ClientComm Data Analysis 4

Gwen Rino and Eric Giannella

8/6/2018

OVERVIEW AND KEY FINDINGS

These are the results of our investigation of the question: What characteristics of initial messages sent by Baltimore POs are associated with clients replying?

Answering this question is important because clients who participate in two-way texting relationships with their POs are significantly less likely ($p < .01$) to experience supervision failure than those who don't. Furthermore, among clients who ever send a text to their PO, the majority (54%) send their first one in response to the initial message they receive.

We studied 3342 initial PO messages, of which 889 received responses and 2453 did not, and we discovered a number of characteristics that are associated with client response rates.

Key Findings

Initial message characteristics that are significantly and positively associated with client responses:
- A request for a response to the message
- A request for employment verification or other documentation
- A reminder of the date of an upcoming appointment (but not an upcoming court date)
- A greeting (but not a closing)
- The PO's name
- Sent in the afternoon (vs. morning, evening, or night)

Initial message characteristics that are positively (but not significantly) associated with client responses:
- "Polite" language (please, thank you) in individualized messages
- The client's name

Initial message characteristics that are significantly and negatively associated with client responses:
- Messages that consist mostly of generic (vs. individualized) language
- "Problem" language (violation, warrant, compliance, failure)
- "Urgent" language (ASAP, immediately, right away, imperative)
- "Polite" language (please, thank you) in generic (vs. individualized) messages

Initial message characteristics that are negatively (but not significantly) associated with client responses:
- Messages about office closures or other generic issues
- "Yelling" (all caps)

ANALYTICAL STORY

In general, clients are more likely to respond to initial PO messages that are individualized than to messages that are generic. Not surprisingly, clients rarely respond to texts that are clearly mass messages, such as the reminder, "For all who have court this week, please be on time." Individualizing a text increases the chance that the client will respond, and the more individualized the text is, the more likely the client is to respond. Messages can be individualized enough to affect the client response rate just by including a future appointment date that is specific to the client or by including the PO's and/or client's name. Even modest personalizations like these can increase the client response rate by as much as 18%.

PO messages that request something of the client are significantly more likely to receive a client response than those that don't. Clients are 26% more likely to respond to a message that includes a request for a paystub or other document, and 36% more likely to respond to a message that directly requests a reply, than

to messages that do not include a request. This PO text demonstrates what a request can look like (it is also individualized with the client's name, the PO's name, and an appointment date):
- "Good Afternoon Mr. Harris. This is Agent Stephens with Pretrial Release Services. This is to remind you of our conversation regarding your report date for Intake on January 22, 2018 at 8:30 am at the Clarence Mitchell Courthouse Room 442. Please bring a photo ID and most recent paystubs or documentation of employment. Please confirm this text."

Urgent or problem language is strongly negatively associated with client responses, and "yelling" (writing in all caps) is also negative for client response. These message characteristics can reduce the client response rate by as much as 35%. Perhaps this is because the initial message sets the tone of the relationship. A client may feel less inclined to begin a texting relationship in a context that seems accusatory, negative, or hopeless from the beginning. The following PO texts have characteristics that are negatively associated with client response:
- "This is your pre-trial agent. If you have received this message that means you have not reported for pre-trial today. You need to report ASAP in order to stay in compliance."
- "FROM AGENT ROGERS, MR. GRANT, YOU ARE INSTRUCTED TO REPORT IN PERSON IN ORDER TO BRING YOUR RECORDS UP-TO-DATE. FAILURE TO DO SO MAY RESULT IN THE ISSUANCE OF A BENCH WARRANT FOR YOUR ARREST. YOU MUST REPORT IN ON 12/28/17, BY 10:00 AM."

In contrast, polite language and greetings such as "Good morning" are message characteristics that are positively associated with client response. Perhaps clients are more inclined to begin a texting relationship when they experience the initial message as respectful and supportive of their success. Importantly, the positive effect of polite language is not as strong as the negative effect of generic messages. Only individualized polite messages are positively associated with client responses, while mass messages, even polite ones, are strongly negatively associated with client responses.

METHODOLOGY

Data

The data used in this analysis is all the initial outbound messages sent by Baltimore ClientComm users between (dates).

Independent variables

With two exceptions, all of the independent variables are binary indicators built using pattern-matching functions to locate words and phrases we identified as indicative of a particular message characteristic by hand-coding hundreds of user messages.
The exceptions are:
- send_at_ToD_bins identifies the time of day that the PO message was sent at one of four levels (morning, afternoon, evening, night).
- max_reuse_score is a continuous variable that quantifies how generic or individualized a message is by comparing it to all other user messages using the function stringdist::stringsim. A message's max_reuse_score is its highest similarity score, where a score of 0 indicates a message that is completely unique within the dataset and a score of 1 indicates a message that is exactly the same as another message within the dataset.

There is one interaction between variables included in the model, max_reuse_score:polite. This interaction clarifies that politeness is positively associated with client response only in messages that are individualized; the positive influence of politeness is overridden by the negative influence of generic messages.

Dependent variable

The dependent variable is a binary indicator of whether or not the PO message received a client response.

Model

We built a generalized linear mixed model to estimate the influence of various message characteristics on the likelihood of a client replying to the message. The model holds the POs as random effects in order to focus on the influence of the message characteristics as fixed effects.

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## msg_replied_i ~ (1 | PO) + court_date_reminder + appointment_date_reminder +
##     pls_respond + business + max_reuse_score + info + problem +
##     urgency + polite + greeting + yelling + has_client_name +
##     has_user_name + polite:max_reuse_score + send_at_ToD_bins
##    Data: initial_msgs_qualities
## Control: glmerControl(optimizer = "bobyqa")
##
##      AIC      BIC   logLik deviance df.resid
##   3678.9   3794.9  -1820.4   3640.9     3296
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.3543 -0.6089 -0.4979  1.0254  3.5202
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  PO     (Intercept) 0.1607   0.4008
## Number of obs: 3315, groups:  PO, 23
##
## Fixed effects:
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)                -0.72879    0.31914  -2.284 0.022396 *
## court_date_reminder        -0.10553    0.14523  -0.727 0.467436
## appointment_date_reminder   0.36686    0.10581   3.467 0.000526 ***
## pls_respond                 0.73625    0.30932   2.380 0.017303 *
## business                    0.46717    0.18151   2.574 0.010060 *
## max_reuse_score            -0.55716    0.34347  -1.622 0.104769
## info                       -0.44247    0.25758  -1.718 0.085835 .
## problem                    -0.42238    0.17356  -2.434 0.014947 *
## urgency                    -0.57112    0.28370  -2.013 0.044106 *
## polite                      0.50384    0.38180   1.320 0.186954
## greeting                    0.21916    0.12238   1.791 0.073314 .
## yelling                    -0.37796    0.33715  -1.121 0.262274
## has_client_name             0.06302    0.13872   0.454 0.649641
## has_user_name               0.34305    0.13350   2.570 0.010178 *
## send_at_ToD_binsNight      -0.13964    0.39950  -0.350 0.726685
## send_at_ToD_binsAfternoon   0.15784    0.08798   1.794 0.072813 .
## send_at_ToD_binsEvening    -0.02635    0.21260  -0.124 0.901364
## max_reuse_score:polite     -1.23136    0.45678  -2.696 0.007023 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 18 > 12.
## Use print(x, correlation=TRUE)  or
```

```
##   vcov(x)     if you need it
```