non-Native accent can negatively impact instructor ratings, but does not seem to impact learning from online environments. These findings are consistent with other work that has examined how perceptual fluency can impact instructor appraisal (Carpenter et al., 2013; Mayer et al., 2003). It also is consistent with other work outside this theoretical perspective that has found that students do rate instructors with accents as less effective teachers (Boyd, 2003; Rubin & Smith, 1990). However, the current work further builds upon these results by examining how such presentations also affect attitudes towards referents external to the task-at-hand. Results suggest that changes in attitudes towards the domain task (i.e., programming) and the generalized delivery method (i.e., online classes) are differentially predicted based on the presence of accented narration. In those cases where non-accented narration is included, changes in
attitudes towards programming are predicted by learning, but not by instructor appraisal. However, in the accented narration condition, changes in attitudes towards programming and online coursework are instead predicted by attitudes towards the instructor, and not learning. This suggests that learners who are exposed to accented presentations are basing more general, de-contextualized, attitudes on their appraisal of a given instructor, instead of their ability to learn the material. This finding is somewhat concerning as it suggests that these learners are potentially at risk for a miscalibration of attitude towards their educational goal. Realistically, success at learning the material should be one of the, if not the sole, primary anchor point(s) used for judgement towards the more general task at hand. This was not the case here, as learners with accented narration are perhaps fixating on the instructor characteristics to an unhealthy degree, thus basing their judgements (and perhaps likelihood to continue with such studies) on a foundation that is likely less relevant to their future achievement in a given area. In other words, by using the attitudes towards a given instructor as a basis for attitudes towards the general task or field, this may put individuals at risk for making a decision that is not reflective of their ability to succeed in the field, but instead based on a more subjective transient reaction to a single instructor.

This work thus further highlights the importance of instructional characteristics when providing educational opportunities through online delivery methods. Aspects of narration do appear to be critical for some aspects of student development, and this result is perhaps not surprising given that this auditory feedback is often the only connection distance learners have to an actual instructor. Importantly, it must be recognized that such educational exposures can resonate through the entire educational experience, and potentially produce less desirable, but unattended, changes in students.

Similarly, this work also highlights a potential disparity in evaluation for instructors with non-native accents versus instructors with no perceived accent. As a critical evaluation mechanism for faculty performance is student ratings, this research (and others) suggests that faculty with accents, although they are likely just as effective at teaching students as any non-accented instructor (as demonstrated here), are perceived as less effective than they truly are simply because of the way their speech is articulated. By better recognizing the shortcomings and issues related to the student evaluation of instruction, it may be possible to more effectively evaluate all instructors, regardless of how they speak, and instead, more appropriately, by how well they teach.

Future work should further explore how different levels of accented narration either magnify or attenuate such effects. The accented narration used here was provided by a linguistically fluent speaker with what would likely be termed a mild to moderate accent. Given the cosmopolitan nature of most current higher education institutions, it is certainly worth exploring how more extreme accents affect the current pattern of findings. Further, it would also be of interest to explore how prior experience, either with the content or accents in general, might also modify the current pattern of findings. As this study was conducted on low-knowledge, native-English speakers, it does potentially neglect other critical interactions with existing knowledge that might produce a different pattern of results.

Finally, it is also of interest as to whether there are additional compensatory mechanisms or tools that can be leveraged to minimize such observed effects. For example, perhaps the addition of subtitles or other cues might serve to reduce such negative effects on student attitudes, and the inclusion of such simple characteristics might produce a more effective teaching tool, in terms of both learning and student attitudes.

References


