

RE: ACLU Matter vs. - REF# 1340012232
Analysis of Multiple Version ISR Records
and Coded Narratives
July-December 2016
For Input to
Hon. Arlander Keys' (Ret.)
Second Period Report

**Final
Technical Report**

Ralph B. Taylor, Lallen T. Johnson, Joshua M. Koehnlein and Danielle T. Stanford

DATE: December 12, 2017

Acknowledgments. All the material herein represents only the views of the authors and does not reflect the views or policies of any other organization including the City of Chicago, the Chicago Police Department, or ACLU-Illinois, or their experts. Any mistakes or misinterpretations herein are solely the authors.

Declaration of Conflicting Interests. The authors declare no potential conflicts of interest with respect to the research, authorship and/or dissemination of this work.

Funding. The current work was funded through a contract from the City of Chicago Department of Law to Temple University, with a subcontract to Drexel University.

1 CONTENTS

2	EXHIBITS	4
3	Executive Summary	5
4	Background, Purpose, and Sources.....	5
4.1	Background	5
4.2	Purposes and sources.....	5
4.2.1	Questions about all multiple version Investigatory Stop Reports (ISRs)	6
4.2.2	Questions about sampled Investigatory Stop Reports	6
5	Data processing.....	7
5.1	Full set of multi-version ISRs, with different versions stacked within individual reports and demographic information added	7
5.2	Last version file of multi-version ISR records.....	7
5.3	Sampling equal numbers of multi-version ISRs for each ethnoracial group	7
5.4	Coded legal narratives of multi-version sampled records.....	8
5.5	Sample weights	10
5.6	How many multi-version ISRs are there in Period 2 (July-December 2016)?.....	10
5.7	Coding manifest content from the rejection message field.....	10
5.7.1	Focus	10
5.7.2	Limitations	11
6	Descriptive information: the full set of multi-version ISRs.....	11
6.1	distribution across ethnoracial groups.....	11
6.2	District differences in relative contributions of multiple versus single version ISRs	11
7	Number of versions.....	14
7.1	Cautionary note	14
7.2	How many multi-version reports got revised how many times and where?	16
8	Predicting how many report versions among multi-version ISRs	17
8.1	The Question	17
8.2	Approaches.....	17
8.3	Detainee race, ethnicity, and gender	18
8.4	Beat-level ethnic and racial composition	18
8.4.1	Links	18
8.4.2	Robustness questions	20
8.4.3	Summary of number of versions and race and ethnicity of beat.....	22

9	Examining reasons for rejection	23
9.1	Preliminary content analysis	23
9.1.1	Limitations	24
9.1.2	Selecting key text.....	25
9.1.3	Descriptive statistics for manifest content coding variables.....	25
9.1.4	Links between manifest content and race and ethnicity of detainee and locale	27
10	Overall Assessments of Multi-Version Coded Records	29
10.1	Preface	29
10.1.1	Question sequence and relevant cases: implications for examining differences	29
10.1.2	Reports on weighted and unweighted data	29
10.2	Error involves RAS (Question 20)	29
10.2.1	Frequency of concern.....	29
10.2.2	Patterning across race or ethnicity of detainee or beat	30
10.2.3	Officer correctly articulates RAS (Q21).....	32
10.2.4	Reviewer flags RAS error.....	32
10.2.5	Sending the ISR back (any reviewer)	33
10.2.6	Officer remediation.....	33
10.2.7	Changing RAS	33
10.2.8	Summing up on RAS concerns	34
10.3	Final status codes.....	35
11	Discussion.....	35
11.1	The novelty issue	35
11.2	Broad outline	35
11.3	Racial links with number of versions	36
11.4	Keywords in rejection messages.....	36
11.5	Legal coding and related processing concerns	37
11.6	Changes in other fields	37
11.7	Closing comment.....	38
12	References.....	38
13	Appendix A: Stata programming for capturing some basic phrases from rejection message field 39	
14	Appendix B: coding instrument used for multiple version investigatory stop reports	41

2 EXHIBITS

Exhibit 1 Chicago beats-within-districts 70 percent or more Black and non-Hispanic.....	9
Exhibit 2 Distribution of multiple-Version ISRs across districts and ethnoracial groups within districts.....	13
Exhibit 3 Contrasting district multiple version vs. single version shares	14
Exhibit 4 Numbers of versions of multi-version ISRs.....	16
Exhibit 5 Truncated Poisson models predicting total versions generated from multi-version ISRs: Impacts of beat ethnic and racial composition only.....	17
Exhibit 6. For last version of multi-version ISRs: District average number of versions and average (unweighted) beat racial composition	20
Exhibit 7 Truncated Poisson models predicting total versions generated from multi-version ISRs: Impacts of beat ethnic and racial composition while controlling for beat socioeconomic status and residential stability	22
Exhibit 8. Random sample of rejection messages (n=50) multi-version ISRs	23
Exhibit 9 Descriptive statistics: Dummy variable manifest content code variables for rejection message text field.....	25
Exhibit 10 Correlations between content codes applied to rejection message field	26
Exhibit 11 Connections between rejection message content codes and detainee race or beat racial composition: Gross impacts	28
Exhibit 12 RAS error	30
Exhibit 13. Ethnoracial patterning of RAS errors (weighted data).....	31
Exhibit 14 RAS justification error rates and beat racial composition (weighted data)	32

3 EXECUTIVE SUMMARY

This report examines three different aspects of revising investigatory stop reports (ISRs). Considered first is whether the number of revisions an ISR requires depends on the racial or ethnic composition of the beat where the stop took place, or on the race or ethnicity of the detainee. It turns out it does depend on the racial composition of the beat; the expected number of revisions required proves significantly higher in beats that are 70 percent or more Black and non-Hispanic in residential composition. This statistically significant link, however, does depend on some analytic details. It is also true, descriptively speaking, that some districts' relative contribution of multiple version ISRs is higher than their relative contribution of single version ISRs.

The second area examines the frequency with which key words, arguably relevant to the concerns of the Agreement, appear in the rejection messages supervisors send back when reports need work. Words relevant to pat downs and justifications crop up significantly more frequently in predominantly Black non-Hispanic beats. Here too, the link depends on some analytic details.

The third area considers the Consultant's legal coding of reasonable articulable suspicion (RAS) factors and processing concerns in multi version sampled ISRs. RAS concerns do surface for a sizable fraction of cases, and are linked, but in different ways, to the racial or ethnic composition of the beat where the stop took place. A variety of processing concerns surface as well, most troublingly a lack of final status codes for the majority of sampled records.

The outcomes considered here – number of versions of revised Terry stop reports, key terms communicated by supervisors back to submitting officers, legal assessments of what is happening when reports are revised – are new both to this effort supporting the Agreement, and new to the scholarship on stop and frisk. The meaning, that is, what's behind these indicators, has not yet been established. Given that novelty, more caution than usual should be applied to framing the implications of this report. At the same time, there are points here of potential concern.

4 BACKGROUND, PURPOSE, AND SOURCES

4.1 BACKGROUND

For this reporting period, the second half of calendar year 2016, the Chicago Police Department (CPD) could track different versions of investigatory stop reports (ISRs) for those reports requiring revisions. This capability was not available for the first reporting period, the first half of calendar year 2016.

4.2 PURPOSES AND SOURCES

The current report examines three questions about multiple version ISRs by considering two different data sources.

4.2.1 Questions about all multiple version Investigatory Stop Reports (ISRs)

One avenue uses information about all multiple version ISRs generated during the period. Looking at all versions of multiple version ISRs provides important descriptive information. More specifically, when multiple versions of reports are generated, how many versions get generated?; Are more versions likely to be generated for certain types of locations or for certain types of detainees? Of particular interest given the focus of the Agreement are links between the number of versions and the race or ethnicity of the detainee, or the race or ethnicity of the residents in the beat where the stop took place.

The above questions rely on a data set composed exclusively of multiple version ISRs, including all versions for each ISR, supplied by the CPD upon request in September 2017. That file also contained information about why a report version required revision.

A related key question is what features of a stop report require its revision? CPD personnel provide information, in a field called "rejection message," about which features of the report need improvement. Given that one of the issues relevant to the Agreement are legal justifications for the stops themselves and, if they occur, pat downs and searches, this report describes the results of some extremely simple searches for keywords in this rejection message field. Do certain of these keywords show up more or less frequently depending either upon the ethnicity or race of the detainee, or upon the ethnic or racial composition of the beat where the stop took place?

4.2.2 Questions about sampled Investigatory Stop Reports

Another set of questions relies on a different data source. The Consultant to the Agreement, Hon. Arlander Keys (Ret.), received from CPD, information about a random sample, that these authors generated, of multiple version stops. The random sample included roughly equal numbers of Hispanic, White non-Hispanic, and Black non-Hispanic detainees. He received information about all versions of each sampled multiple version stop, and the rejection message accompanying each version. The forms he received, however, did not include detainee race, detainee ethnicity, or the district where the stop took place. The Consultant coded not only features of the stop, but also assessed reasons requiring report revisions, and whether revised versions addressed stated concerns. The Consultant used a closed ended coding form.¹

Coded information about the sampled records entered into a database was downloaded, checked, cleaned, and merged with the full set of information from the last version of each sampled ISR, the latter extracted from the stacked multi-version ISR file provided by the CPD and described as the first data source. Once linked, information about detainees and the location of stops could be connected to the assessments of each ISR made by the Consultant.

¹ The coding instrument for the multiple version investigatory stop reports appears here as Appendix B.

5 DATA PROCESSING

5.1 FULL SET OF MULTI-VERSION ISRs, WITH DIFFERENT VERSIONS STACKED WITHIN INDIVIDUAL REPORTS AND DEMOGRAPHIC INFORMATION ADDED

In early September, pursuant to our request, the CPD supplied a file containing all versions for all multiple version ISRs generated during the last half of 2016. This file also contained information on why each specific record needed revision. This information was contained in a field called rejection message. Beat level demographic information compiled from the 2011 – 2015 American community survey census data file at the census block group level was later added, using district number as the linking variable.² This file is called the stacked version multi-ISR file.

The file provided by the CPD, which "stacked" sequential versions of each ISR one after the other, generated a total of 2,707 separate ISRs. This number disagrees with the expected number of multi-version ISRs for the timeframe which CPD reported: 2,714. The discrepancy arises because 7 ISRs were from outside the city boundaries (District 31). This number was further reduced because 37 multiple version ISRs reported stops of detainees who were not in one of the three primary ethnorracial groups – Black non-Hispanic, White non-Hispanic, and Hispanic detainees – of interest here.³

5.2 LAST VERSION FILE OF MULTI-VERSION ISR RECORDS

From the multi version ISR file another file was created containing only the last version of each multi-version ISR. This last version file contained 2,700 records. This is called the last version ISR file. An additional 7 records were missing on one or more key variables, leaving a final n of 2,663 final-versions of multi-version ISRs for processing.

5.3 SAMPLING EQUAL NUMBERS OF MULTI-VERSION ISRs FOR EACH ETHNORACIAL GROUP

Using a list of ISR numbers for multiple version ISRs that were checked against the CPD master list of ISR numbers for the second half of 2016, we randomly sampled equal numbers of ISRs for Black non-Hispanic detainees, White non-Hispanic detainees, and Hispanic detainees. The sampling ratios used for each ethnorracial group were the same as the sampling ratio used for the single version ISR samples drawn for legal analysis (see Period 2 legal narratives report). One hundred seventy-six multiple version ISRs were sampled and coded.⁴

² Details on the processing of the American Community Survey 2011-2015 census block group data and constructed beat variables can be found in the Period 2 Post Stop Outcomes report Section 8.5.2.

³ This group included 27 Asians, 7 Indians and 3 Pacific Islanders.

⁴ The sampling ratio was kept the same because the original purpose was to join the sampled and coded multi-version ISRs with the sampled and coded single-version ISRs in the same analysis. As interest deepened in exploring further the changes taking place across versions of multiple version ISRs, however, these authors suggested instead linking the sampled and coded multi-version ISRs with the full set of multi-version ISRs as part of the analysis of changes across versions. Doing this means that the data from the sampled and coded multi-version ISRs appear only in this one report. Stated differently, the sampled and coded ISR data here do not overlap with the sampled and coded ISR data used in the Period 2 legal narratives report.

5.4 CODED LEGAL NARRATIVES OF MULTI-VERSION SAMPLED RECORDS

The CPD directly supplied to the Consultant files for all the 176 sampled multi-version records. Each file contained information on each version of each ISR, along with information about the reasons why earlier versions of reports need revision. Key identifying information, such as race, ethnicity, and district, however, were not included.

Information about the coding of each sampled record was entered into a Survey Monkey database. Because the online program "froze" sometimes during data entry, it was necessary to remove duplicate, partially completed records from the final file. For each ISR where a duplicate or partially complete record was present, the latest coded version of the record was retained. Further processing was required to get the database and the coded file to match up with ISR numbers in the main last version ISR file. After all the deduplicating, cleaning, checking, and linking, completed coding forms for 173 of these 176 records were successfully relinked to the last version ISR file.

These 173 records included stops of

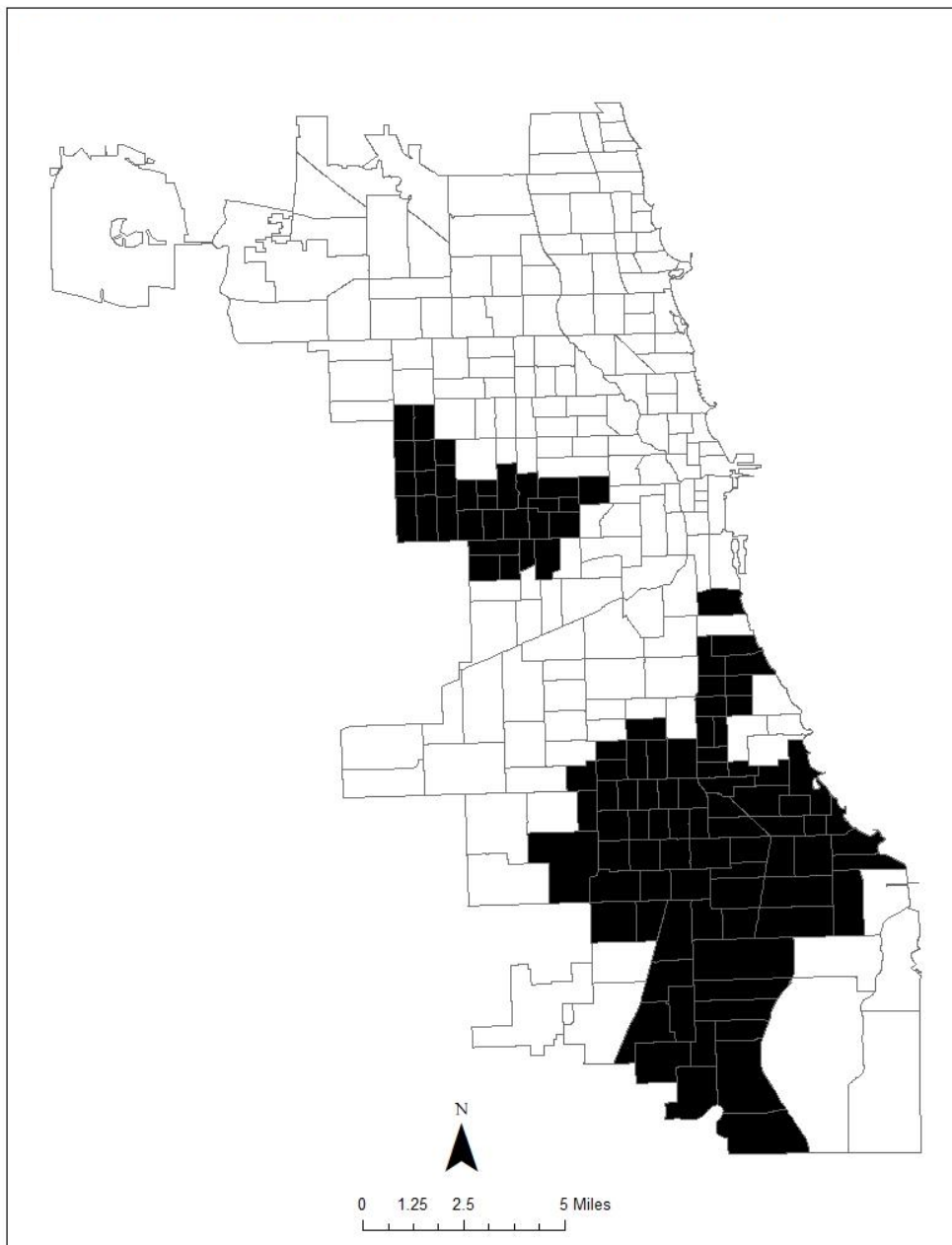
- 48 White non-Hispanic detainees
- 61 Black non-Hispanic detainees
- 64 Hispanic detainees

Of these coded 173 stops

- 113 occurred in beats where the residential composition was less than 70 percent Black non-Hispanic; and
- 60 occurred in beats where the residential composition was 70 percent or more Black non-Hispanic.

The location of these beats that are 70 percent or more Black and non-Hispanic in residential composition is shown in Exhibit 1.

Exhibit 1 Chicago beats-within-districts 70 percent or more Black and non-Hispanic



Note. Beats shaded black are 70 percent or more Black and non-Hispanic, based on re-allocated 2011-2015 U.S. Census American Community Survey census block group data re-allocated to police beats.

Further, of these coded 173 stops

- 129 occurred in beats where the residential composition was less than 70 percent Hispanic; and
- 44 occurred in beats where the residential composition was 70 percent or more Hispanic.

5.5 SAMPLE WEIGHTS

The sampling strategy was geared to generate equal numbers of sampled multi-version records for each of the three key ethnoracial groups: Black non-Hispanic, White non-Hispanic, and Hispanic detainees. Doing so creates a discrepancy between the relative frequency of the three groups in the sampled records and the relative frequency of the three groups in the full set of ISRs the second half of 2016. This discrepancy disappears if we apply weights to the coded legal narrative data. The weights count each sampled Black non-Hispanic ISR about two times, each sampled Hispanic ISR about .6 times, and each sampled White non-Hispanic record about .27 times. If the sampled and coded multi-version ISRs are weighted, then the number of coded records for each ethnoracial group is as follows:

- 13 White non-Hispanic sampled detainees;
- 127 Black non-Hispanic sampled detainees;
- 39 Hispanic sampled detainees;
- for a weighted total of 179 sampled records.

Some analyses will report based on unweighted data, which means the results apply *only* to the sample itself and not the population of records from which the sample was drawn. Some analyses will report analyses based on weighted data, meaning the results are intended to apply to the population of records from which the sample was drawn, that is, all Period 2 multi-version ISRs. All the statistical significance tests applied here permit inference from the sample back to the full set of records, and tell us about differences in that full set. Some analyses will report using both unweighted and weighted data.

5.6 HOW MANY MULTI-VERSION ISRs ARE THERE IN PERIOD 2 (JULY-DECEMBER 2016)?

The Period 2 report on post stop outcomes did include in its analyses records representing the last version of each multiple version investigatory stop report. For the total 51,538 ISRs stops available in the second half of the calendar year and analyzed in that report, 2,707 of those reports (5.2 percent of the total) were multiple version ISRs. This leaves out 7 stops that did not take place in a police district within the City of Chicago.

5.7 CODING MANIFEST CONTENT FROM THE REJECTION MESSAGE FIELD

5.7.1 Focus

This report examines the manifest content of the rejection message field using simple searches for key terms that reflect some of central concerns about investigatory stops.

- How often did reasonable articulable suspicion or related terms get mentioned?
- How often did the revision guide mention either the pat down, if it occurred, or the search?
- How often was the concern about justifying some feature in the narrative generated by the officer?

Given time constraints, the choice of these key words was made quickly and without consultation with the parties to the Agreement. Authors anticipate that if they are directed to develop latent

coding protocols to examine themes in these rejection messages, they will receive important input from the parties and will look to communicate with parties about the themes investigated.

5.7.2 Limitations

This manifest content analysis is extremely crude, and its results should be interpreted with considerable caution. One must always "proceed cautiously when making inferences solely from crude summary measures. They harbor potential problems."⁵ Most importantly, the meaning of a specific word can depend on the context. What is done here is **not** a "keyword in context" approach. More simply, although we can count how many times a certain word appears, we don't know what the **meaning** is of that specific word in every instance.

Scientifically assessing the meaning of specific words or phrases requires assessing latent content using qualitative content analysis and multiple trained coders. **That analysis has not yet been undertaken and is not feasible to complete in time for the Period 2 reports. We plan, if directed, to use a more adequate latent content coding strategy when examining data for calendar year 2017.**

6 DESCRIPTIVE INFORMATION: THE FULL SET OF MULTI-VERSION ISRS

6.1 DISTRIBUTION ACROSS ETHNORACIAL GROUPS

Exhibit 2 shows the distribution of multiple version ISRs across districts, and across ethnoracial groups within districts.⁶ The numbers shown here are based on the last version file of multi-version ISRs. The very bottom row shows the proportion of *single* version ISRs associated with each of the primary ethnoracial groups. The single version percentages roughly correspond, within a couple of percentage points, to the corresponding percentages for each ethnoracial group's contribution to multiple version ISRs.

In short, multiple version ISRs appear for each ethnoracial group in rough proportion to the fraction of single version ISRs generated for each group. The current report does not test whether these discrepancies are statistically significant or not.

6.2 DISTRICT DIFFERENCES IN RELATIVE CONTRIBUTIONS OF MULTIPLE VERSUS SINGLE VERSION ISRS

Exhibit 3 considers if districts differ in how often they contribute multiple version ISRs, in comparison to how often they contribute single version ISRs. In other words, how do multiple version ISRs, as a "share" of all a district's submitted ISRs, vary across districts?

The investigation begins by calculating for each district its relative contribution to all the single version ISRs generated during the period, and separately, its relative contribution to all of the multi-version ISRs generated during the latter half of 2016. In short, each district contributes its own "share" to both the full set of single version ISRs, and the full set of multiple version ISRs. Each multiple version ISR is counted only once, regardless of how many versions it spawns.

⁵ (Taylor, 1994: 320)

⁶ The number is lower than 2,707 because 37 ISRs were not in one of the three primary ethnoracial groups, and 7 ISRs were from outside the city boundaries (District 31)

Those two relative contributions or shares are then contrasted for each district. That is, we take the proportion of each district's overall multiple version ISR contribution and divide it by the proportion of each district's overall single version ISR contribution. Stated differently, we are looking at each district's (multi-version share / single version share). This creates a ratio.

If the ratio is above one, that suggests a district's *relative* contribution to the period total is greater for multi-version ISRs than it is for single version ISRs. Its multi-version share is bigger than its single version share. If the ratio is below one, that suggests a district's relative contribution to the period total is smaller for multi-version ISRs than it is for single version ISRs. Its multi-version share is lower than its single version share.

The districts are then sorted from lower to higher scores on this ratio. To oversimplify, districts toward the bottom of the list are less likely to get the ISR paperwork properly completed the first time around. There could be *any number of reasons for* this ranging from more complex stops to more scrupulous supervisors, to mention just a couple. Construct validity, that is, the meaning, of the indicator examined is not yet established.

Three districts on the South side of Chicago – districts 5, 22, and 17 – appear at the very bottom of the ordering. It is noted in passing that all three of these districts are somewhat close to one another. Conversely, one could argue that districts where the ratio is low – districts 16, 15, and 20 for example – are districts where personnel are more successful at getting the paperwork properly submitted the first time around. These districts are all located on the north side of Chicago.

To provide a simpler description of what is being looked at here, consider the case of district 5. Officers here generated 2,233 single version ISRs during the timeframe, and these records represented 4.65 percent of all single version ISRs during the timeframe. The same district also generated 209 multi-version ISRs, and these represented 7.8 percent of all multi-version ISRs generated during this timeframe. The ratio of these two percentages – 7.8 percent divided by 4.65 percent – generates the ratio of 1.69. Given that ratio one could say that the district's share of city-wide multi-version ISRs was 69 percent higher than the district's share of city-wide single version ISRs.

Exhibit 2 Distribution of multiple-Version ISRs across districts and ethnoracial groups within districts

District	White NH		Black NH		Hispanic		Total		Percent of total
	N	Percent	N	Percent	N	Percent	N	Percent	
1	8	23.53	25	73.53	1	2.94	34	100	1.28
2	2	1.16	169	97.69	2	1.16	173	100	6.50
3	2	2.33	84	97.67	0	0	86	100	3.23
4	1	0.47	155	73.46	55	26.07	211	100	7.92
5	2	0.96	199	95.22	8	3.83	209	100	7.85
6	1	0.74	134	99.26	0	0	135	100	5.07
7	3	1.23	237	97.13	4	1.64	244	100	9.16
8	15	12.61	55	46.22	49	41.18	119	100	4.47
9	18	7.96	75	33.19	133	58.85	226	100	8.49
10	8	4.91	95	58.28	60	36.81	163	100	6.12
11	23	7.17	284	88.47	14	4.36	321	100	12.05
12	8	9.3	53	61.63	25	29.07	86	100	3.23
14	3	6	18	36	29	58	50	100	1.88
15	3	3.41	82	93.18	3	3.41	88	100	3.30
16	9	34.62	6	23.08	11	42.31	26	100	0.98
17	12	18.75	13	20.31	39	60.94	64	100	2.40
18	3	6.12	43	87.76	3	6.12	49	100	1.84
19	13	22.03	33	55.93	13	22.03	59	100	2.22
20	8	29.63	13	48.15	6	22.22	27	100	1.01
22	10	13.16	66	86.84	0	0	76	100	2.85
24	21	20.59	50	49.02	31	30.39	102	100	3.83
25	12	10.43	35	30.43	68	59.13	115	100	4.32
Total	185	6.95	1,924	72.25	554	20.8	2,663	100	100
Single version ISR (% of total)		8.57		71.84		19.59			

Note. NH = non-Hispanic. Source: Period 2 multi-version ISR file provided by CPD 9/26/2017. Each multi-version ISR counted only **once** regardless of how many versions ultimately submitted. Each row shows, for each specific ethnoracial group in each district, the number of multiple version ISRs, and the percentage of multiple version ISRs in that district contributed by that civic ethnoracial group. The total line shows, across all districts, that 6.95 percent of multi-version reports involved white non-Hispanic detainees; 72.25 percent involve black non-Hispanic detainees; and 20.8 percent involved Hispanic detainees. Directly under the "total" row, numbers show the percentage of single version ISRs contributed by each ethnoracial group. The percentage of single version ISRs contributed by each ethnoracial group can be compared to the percentage of multiple version ISRs contributed by each ethnoracial group.

Exhibit 3 Contrasting district multiple version vs. single version shares

District	Ratio: district multi ISR contribution / district single version ISR contribution	District contribution to all multi-version ISRs (as percent of multi-version ISRs)	Number of multi-version ISRs	District contribution to all single version ISRs (as percent of single version ISRs)	Number of single version ISRs
16	0.31	0.98	26	3.17	1,523
15	0.46	3.30	88	7.24	3,474
20	0.57	1.01	27	1.79	858
25	0.79	4.32	115	5.46	2,620
8	0.79	4.47	119	5.63	2,700
10	0.82	6.12	163	7.50	3,600
19	0.82	2.22	59	2.69	1,292
12	0.83	3.23	86	3.90	1,870
11	0.91	12.05	321	13.23	6,346
14	0.96	1.88	50	1.96	942
18	0.96	1.84	49	1.91	918
3	1.04	3.23	86	3.10	1,487
1	1.05	1.28	34	1.22	583
2	1.09	6.50	173	5.97	2,864
24	1.11	3.83	102	3.44	1,652
9	1.16	8.49	226	7.32	3,514
6	1.31	5.07	135	3.88	1,861
4	1.36	7.92	211	5.84	2,802
17	1.38	2.40	64	1.74	834
7	1.42	9.16	244	6.44	3,092
22	1.50	2.85	76	1.90	912
5	1.69	7.85	209	4.65	2,233
Total		100.00	2,663	100.00	47,977

Note. July-December 2016 data on investigatory stop reports. Districts sorted by ratio of (multi ISR share / single ISR share). **See text for cautions on interpreting ratio differences.**

7 NUMBER OF VERSIONS

7.1 CAUTIONARY NOTE

As far as the authors are aware, stop and frisk researchers have not performed the type of analysis done here, seeking to describe and observe the correlates of how many times an investigatory stop report is revised. Consequently, the meaning or, more technically, the construct validity of this indicator, at either the level of the individual report or at the level of the police-beat-within-district is unknown.

It is not unusual to have construct validity questions about features of criminal justice processing (Goldkamp, 2010). It is also not unusual to have questions about the construct validity of ecological indicators in community criminology research (Taylor, 2011). Because we are

examining beat-level features of revisions, these become ecological indicators. Extensive future research will be required to learn more about what this indicator means. It could well mean different things in different circumstances.

These unanswered research questions about the meaning of the number of report revisions exist alongside operational questions about and variation in report review procedures within the CPD. From an operational standpoint, **because CPD personnel may decide for any number of reasons that a particular ISR does not merit further revision, the meaning of "number of versions" is ambiguous pending further investigation.**

It would be wrong to simply think of several versions as the number of tries required before the submitting officer "got it right." The final state of an ISR can mean several different things. It may mean that an officer has satisfactorily resolved ambiguities previously noted by supervisors. Alternatively, it may mean that supervisors have decided that the report is deficient and simply cannot be fixed. There are other possible end states as well. In other words, once a supervisor decides that an ISR has been revised to the point where it is "good enough", "good enough" could mean several different things.

That said, it may be possible in a versioning report for Period 3, to reduce these operational ambiguities somewhat by paying attention to final status codes and separating out records by those codes.

But another source of ambiguity arises from the variations across investigative stops in the complexity of the situations confronted by investigating officers. We simply don't know if reports with more challenging stop features – more persons involved, more ambiguous situations, for example – inherently take longer to get to one of the acceptable final states.

These ratios are problematic in at least one regard. The underlying share numbers are based on extremely different numbers of records in different districts. Therefore, the underlying share numbers are more precise in some districts and less precise others. These ratios also may be flawed in other ways. Further, there could be many reasons contributing to the share differentials.

These ratios of (multi-version share of city total multi-version ISRs / single version share of city total single version ISRs) are presented **for descriptive purposes only**. It is not clear how these ratios should be interpreted. Different districts confront enormously different investigative situations, resource levels relative to need, and types of crime and disorder confronted daily. Personnel reviewing ISRs undoubtedly vary in important ways across districts as well. Intertwined with these differences is district to district variation in normative approaches to different aspects of police work (Klinger, 1997). Simply put, it is not yet clear what these ratios mean. We are introducing a new indicator to the stop and frisk research literature. We need to learn more about it. The variation multiple/single version ratios, and the geographic patterning of that variation in these ratios, is simply noted.

In sum, the outcome examined here, to the best of the authors' knowledge, is a new indicator as far as stop and frisk research is concerned; that novelty means there are important unresolved questions about the meaning – the construct validity – of the indicator. In addition, CPD

personnel reviewing ISRs can decide to ask for a further revision, or decide a further revision isn't needed, for several reasons. These complications make it even harder to decide what this indicator means.

7.2 HOW MANY MULTI-VERSION REPORTS GOT REVISED HOW MANY TIMES AND WHERE?

Exhibit 4 shows how often each multiple version ISR got revised. For most of these multi-version reports, 87.8 percent, one revision was “enough.” For another 7 percent of these records, two revisions were required for a total of three versions. So, for 95 percent of these records, two revisions proved sufficient before the record could be assigned one of several ultimate status codes.

Twelve reports got revised a total of five times. Five of these 12 came from District 7. The beats in District 7 range from 88 percent Black non-Hispanic to almost 98 percent Black non-Hispanic. The (unweighted) average beat has 94.4 percent of its residents identified as Black non-Hispanic. The typical (median) beat is 94.8 percent Black non-Hispanic in residential composition.

The maximum number of versions for any ISR was six versions; this happened with two reports. One of these reports was from district 6, one was from district 2. In District 6, in a typical beat, residents are 96.6 percent Black non-Hispanic. In District 2, residential composition across beats ranges from 22 percent to 97 percent Black non-Hispanic; the (unweighted) average beat is 83 percent Black non-Hispanic in residential makeup, and the typical (median) beat is 90.4 percent Black non-Hispanic in residential makeup.

It appears therefore, that the districts generating the ISRs which got revised most frequently within this timeframe came from districts which were very heavily Black non-Hispanic in their residential makeup. At this juncture, this feature is just noted descriptively.

Exhibit 4 Numbers of versions of multi-version ISRs

Maximum number of versions generated for multi-version ISRs	N	Percent
	-----	-----
2	2,371	87.81
3	196	7.26
4	119	4.41
5	12	0.44
6	2	0.07
Total	2,700	100

Note. ISR multiple version ISR data for Period 2, July-December 2017.
Source: CPD.

8 PREDICTING HOW MANY REPORT VERSIONS AMONG MULTI-VERSION ISRS

8.1 THE QUESTION

Among the multi-version ISRs, did features of the detainee or the stop influence how many times the report had to be revised? Of particular interest are impacts of detainee race, ethnicity, and gender. Racial and ethnic composition of the beats-within-district where the stop took place are also of interest. This analysis uses the final version multi-ISR file with the last version of each multi-version ISR combined with information about the number of versions, and information about beat composition.

8.2 APPROACHES

This question is examined different ways. First, several types of models tried to predict how many versions had been created for each stop report.⁷ Some models examined individual detainee attributes while others examined the demographic factors associated with the beat. For racial and ethnic composition, these models used categorical variables based on the scholarship of Peterson and Krivo (Peterson & Krivo, 2010).⁸ Following their guideline, the beat was classified as predominantly Black and non-Hispanic if at least 70 percent of the beat population were in this group. A beat was classified as predominantly Hispanic if its population was at least 70 percent Hispanic.

Models were run separately for two different 50 percent random samples of multi-version records. A statistically significant result could be considered more robust if it replicated internally across both random samples of the data.

Exhibit 5 Truncated Poisson models predicting total versions generated from multi-version ISRs: Impacts of beat ethnic and racial composition only

Variable	Name	Robust				95 % Confidence intervals				
		IRR	b	se	z	p =	LCL	UCL	LCL	UCL
Sample 1										
Beat >= 70 % Black NH (1, 0 else)	p_k_bla	1.595	0.467	0.211	2.22	0.027	0.05391	0.8801	1.055	2.411
Beat >= 70% Hispanic (1, 0 else)	p_k_his	1.804	0.590	0.282	2.09	0.036	0.03778	1.14268	1.039	3.135
Constant		0.338	-1.085	0.183	-5.94					

Note. Period 2 (July-December 2017) data. N = 1,307. Errors adjusted for 241 beat clusters. Wald Chi squared (df=4) = 6.05, p < .05.

⁷ The types of models examined included truncated Poisson single level models with standard errors clustered by beat within district, truncated Poisson single level models with standard errors clustered by district, Poisson models with each of the two above types of clustering, and mixed effects multilevel Poisson models with the clustering of stops either by district or by beat.

⁸ For scholarly background on the reasoning behind these categorical variables capturing community racial and ethnic composition, see the Period 2 Legal Narratives Report.

Sample 2

Beat >= 70 % Black NH (1, 0 else)	p_k_bla	2.261	0.816	0.216	3.78	0.001	0.39301	1.23876	1.481	3.451
Beat >= 70% Hispanic (1, 0 else)	p_k_his	1.793	0.584	0.285	2.05	0.041	0.02512	1.14242	1.025	3.134
Constant		0.270	-1.309	0.188	-6.96					

Note. Period 2 data (July-December 2017). N = 1,393. Errors adjusted for 242 beat clusters. Wald Chi squared (df=2) = 14.32, p < .001.

Note. Results from truncated Poisson regression models, with robust standard errors clustered by beat within district. Outcome is the total number of revisions required for each multi-version ISR and ranges from 2-6. *Only* the last version of each ISR, and *only* multi-version ISRs are examined. Results reported separately for two random samples. Because variables for percent Asian, and percent other are not entered, the contrast or reference string is just beats that are both less than 70 percent Black non-Hispanic and less than 70 percent Hispanic.

8.3 DETAINEE RACE, ETHNICITY, AND GENDER

No robust connections appeared within the multi-version ISRs between the total number of revisions and detainee race, ethnicity, and gender (results not shown). Models with detainee predictors showed only one significant connection. The total number of versions generated for a stop report was significantly lower (p < .05) for stops involving male detainees as compared to female detainees. This connection appeared, however, only in one of the two random samples of the multiple version ISRs. Given this result surfaced in only one of the two random samples of records, it is not viewed as robust.

8.4 BEAT-LEVEL ETHNIC AND RACIAL COMPOSITION

8.4.1 Links

In some models with beat level demographic factors entered as predictors, by contrast, some robust connections **did** appear. More specifically, in truncated Poisson models with error terms clustered at the beat level, when only beat racial composition and ethnic racial composition categorical variables were entered as predictors, number of versions was significantly (p < .05) higher in predominantly Black non-Hispanic beats, and in predominantly Hispanic beats, as compared to beats that were *both* less than 70 percent Black non-Hispanic and less than 70 percent Hispanic. See Exhibit 5.

To more specifically interpret a particular finding, consider the impact of the categorical variable contrasting predominantly Black and non-Hispanic beats with beats that were both less than 70 percent Black non-Hispanic and less than 70 percent Hispanic, i.e., predominantly White and non-Hispanic. In Exhibit 5, in sample 1, the incidence rate ratio (IRR) of 1.60 suggests that stops taking place in beats where the residential population is 70 percent or higher Black and non-Hispanic, were predicted on average to have an expected number of versions that was 60 percent

higher,⁹ compared to the expected number of versions for multi-version ISRs taking place in beats that were *both* less than 70 percent Hispanic and less than 70 percent Black non-Hispanic.¹⁰

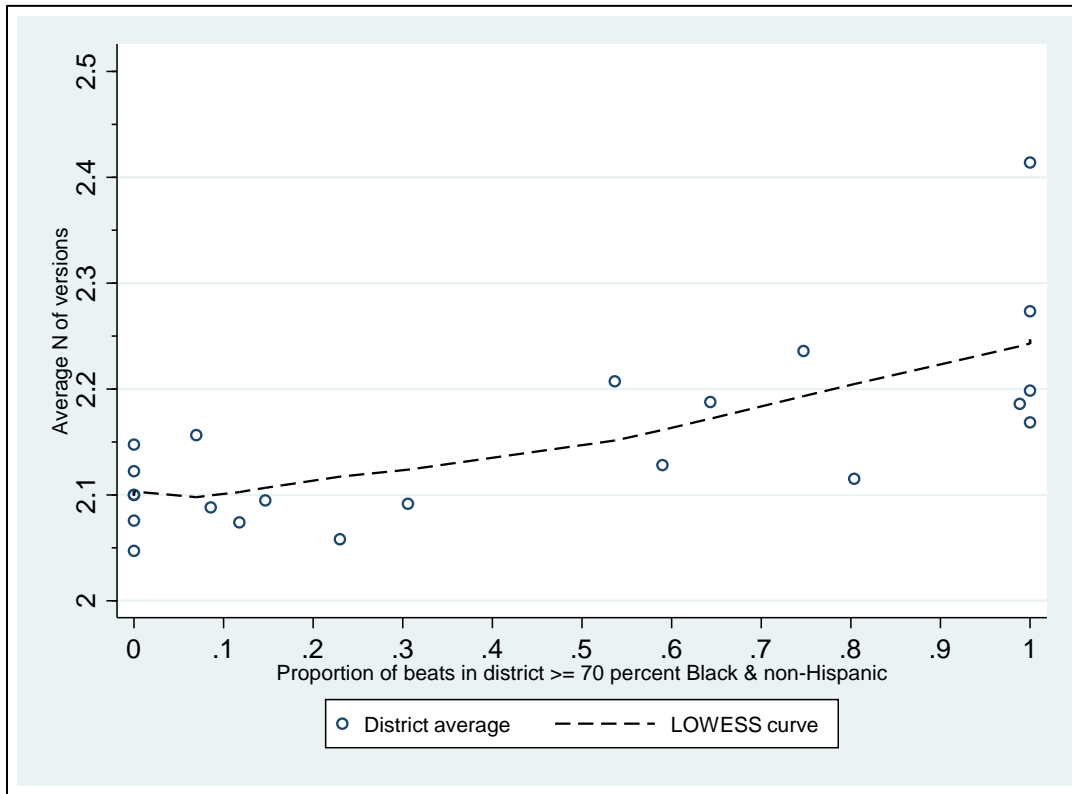
Exhibit 6 shows the relationship between beat-average (unweighted) racial categorical composition of a district, and the average number of versions that the multiple version ISRs went through in that district. Districts where a higher (unweighted) fraction of beats within the district were predominantly (70 percent or higher) Black and non-Hispanic on average had more iterations for their multi-version ISRs. The averages were

- highest in District 5 (average number of versions for multi-version ISRs = 2.41; proportion 70 percent or higher Black non-Hispanic beats = 1),
- next highest in District 7 (average number of versions for multi-version ISRs = 2.27; proportion = 1), and
- third highest in District 2 (average number of versions for multi-version ISRs = 2.24; proportion = .75).

⁹ This is derived by taking the incidence rate ratio (IRR) and subtracting one from it because the number is above one.

¹⁰ The contrasted beats also include, in addition to predominantly White non-Hispanic beats, some beats with Asian population, and up to 5 percent of residents classified ethnographically as other.

Exhibit 6. For last version of multi-version ISRs: District average number of versions and average (unweighted) beat racial composition



Note. District average number of versions, only for multiple version ISRs, appears on vertical axis. In effect, this is the district-wide average of how many versions it took until a multiple-version ISR was no longer revised. Horizontal axis shows the proportion of beats within each district classified as 70 percent or more Black non-Hispanic. Beats within districts *not* weighted. LOWESS curve shows a locally weighted regression line.¹¹ Kendall's Tau-b = .48; $p < .01$. Data from July-December 2016 CPD ISRs. Single version ISRs *excluded*.

8.4.2 Robustness questions

Both the categorical race and ethnicity beat indicators connect significantly to the outcome in both random samples of stops. This suggests the links are somewhat robust in that they replicate across different random subsets of ISR records.

But the links did *not* prove robust across different analytic approaches. If stops were clustered by district rather than beat, the significant links failed to emerge (results not shown). Given the dependency of the significant links on the specific type of model used, they should not be considered robust across statistical models that organize the data in different ways.

Returning to the ISR-level analyses, however, shows that this connection is **not robust in other ways**. If additional demographic features of the beat are incorporated in the models, then the

¹¹ (Cleveland, 1979)

connections between ethnoracial composition and number of versions becomes less robust. The racial composition variable (Black NH \geq 70 percent) links to number of versions in **only one** random sample, and the Hispanic composition variable (Hispanic \geq 70 percent) links to number of versions in **only one** random sample.¹² See Exhibit 7.

Further, these connections seen in the truncated Poisson models did not consistently emerge if multilevel mixed effects models were used. So, the results are **not robust across alternate analytics** either.

Moreover, the results described here are specific to the **multiple version ISRs**. If attempts were made to look at the relationship between number of versions while including the single version ISRs as well, the connections described above did not appear.

Finally, if ISRs were dichotomized into two groups – single version and multiple version – again, neither detainee race and ethnicity, nor beat ethnic or racial composition significantly predicted which type of record was generated.

¹² These significance tests are two-tailed, as are all statistical significance tests in this and the other reports. With a one-tailed significance test the race and ethnic composition links to versions shown in Exhibit 6 would be significant at $p < .05$, and one could then argue that the links prove robust across models with different sets of predictors. We do not opt to do that because, as explained above, the outcome examined here is novel.

Exhibit 7 Truncated Poisson models predicting total versions generated from multi-version ISRs: Impacts of beat ethnic and racial composition while controlling for beat socioeconomic status and residential stability

Sample 1

Variable	Name	95 % Confidence intervals									
		Robust					b				
		IRR	b	se	z	p =	LCL	UCL	LCL	UCL	
Beat >= 70 % Black NH (1, 0 else)	p_k_bla	1.818	0.598	0.292	2.05	0.04	0.026	1.170	1.027	3.221	
Beat >= 70% Hispanic (1, 0 else)	p_k_his	1.670	0.513	0.292	1.76	0.079	-0.059	1.085	0.942	2.961	
Residential socioeconomic status ^(d)	sesindx2	0.922	-0.081	0.208	-0.39	ns	-0.489	0.327	0.613	1.387	
Residential stability ^(c)	stabindx	1.246	0.220	0.133	1.65	ns	-0.041	0.482	0.960	1.619	
Constant		0.350	-1.051								

Note. Period 2 (July-December, 2017) data. N = 1,371. Errors adjusted for 238 beat clusters. Wald Chi squared (df=4) = 18.8, p < .01.

Sample 2

Beat >= 70 % Black NH (1, 0 else)	p_k_bla	1.866	0.624	0.332	1.88	0.06	-0.027	1.274	0.973	3.576
Beat >= 70% Hispanic (1, 0 else)	p_k_his	1.907	0.646	0.285	2.26	0.024	0.087	1.204	1.091	3.335
Residential socioeconomic status ^(a)	sesindx2	0.933	-0.069	0.250	-0.28	ns	-0.559	0.420	0.572	1.522
Residential stability ^(b)	stabindx	1.116	0.110	0.186	0.59	ns	-0.255	0.475	0.775	1.608
Constant		0.260	-1.347							

Note. Period 2 data (July-December, 2017). N = 1,328. Errors adjusted for 240 beat clusters. Wald Chi squared (df=4) = 8.67 (ns)

Note. Results from truncated Poisson regression models, with robust standard errors clustered by beat within district. Outcome is the maximum number of revisions and ranges from 2-6. *Only* the last version of each ISR, and *only* multi-version ISRs are examined. Results reported separately for two random samples.

^(a) Socioeconomic status was captured with an internally consistent (Cronbach's $\alpha = .91$) multi-item index using median house value, median income, the percent of households reporting less than \$20,000 in earnings (reversed), and the percent of households reporting greater than \$60,000 in earnings

^(b) Residential stability was captured with an internally consistent (Cronbach's $\alpha = .73$) multi-item index using percent owner occupied households; percent of all households, currently owner occupied, where the household moved in before 1990; and percent of all households, currently renter occupied, where the household moved in before 2000. Multi-collinearity was not problematic.

8.4.3 Summary of number of versions and race and ethnicity of beat

In sum, using one particular type of analysis, with one particular set of beat-level demographic predictors, connections appear between racial and ethnic classification of the beat in which the stop occurred, and the number of versions ultimately generated for multiple version investigative stop reports originating in that beat. This connection is visible (Exhibit 6) at the district level. That said, the connection between racial composition and generating more versions of multiple version ISRs, **although robust across repeated random**

samples, is *not robust* across different analytic approaches or models with different sets of predictors. Given this lack of robustness, **these connections should be viewed *cautiously*** given that they are dependent upon one particular type of analytic approach, and are markedly reduced if other beat demographic get included in models.

In short, there is a suggestion that, for reasons unknown, multiple version ISRs required more reworking if the stop took place in a beat that was predominantly Black and non-Hispanic, or predominantly Hispanic. Much more work is needed on the outcome analyzed here, the number of versions generated for multiple-version ISRs, because currently the meaning of this indicator is not clear.

9 EXAMINING REASONS FOR REJECTION

9.1 PRELIMINARY CONTENT ANALYSIS

Each **penultimate** version of each multi-version ISR has an accompanying rejection message, indicating which features of that version deserve additional attention. These messages are sometimes lengthy, sometimes short. Exhibit 8 provides a random sampling of rejection messages associated with 50 multiple version ISRs to illustrate the types of concerns surfacing.

Exhibit 8. Random sample of rejection messages (n=50) multi-version ISRs

It appears as if this stop is based on on-view probable cause (drinking on public way). if this is indeed the case then an ISR does not need to be completed for this stop. However, if the basis for this stop was to conduct an investigation to determine if this subject was drinking on the public way then an ISR is required. Please clarify and resubmit if necessary.

enforcement action --yes, curfew

Forgot to check a box

subjects name spelled incorrectly

no receipt

add age info

receipt no change wording

search beyond protective patdown-incident to arrest

please add the part about how Dep. Chief Georgas stated they had been observed flying low in a manner that posed a risk to public safety.

forgot to put "asked to leave"

wording

Returned to officer at his request; accidentally checked incorrect search related boxes

missing check box

Pat down factors?

INFO

Enforcement action should have curfew listed in this section. The narrative has this information.

she is listed as a MALE on the report. switch her to FEMALE and resubmit.

why no receipt?

Returned at submitting officer's request in order to correctly check the box indicating that a custodial search was done post arrest.

Put in the charge under Cited Charges for DOPW.

receipt box not checked

Administrative observation

WHAT TYPE OF SEARCH WAS INITIALLY CONDUCTED? I DONT BELIEVE YOU CAN PERFORM A CUSTODIAL SEARCH AND THEN "GIVE A VERBAL WARNING"

per convo

Returned at officer's request for corrections

Please, include more clarification in the narrative as to what "Other Reasonable Suspicion of Weapons" includes.

Please remove verbal warning from enforcement action taken

correction

R/O informed Sgt that she submitted this report and forgot to add her partner and inadvertently omitted a sentence in her narrative for the reason of her search.

change the sex of the subj

give more info

add info

Spelling, no receipt needed if no pat down

CHECK SPELLING OF CAUSE

date /correction

the stop was based on probable cause (adult in the playground) therefore no isr is needed.

change race code

add gang info

See me

add partner

routine traffic stop

Officer inadvertently failed to check the "receipt given" box

last name typed twice as first name

you have to articulate why she was patted down. see me if you have questions

might have checked the wrong boxes. a receipt is not given if you do not do a pat down. you have no pat down but receipt given.

NARRATIVE IS ABOUT THE DRIVER - NOT THE PASSENGER. DID THE INVESTIGATION INVOLVE THE PASSENGER ?

walking away from the group

need charge

vehicle/detailed info given in ticket

Note. Random sample drawn only from first version of multi-version ISRs, July-December 2016.

9.1.1 Limitations

The preliminary content coding reported here considers only the frequency with which a small number of select words or phrases appear in the rejection message field. This approach is necessarily limited. A specific word or phrase can mean different things depending on the context. For example, if "probable cause" appears in this rejection message text field, it could mean, among many other things, that the ISR was not needed because the stop was a probable cause on view stop; or it could mean that the officer reviewing the form wanted more details on the probable cause for the search which was conducted. It is not possible to determine the

meaning of the word without taking the context into account. This can only be done if raters are trained to systematically code the different phrases that appear in the text.

The implication is that the preliminary results reported here should be considered **extremely tentative**, in that it is not yet clear how to interpret the patterns that are observed. **The meaning of the terms extracted is not clear.**

In a wider context, controversy about how to analyze content continues. In content analysis, researchers have historically disagreed about focusing on manifest vs. latent content, the two approaches aligning to some extent, respectively, with a quantitative vs. a qualitative approach (Holsti, 1969). The “debate of how content analysis should be approached” continues (Grüber, 2015: 171). “According to the notion of manifest content analysis, the researcher should only focus on items that are physically present and can actually be counted. In contrast, latent content analysis extends the concept to an analysis of the meaning that is conveyed by the physical data ... Hence, the researchers interpret messages rather than only counting information items” (Grüber, 2015: 171). The view of the authors is that the strongest approach probably combines manifest and latent content coding (Taylor, 1994: 321).

9.1.2 Selecting key text

Text strings were searched for specific words and phrases. The words/phrases searched for included:

- patdown
- search
- justification
- reasonable articulable suspicion
- as well as variations on these (e.g., “justified”, “pat down”, “RAS”) ¹³

9.1.3 Descriptive statistics for manifest content coding variables

Dummy variables, coded zero when the text in question was not present, and one when the text in question was present in any form (0,1), were created for each of these four topics.

Exhibit 9 Descriptive statistics: Dummy variable manifest content code variables for rejection message text field

Variable	Variable name	N	Min	Max	Mean	SD
Rejection message contains text related to pat downs	rejpatany	3,177	0	1	0.1294	0.1127
Rejection message contains text related to search	rejseaany	3,177	0	1	0.0444	0.0424
Rejection message contains text related to justification	rejdnyany	3,177	0	1	0.0179	0.0176
Rejection message contains text related to reasonable articulable suspicion	rejasany	3,177	0	1	0.0167	0.0164

Note. See Appendix B for programming of codes. Final version of each ISR excluded.

¹³ See Appendix A for the programming.

Descriptive statistics for these dummy (0,1) variables appear in Exhibit 9. The descriptive statistics are based on all versions of each multi-version ISR *save* the last version. This is because the last version had no rejection message attached, since it was accepted.

Referring to each separate version of an ISR as a contact card, the table shows that, of the topics coded, the most frequently occurring one concerned pat downs. Words related to pat downs appeared in 12.9 percent of the contact cards. By contrast, words related to searches appeared in only 4.4 percent of the contact cards. This is understandable because searches occur far less frequently than do pat downs.¹⁴

Words referring specifically either to RAS or to justification each appeared in one to two percent of the records. Given the very low rates at which these two sets of words got mentioned, multivariate modeling difficulties might surface.

How often, at the record level, do these different topics co-occur in a rejection message? Exhibit 10 shows the relevant correlations. The content code addressing pat down mentions in rejection messages correlates quite strongly (tetrachoric correlations > .50) with each of the other three codes. Further, calculation of the 99 percent confidence interval for each correlation (details not shown) indicated that the correlation between pat down concerns and justification concerns was stronger than the correlation between pat down concerns and the other two codes. When pat down concerns got mentioned, of the topics coded, the most likely co-occurring code referenced justification.

Exhibit 10 Correlations between content codes applied to rejection message field

Manifest content dummy variable	Variable name	Correlations			
Rejection message contains text related to pat downs	rejpatany	1.0000			
Rejection message contains text related to justification	rejdnyany	0.762***	1.0000		
Rejection message contains text related to search	rejseaany	0.5468***	0.0337	1.0000	
Rejection message contains text related to reasonable articulable suspicion	rejasany	0.5295***	0.3918***	-0.1434	1.0000

Note. CPD multi-version stacked ISR file for Period 2 (July-December, 2016). File provided September, 2017. Tetrachoric correlations between content codes in multi-version records. Each record = one contact card or version of an ISR, not one ISR. Ultimate version of each ISR excluded. N=3,177. *** = p < .001 (two tailed).

One implication of the strong correlations among the different codes is that the significance levels in the multivariate models, where each outcome is treated as a separate independent outcome, need adjustment to avoid an inflated experiment-wise type I error rate. A Bonferroni adjusted alpha level of $p < .0125$ rather than $p < .05$ is used.

¹⁴ See Period 2 Post Stop Outcomes report.

9.1.4 Links between manifest content and race and ethnicity of detainee and locale

The connection between the presence or absence of each topic, and detainee race and ethnicity, and beat-within-district racial and ethnic composition, were examined in two different ways.

To get at the gross impacts of race and ethnicity, the connection between each content variable and each of the four key race and ethnicity variables – detainee Black and non-Hispanic or not, detainee Hispanic or not, beat predominantly Black non-Hispanic or not, and beat predominantly Hispanic or not – is investigated using simple two-way cross tabulation tables.

To get at net impacts of race and ethnicity, after controlling for other factors, mixed effects logit models were conducted for each content variable. Where feasible, these models simultaneously entered detainee race and ethnicity variables, beat racial and ethnic composition variables, additional detainee factors (age, gender), stop factors (weekend occurrence, pedestrian stop, time of day, event code missing), and additional beat demographic factors (residential socioeconomic status, residential stability). When all these predictors were entered simultaneously, all variance inflation factors were less than 4.0, suggesting multicollinearity was not a concern.

9.1.4.1 *Gross impacts of race and ethnicity of detainee and beat: Bivariate models*

The gross impact of race and ethnicity on each of the rejection message content codes employ "flat" models based simply on the 2 x 2 contingency tables. A simple likelihood ratio chi-squared value indicates whether the connection between the topic mentioned, and race or ethnicity of detainee, or between the topic and racial or ethnic composition of beat, was stronger than would be expected just due to chance association.

Exhibit 11 provides an overview of the gross impacts of detainee race or ethnicity, or beat race or ethnicity, on the themes captured from text in the rejection messages field. Three statistically significant gross impacts surface. In beats that were 70 percent or more Black non-Hispanic in their residential composition, text concerned with justification, and text concerned with pat downs, both appeared significantly ($p < .001$) more frequently. Both these ecological connections also appeared, at or below the adjusted significance level ($p < .0125$) if tables were constructed for two separate random halves of the data (detailed results not shown). Therefore, these two ecological connections between beat racial composition and pat down and justification themes prove robust across two separate sets of cases. The percentage differences, described in the notes to the exhibit, are on the order of differences of two to three percent. Whether those size differences merit attention as a practical matter certainly deserves discussion. That discussion should simultaneously consider the relatively low base rate at which both the pat down and justification themes appeared in the rejection message field.

Detainee race linked to one theme, whether the pat down was mentioned in the rejection message. This theme surfaced significantly more frequently if the detainee was Black and non-Hispanic compared to when he or she was not. This significant connection, however, did not replicate when two random, independent halves of the data set were analyzed separately. Therefore, the link between the pat down theme and detainee race failed to prove robust across two independent subsets of cases.

Theme mentioned	Feature of detainee or stop context considered			
	Detainee		Beat	
	Detainee BNH	Detainee Hispanic	Beat predominantly BNH	Beat predominantly Hispanic
RAS	p=.55 (ns)	p = .42 (ns)	p =.04 (ns)	p = .19 (ns)
Search	p = .13 (ns)	p = .17 (ns)	p = .03 (ns)	p = .36 (ns)
Justification / lack	p = .06 (ns)	p = .05 (ns)	p < .001 ^a	p = .30 (ns)
Pat down	p = .010 ^b	p = .10 (ns)	p < .001 ^c	p = .73 (ns)

Note. July-December 2016 multi-version ISRs. Last version of each multi-version ISR excluded. N = 3,177 individual records (individual versions or contact cards).

Entries show the level of statistical significance associated with the Likelihood Ratio χ^2 (df=1) for the 2x2 table, using all versions, save the last, for all ISRs.

Expected frequencies may be less than 5 in some of the 2x2 contingency tables, making the χ^2 test suspect.

Theme mentioned was extracted from the rejection message field.

BNH = Black and non-Hispanic. Predominantly means beat was 70 percent or more BNH or 70 percent or more Hispanic. Bonferroni adjusted alpha level = .0125. Tables showing a significant gross impact at the adjusted alpha level in **bold**.

^a In the predominantly BNH beats, this theme appeared in 2.5 percent of the rejection messages, but in the beats that were not predominantly BNH, it occurred in 0.8 percent of the rejection messages. If the same table is run for two separate random halves of the data, p = .004 in the first random sample and .006 in the second random sample.

^b This theme occurred in 13.9 percent of the rejection messages associated with BNH detainees, but in only 10.5 percent of the rejection messages associated with detainees who were *not* BNH. If the same table is run for two separate random halves of the data, p = .02 in the first random sample and p = .157 in the second random half.

^c In the predominantly BNH beats, this theme appeared in 15 percent of the rejection messages, but in the beats that were not predominantly BNH, it appeared in 10.1 percent of the rejection messages. If the same table is run for two separate random halves of the data, p = .001 in the first random sample and p = .008 in the second random sample.

9.1.4.2 Net impacts of race and ethnicity of detainee and beat: Multivariate models

When all detainee features, stop features, and beat features are considered, no significant impacts of detainee race, detainee ethnicity, beat racial composition, or beat ethnic composition emerged on any of the four rejection message themes examined (detailed results not shown). The *net* impacts of detainee and beat ethnicity and race are not noteworthy.¹⁵

¹⁵ Several estimation difficulties emerged with these models. Mixed effects models would sometimes not converge. And single level logit models with error terms clustered at the ISR report level sometimes provided non-sensical results. All these models were estimated excluding the final version of each ISR, since that final version contained no rejection message.

10 OVERALL ASSESSMENTS OF MULTI-VERSION CODED RECORDS

10.1 PREFACE

10.1.1 Question sequence and relevant cases: implications for examining differences

The logic and sequencing of the questions addressed in the multi-version coding form (Appendix B) merits mention. At the end of the series of questions that applied to single version and multiple version ISRs, the Consultant then turned to a separate series of questions *specific* to the multiple version ISRs. These start at Question 20 on the form. This section looks at the answers to these questions across the 173 sampled, coded and re-linked multi-version ISRs.

The first multi-version question (Q20) applied to all coded multi-version ISRs; responses indicated whether there was a problem in determining RAS from the narrative. This question acted as a filter. If a problem was not identified, the Consultant went to the next-to-last question on the survey, which just reported the finalized status code of the ISR (Q25).

Given the above skip pattern, and given that most sampled and coded multi-version ISRs did *not* reveal a problem with RAS determination, the numbers of ISRs coded on the questions after the initial multi-version question, Question 20, are quite small. Given those small numbers of coded records for questions later in the multi-version question series, links between detainee or beat race and ethnicity only are described for this first question, Question 20. After that, given the small numbers, we just report how various subsets of multi-version ISRs scored on the various questions.

10.1.2 Reports on weighted and unweighted data

A note on weighted and unweighted data. When data are weighted, each ISR is given a weight that makes its relative contribution to all the sampled ISRs proportional to the relative contribution of the specific ethnoracial group in question (Black non-Hispanic, White non-Hispanic, Hispanic) in the full set of records from which these records were sampled. In other words, with weighted data the patterns seen apply to the full set or population of records from which this sample was drawn. When results mention unweighted data, that references only the sample records themselves, and not the full population of ISRs for these three ethnoracial groups for the period.

10.2 ERROR INVOLVES RAS (QUESTION 20)

The first question in the multi-version question series asks, “Does the identified error in the final version concern the legal narrative RAS determination?”

10.2.1 Frequency of concern

The assessment appears in Exhibit 12. About a quarter to a fifth (24 percent weighted; 21 percent unweighted) of the errors related to reasonable articulable suspicion (RAS).

		Weighted		Unweighted	
		N (weighted)	Percent	N	Percent
Does the error in the final version concern RAS determination (Q20)?	No	134.1	75.0	132.0	76.3
	Yes	43.7	24.4	37.0	21.4
	Missing	1.1	0.6	4.0	2.3
	Total	178.9	100.0	173.0	100.0

Source: Sampled and coded multiple-version ISRs for Period 2, July-December, 2016. Weighted numbers rate ISRs for each ethnoracial group in proportion to their contribution to all ISRs for the period.

Because this is the first time this number has been examined, there is no way to benchmark whether this is too high or too low. The number that describes the full set of multi-version ISRs, for the period, 24 percent, is just an estimate. There is uncertainty around that estimate, because of sampling error. We are not showing that uncertainty at this point. **The only suggestion being made at this juncture is that when Period 2 ISRs required revision, in about one out of four cases, roughly, RAS-related errors remain in the final version.**

10.2.2 Patterning across race or ethnicity of detainee or beat

Were multi-version ISRs from different ethnic or racial groups, or in beats with different ethnic or racial composition, more or less likely to have errors concerning RAS determination? This question is examined *descriptively only* given the small numbers of ISRs here. Further, given the small numbers, the examination considers just one race or ethnicity factor at a time, and does not control for additional factors. In short, gross impacts of race or ethnicity, and gross impacts of racial or ethnic composition of the beat, are just described. Results appear in Exhibit 13.

The percentage of ISRs with this concern seems about five percent *higher* if a Black non-Hispanic detainee is contrasted with other detainees (26 percent vs. 20.5 percent based on weighted data). It seems about seven percent higher if stops in predominantly Black non-Hispanic beats are contrasted with stops in other beats (28 percent vs. 21 percent based on weighted data). This latter difference also is graphed in Exhibit 14.

Exhibit 13. Ethnoracial patterning of RAS errors (weighted data)

		Does the error in the final version concern RAS determination (Q20)?	
		No	Yes
Detainee			
Black non-Hispanic	No	79.5	20.5
	Yes	74	26
Hispanic	No	75	25
	Yes	78	22
Beat			
70 percent or more Black non-Hispanic	No	79	21
	Yes	72	28
70 percent or more Hispanic	No	74	26
	Yes	81	19

Source: July-December 2016 sampled and coded multi-version ISRs.

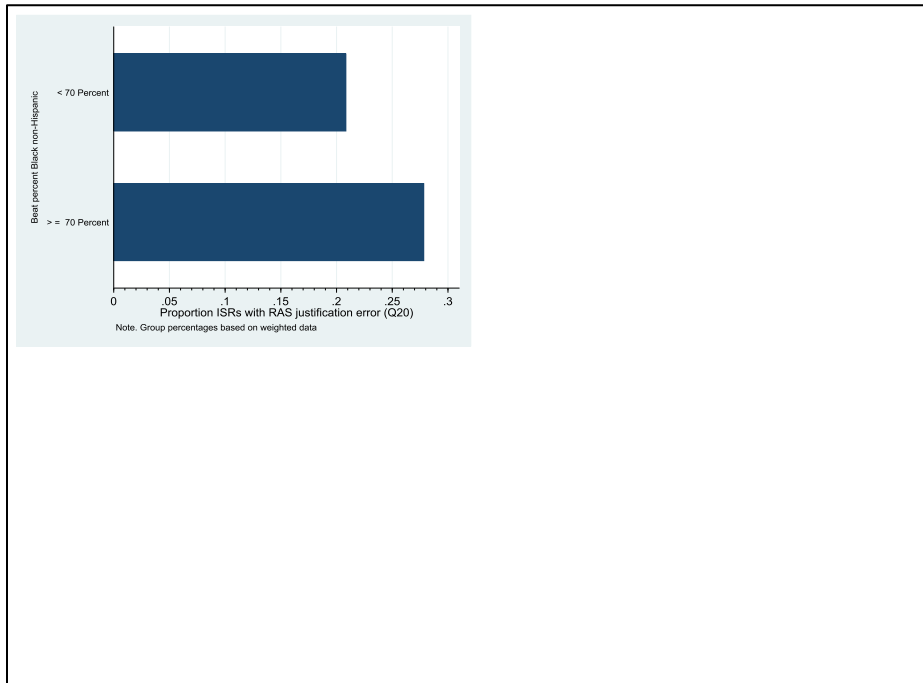
Note. Percentages based on weighted data.

Note. Percentages sum to 100 for each group (row percentages)

Speaking descriptively, if race is in question, either for detainee groups of interest or the places of interest, this concern pops up about five to seven percent more often.

But the reverse appears for ethnicity. The RAS concern surfaces **less frequently** if the detainee is Hispanic (22 vs. 25 percent based on the weighted data) or the beat where the stop took place is predominantly Hispanic (19 vs. 26 percent based on weighted data).

Exhibit 14 RAS justification error rates and beat racial composition (weighted data)



Source: July-December 2016 sampled and coded multi-version ISRs.
Note. Proportions based on weighted data.

10.2.3 Officer correctly articulates RAS (Q21)

At this juncture, as mentioned above, the numbers get quite small, because of some of the following questions applied only if the overall RAS question (Q20) flagged a concern.

For those ISRs where RAS determination was a concern, **Question 21** asked “Did the submitting officer correctly articulate RAS?”

About three quarters of the time the officer’s articulation proved **unsatisfactory** (72 percent, 31 out of 43 based on weighted data; 78 percent, 29 out of 37 ISRs based on unweighted data). For this particular set of ISRs, officers only got it right about a quarter to a fifth of the time (28 percent, 12 out of 43 based on weighted data; 22 percent, 8 out of 37 based on unweighted data).

10.2.4 Reviewer flags RAS error

Question 22 asks “Did the 1st Reviewer correctly identify the RAS error?”¹⁶

¹⁶ The form noted:

The “RAS error” can refer to the Terry Stop, protective pat down or the search. The objective here is simply to determine if the reviewing supervisors are correctly identifying the legal errors being made by some officers; the objective at this time is NOT to identify the types of errors being made by the submitting officers in the ISR versions prior to the final version.

Most of the time the 1st reviewer did flag up an RAS error for the subset of records where RAS justification remained a concern in the final version (Q20=yes). In 73 percent of the weighted cases (29 out of 39) and 77 percent of the unweighted cases (27 out of 35) reviewers had flagged the RAS concern.

In short, when the Consultant saw a **final version problem** with RAS determination based on the legal narrative, the first CPD personnel reviewing the record also saw the RAS error about three quarters of the time.

This also means, conversely, in about one out of four cases where the Consultant saw an RAS error, the 1st reviewer did *not* flag the same concern.

10.2.5 Sending the ISR back (any reviewer)

Question 23 asks: “Did any Reviewer send the ISR back to the submitting officer?”¹⁷

Roughly nine times out of ten the reviewer did send it back (91.6 percent, 36 out of 39 based on weighted data; 91.4 percent, 32 out of 35 based on unweighted data).

Of course, the question comes up: what is happening in these roughly one out of ten cases where RAS determination is an issue but the ISR was not sent back?

10.2.6 Officer remediation

Question 24 asks: “Did the officer fix the identified error?”¹⁸

So, if there is a problem with RAS determination, and if the ISR is returned to the submitting officer for correction, is the form corrected?

Yes, in roughly two cases out of three (65 percent, 23 out of 36 cases with weighted data; 72 percent, 23 out of 32 with unweighted data).

But in roughly one out of three instances, for this set of records, the submitting officer failed in remedying the error.

10.2.7 Changing RAS

Question 26 asks: “Did the officer supplement or change the facts concerning RAS with information that did not appear elsewhere on the ISR as originally submitted?”

Among the set of ISRs coded on this question, **the Consultant saw this as an issue in about two thirds of the cases.**

With the weighted data, 63 percent of ISRs (23 out of 36) were coded affirmatively by the Consultant on this. With the unweighted data the number was 68 percent (21 out of 31).

¹⁷ The form noted:

If the first reviewer sent the ISR to the IU for a determination before resubmission to the officer, note only whether the officer ever had an opportunity to correct the ISR – it doesn’t matter who sent the ISR back to the officer, only that the officer revised the initial submission.

¹⁸ The form noted:

It does not matter how many “versions” it took to reach the correct determination. The # of versions can be determined apart from coding an answer.

This suggests information that did not appear elsewhere on the ISR as originally submitted **appears in later versions of ISRs with considerable frequency as officers seek to address concerns about RAS flagged by reviewers.**

10.2.8 Summing up on RAS concerns

This investigation of RAS concerns in multi-version sampled and coded ISRs relies on a small but nonetheless representative cross-section of records. Because the sample size is small, and for most questions gets markedly smaller because of skip patterns, inferences back to the full set of records from which the sample was drawn are extremely rough. Descriptive estimates reported here and based on weighted data do describe the full set of records. But those estimates have a lot of uncertainty built in to them - uncertainty that naturally arises from sampling error. That's why the numbers reported here should be considered extremely crude.

Results suggest RAS concerns do contribute to and **persist in** ISR reports being revised. In roughly a quarter to a fifth of the multi-version records requiring revision, concerns did surface about RAS, given the narrative available. It is not clear currently why RAS determination remains an ongoing concern for a subset of stop reports.

That said, the frequency with which the Consultant flagged RAS concerns linked descriptively but in opposite ways to race and ethnicity. When the ISR was either about a Black non-Hispanic detainee, or a stop taking place in a predominantly Black non-Hispanic beat, RAS concerns were noted for slightly higher percentages of the corresponding group; the differences were on the order of 5 to 7 percent. But, when the ISR was either about a Hispanic detainee (a 3 percent difference), or a stop taking place in a predominantly Hispanic beat (a 7 percent difference), RAS concerns were flagged *less* frequently.

The importance of these opposite direction race and ethnicity links is not currently clear. The differences are just noted descriptively. If deemed worth pursuing, that pursuit could be reflected in a modified sampling strategy for multi-version records for the next reporting period.

Digging deeper and focusing just on the subset of records where RAS was a concern, revealed the following.

- In about three out of four cases, the submitting officer failed to satisfactorily articulate RAS.
- In about three out of four cases, the Consultant agreed that the initial CPD personnel reviewing the reports did highlight the RAS error. Conversely, in about one out of four cases the CPD reviewer did not see what was a concern to the Consultant.
- If there was an RAS concern, about nine times out of 10 the CPD reviewer did refer the ISR report back to the submitting officer for remediation.
- If it was referred back to the submitting officer, that officer was able to fix it in about two out of three cases, and not able to fix it in about one out of three cases.
- When officers sought to remedy the RAS concerns, about two thirds of the time later versions include information *that did not appear in the original submission.*

10.3 FINAL STATUS CODES

CPD uses a number of codes to indicate the final disposition of each ISR. These are listed under Question 25 in the appendix with the multi-version coding form.

Most of the data were missing. Only a small group of multi-version ISRs had a finalized status code.

With the unweighted data (n=173),

- 81 percent (141) were missing
- 18 percent (31) were coded APR; and
- 0.6 percent (1) was coded REJ.

With the weighted data (n=179),

- 80 percent (143) were coded missing
- 19 percent (34) were coded APR and
- one percent (2) were coded REJ

11 DISCUSSION

11.1 THE NOVELTY ISSUE

In thinking about the questions posed here and answers provided, readers should bear in mind the novelty of the current undertaking. Over the decades, scholarship on driving while black or stop and frisk practices have analyzed a wide range of indicators reflecting how the police are doing. But this report introduces at least one new indicator to that literature: the number of versions generated for each multi-version report. Because this is a new indicator, were not sure what it means. Scholarship required to establish its meaning – its construct validity – has not yet been done. Consequently, patterns suggesting racial or ethnic differentials in the number of versions created for multi-version reports should be interpreted with considerable caution. We don't yet have a comprehensive picture of the factors that lead to multi-version reports being revised more times.

Also new to the literature on stop and frisk and questions of racial disparities are the legal assessments made by the Consultant about the processing of RAS related report deficiencies. The authors are not aware that other ongoing or recent cases investigating this topic, such as the reports by Rudovsky and his team pursuant to the Bailey et al. v. City of Philadelphia consent decree, are making these kinds of legal assessments (Rudovsky, Messing, & Lin, 2017).

11.2 BROAD OUTLINE

Broadly speaking, the current investigation worked on three sets of issues: racial or ethnic links with varying numbers of versions for revised ISR reports; racial or ethnic links with particular

text reflecting why particular reports needed revision; and legal assessments of RAS deficiencies, and their processing, made by the Consultant.

11.3 RACIAL LINKS WITH NUMBER OF VERSIONS

Most multi-version reports, roughly 95 percent, required only one or two revisions for a total of 2 to 3 versions. A small number required up to five or six versions. The number of versions links significantly, and in both random samples, to both beat ethnic and racial composition. The expected number of versions was significantly higher if the beat was 70 percent or more Black and non-Hispanic in residential makeup as compared to less than 70 percent Black and non-Hispanic in residential makeup. The expected number of versions was also higher if the beat was predominantly Hispanic as compared to not predominantly Hispanic. See Exhibit 5. Although both findings are robust across both random samples of data, they are limited in other ways. The significance pattern depends on the model type, and the particular set of predictors included. When more demographic predictors are included, the racial composition link or the ethnic composition link shows up in only one sample. See Exhibit 7.

From a theoretical point, we would want to know if the links persist after we control for concentration effects. For example, does the link arise because the beat is predominantly Black and non-Hispanic, or does it arise because the beat is surrounded by other beats that are predominantly Black and non-Hispanic? At this juncture, we don't know if the link is because of the beat racial feature itself, or because of the types of surrounding beats, or both.

But, as mentioned above, before flagging this as either a policy and practice concern, we need to know much more about what it means to generate more versions of a report. Future work could employ stratified sampling protocols, sampling large numbers of multi-version ISRs from predominantly Black and not predominantly Black beats, and predominantly Hispanic and predominantly non-Hispanic beats, and learn more about the changes those ISRs go through as they are revised.

11.4 KEYWORDS IN REJECTION MESSAGES

The current work explored the frequency with which keywords showed up in the rejection message field, to try to learn more about some of the issues the supervisors were communicating to officers who originally submitted reports which needed revisions. The content analysis here was quite crude, just looking at how frequently concerns related to RAS, justification, patdown, or search appeared in these rejection messages. Because the analysis just looks at keywords, and not at keywords in context, there are construct validity or meaning questions here as well.

Words in the rejection message addressing either justification, or pat downs, appeared significantly more frequently for stops taking place in predominantly Black non-Hispanic police beats. See Exhibit 11. These two links proved robust across two random samples of the data. But they did not remain significant when additional factors either about the beat or about the detainees were added to the models. So, this seems to be just a bivariate connection between beat racial composition and reviewers touching on these two concerns. Racial composition of beat does *not* link to these keywords after we control for several other beat features.

Future work on this issue requires complementing the manifest content analysis here with a latent content analysis using trained coders looking for defined themes. It would also prove helpful in framing such a future content analysis if the parties weighed in on the types of words and themes they think deserve to be investigated in these rejection message fields. Of course, such future efforts should also look to the data themselves to see what themes are suggested from the source.

Such a future effort combining manifest and latent content could advance our understanding of the meaning of various messages communicated in the rejection message field. At the same time, if a big deal is made about analyzing what is said in these fields, that could have a shaping effect on what reviewers are willing to put down in writing in these fields. Careful consideration is needed for framing how to proceed on this matter.

11.5 LEGAL CODING AND RELATED PROCESSING CONCERNS

Turning to the legal coding of the sampled multi-version ISRs reviewed by the Consultant, the results show a range of potential concerns. To just briefly summarize: RAS determinations are problematic in about a fifth to a quarter of the sample records; further, among the records with RAS concerns, several additional processing shortcomings surface for some fraction of records. Finally, patterns of RAS problems have a disadvantaging racial link but an advantaging ethnic link.

Since this is the first time this assessment has been done of multiple versions of ISR reports, it's not clear if officers now are doing better than they were previously. It is also not clear if, with further training protocols, the remaining processing shortcomings can be remedied. In addition, the number of sampled multi-version records was too small to explore geographic variation with any precision. It might well be that these processing shortcomings surface most frequently in particular districts, and that further training efforts should be concentrated in just some locations.

To learn more about the locations, causes, and remedies of either RAS deficiencies or processing shortcomings, future work will need to sample and code much larger numbers of multi-version ISR records, constructing samples in such a way that it will be easier to address key questions like geographic concentration. This probably will require discussion among the parties.

Most troubling with the sample multi-version records was the absence of finalized codes for the majority of those ISRs.

11.6 CHANGES IN OTHER FIELDS

In addition to the initiatives warranted in future investigation to clear up some of the questions surfacing here, another future avenue could examine the frequency with which particular fields in the multi-version ISR reports change as we move from earlier to later versions. At this juncture the focus has been on changes in the narrative field, and in select content appearing in the rejection messages field. It's not yet clear how often changes are happening in other fields in the report that are of substantive interest given the focus of the Agreement.

11.7 CLOSING COMMENT

In closing, because this report investigates indicators which are new in two ways – new to the data analysis providing support to monitoring the Agreement, and new to this field of scholarship more generally – it's extremely difficult to translate what we see here into a grade about how the CPD is doing. The field of scholarship provides no comparable benchmarks, nor are we aware of benchmarking data available from other departments.

Clearly, the lack of final status codes for most of these records is a strong concern. But beyond that, it's hard to frame the practical and policy significance of the levels and patterns seen here.

This is not to say that some of the patterns seen here are not concerning. They are.

But the challenge is trying to decide how much concern and, concomitantly, how much follow up investigatory effort should be mounted.

12 REFERENCES

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74, 829-836.

Goldkamp, J. S. (2010). Construct validity: The Importance of understanding the nature of the intervention under study. In A. R. Piquero & D. Weisburd (Eds.), *Handbook of Quantitative Criminology*. New York: Springer Science and Business Media.

Grüber, S. (2015). Empirical Analyses of the Perspective of Financial Analysts on Information about Intangibles *Intangible Values in Financial Accounting and Reporting: An Analysis from the Perspective of Financial Analysts* (pp. 166-400). Wiesbaden: Springer Fachmedien Wiesbaden.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, Mass.: Addison-Wesley.

Klinger, D. A. (1997). Negotiating order in patrol work: An Ecological theory of police response to deviance. *Criminology*, 35(2), 277-306.

Peterson, R. D., & Krivo, L. J. (2010). *Divergent Social Worlds: Neighborhood Crime and the Racial-Spatial Divide*. New York: Russell Sage

Rudovsky, D., Messing, P., & Lin, S. (2017). *Plaintiffs' seventh report to court and monitor on stop and frisk practices: Fourteenth Amendment racial disparity issues. Mahari Bailey et. al, v. City of Philadelphia, et al., in the United States District Court for the Eastern District of Pennsylvania, C.A. No.10-5952*. Retrieved from

Taylor, R. B. (1994). *Research Methods in Criminal Justice*. New York: McGraw Hill.

Taylor, R. B. (2011). Communities and crime: construct validation deficits, a paucity of comparisons, and a Boudon-Coleman metamodel resolution. In J. MacDonald (Ed.), *Measuring Crime and Criminality* (pp. 71-100). New Brunswick: Transaction Publishers.

13 APPENDIX A: STATA PROGRAMMING FOR CAPTURING SOME BASIC PHRASES FROM REJECTION MESSAGE FIELD

```
* DIFFERENT VERSIONS OF REASONABLE ARTICULABLE SUSPICION
gen rejras1=strpos(rejectionmessageactionstake,"RAS " )
gen rejras2=strpos(rejectionmessageactionstake," ras " )
gen rejras3=strpos(rejectionmessageactionstake,"reasonable articulable suspicion" )
gen rejras4=strpos(rejectionmessageactionstake,"REASONABLE ARTICULABLE SUSPICION" )

gen rejrasany=0
replace rejrasany=1 if((rejras1!=0)|(rejras2!=0)|(rejras3!=0)|(rejras4!=0))
label var rejrasany "ras mentioned in rejection message"
tab2 rejrasany rejras1 rejras2 rejras3, firstonly missing

* JUSTIFICATION
gen rejdnj1=strpos(rejectionmessageactionstake,"does not justify")
gen rejdnj2=strpos(rejectionmessageactionstake,"DOES NOT JUSTIFY")
gen rejdnj3=strpos(rejectionmessageactionstake,"NOT JUSTIFIED")
gen rejdnj4=strpos(rejectionmessageactionstake,"not justified")
gen rejdnj5=strpos(rejectionmessageactionstake,"JUSTIFICATION")
gen rejdnj6=strpos(rejectionmessageactionstake,"justification")
gen rejdnj7=strpos(rejectionmessageactionstake,"justify")

gen rejdnjany=0
replace rejdnjany=1
if((rejdj1!=0)|(rejdj2!=0)|(rejdj3!=0)|(rejdj4!=0)|(rejdj5!=0)|(rejdj6!=0)|(rejdj7!=0))
tab2 rejdnjany rejdj1-rejdj6, firstonly missing
label var rejdnjany " justification / lack mentioned in rejection message"

* PAT DOWN MENTION
gen rejpat1=strpos(rejectionmessageactionstake,"pat down")
gen rejpat2=strpos(rejectionmessageactionstake,"PAT DOWN")
gen rejpat3=strpos(rejectionmessageactionstake,"pat-down")
gen rejpat4=strpos(rejectionmessageactionstake,"PAT-DOWN")
gen rejpat5=strpos(rejectionmessageactionstake,"patdown")
gen rejpat6=strpos(rejectionmessageactionstake,"PATDOWN")
gen rejpat7=strpos(rejectionmessageactionstake,"patted down")
```

```
gen rejpat8=strpos(rejectionmessageactionstake,"PATTED DOWN")

gen rejpatany=0
replace rejpatany=1
if((rejpat1!=0)|(rejpat2!=0)|(rejpat3!=0)|(rejpat4!=0)|(rejpat5!=0)|(rejpat6!=0)|(rejdj7!=0)|(re
jpat8!=0))
tab rejpatany
tab rejpatany rejdjany
label var rejpatany "pat down mentioned in rejection message"

* SEARCH MENTION
gen rejsea1=strpos(rejectionmessageactionstake," SEARCH ")
gen rejsea2=strpos(rejectionmessageactionstake,"search ")
gen rejseaany=0
replace rejseaany=1 if((rejsea1!=0)|(rejsea2!=0))
tab2 rejseaany rejsea1 rejsea2, firstonly
label var "search mentioned in rejection message"
```


14 APPENDIX B: CODING INSTRUMENT USED FOR MULTIPLE VERSION INVESTIGATORY STOP REPORTS

The multi-version coding form appears in a different file:

[2RP_legal_narratives_PDF_v_10_extended_w_multi_v.pdf](#)

That file will be linked for the final pdf.