STATE OF CALIFORNIA
DEPARTMENT OF MOTOR VEHICLES

California's 3-Tier Pilot Process Analysis Appendix

March 2010

Author: Bayliss J. Camp, Ph.D.
Research and Development Branch
Licensing Operations Division

RSS-09-229

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2010	Final Report

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
California's 3-Tier Pilot Process Analysis Appendix	
6. AUTHOR(S)	
Bayliss J. Camp, Ph.D.	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
California Department of Motor Vehicles Research and Development Branch P.O. Box 932382 Sacramento, CA 94232-3820	CAL-DMV-RSS-09-229
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
Office of Traffic Safety; 2208 Kausen Drive., Suite 300;Elk Grove, CA 95758- 7115	
11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report presents the descriptive and predictive analyses of: (i) the results of a survey (n = 130) conducted of California Department of Motor Vehicles (CA DMV) Field Office and Driver Safety Branch staff and managers participating in the 3-Tier Pilot project; (ii) interviews (n = 49) conducted of CA DMV Field Office and Driver Safety Branch staff and managers participating in the 3-Tier Pilot; (iii) the results of a survey (n = 5,777) conducted of customers participating in the 3-Tier Pilot; and (iv) the robustness of the Pelli-Robson contrast sensitivity chart by location and technician (n = 9,934). These analyses form the primary evidentiary basis for some of the findings and conclusions presented in the 3-Tier Pilot Process Analysis Report.

14. SUBJECT TERMS	15. NUMBER OF PAGES
3-Tier Pilot, Driver Competency, License Renewals, Licensing Tests, Licensing Assessment System, Functionally Limited Drivers	204
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	None

PREFACE

This report is issued as an internal monograph of the California Department of Motor Vehicles Research and Development Branch rather than as an official report of the State of California. The opinions, findings, and conclusions expressed in this report are those of the author and not necessarily those of the state of California, the California Office of Traffic Safety, or the National Highway Traffic Safety Administration.

ACKNOWLEDGEMENTS

This project was made possible with the support of the California Office of Traffic Safety and the National Highway Traffic Safety Administration (OTS Grant #TR0703).

This study was conducted under the general direction of David DeYoung, Chief of the Research and Development Branch of the California Department of Motor Vehicles, and under the supervision of Leonard Marowitz, Manager of the Alcohol and Drug Projects Section.

The author would like to thank Debby Atkinson, retired annuitant; Gwen Bridges, Manager in Field Operations; and E. Torricel Taylor, Manager in Administrative Services, for their comments and critiques on previous drafts of the constituent modules of this appendix.

The author would also like to thank Doug Rickard, Douglas Luong, and Debbie McKenzie for their assistance at various stages of this project. Their expertise was particularly helpful in data collection, database management, and preparing the manuscript for publication.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

TABLE OF CONTENTS (continued)

LIST OF TABLES (continued)

INTRODUCTION

In accordance with California Vehicle Code Section 1659.9, the California Department of Motor Vehicles (CA DMV) recently (5/1/07-12/31/07) conducted a field pilot of several new driver competency assessment tools that have shown promise for predicting traffic safety outcomes in prior studies (Hennessy & Janke, 2005). This pilot took place in six DMV field offices as well as one Driver Safety office in Northern California. The pilot study and its associated outcome and process reports serve two closely related purposes. The first purpose is to use the data collected during the pilot period to determine the large-scale predictive validity of these new assessment tools for reducing the crash risk and violation rates of California drivers. This outcome report is scheduled for publication in 2011. The second purpose of the 3-Tier Pilot was to test the feasibility of implementing, in a production setting, these new assessment tools. In order to determine the feasibility of implementation, the Research and Development Branch of CA DMV has undertaken the task of producing a process report, of which this document serves as the appendix.

This appendix contains four modules. Module #1 presents the descriptive analysis of the results of a survey (n = 130) conducted of CA DMV Field Office and Driver Safety Branch staff participating in the 3-Tier Pilot project. Module #2 presents the descriptive analysis of interviews (n = 49) conducted of CA DMV Field Office and Driver Safety Branch staff participating in the 3-Tier Pilot. Module #3 presents the descriptive and predictive analyses of results of a survey (n = 5,777) conducted of customers participating in the 3-tier Pilot. Module #4 presents a hierarchical logistic regression analysis of individual outcomes on the Pelli-Robson contrast sensitivity assessment, as predicted by chart location, technician, customer age, and possession of a previously-identified vision condition (n = 9,934). Together, these analyses form the evidentiary basis for many of the findings and conclusions presented in the 3-Tier Pilot Process Analysis Report.

MODULE #1: ANALYSIS OF THE STAFF SURVEY

ANALYSIS OF THE STAFF SURVEY

Introduction

At the termination of the field office portion of the 3-Tier Pilot, the Research and Development Branch (R&D) of California's Department of Motor Vehicles (CA DMV) distributed surveys to the staff and managers who implemented the project. The collection of data through this survey served four goals. Explicitly, the survey goals included: (a) gathering suggestions regarding revisions and improvements to the 3-Tier process itself, (b) gauging CA DMV staff perceptions of 3-Tier's impact on customer service, and (c) identifying potential areas of improvement for training. In addition, the distribution of this survey was implicitly intended to improve communication flows between R&D and both Field Operations Division and Driver Safety Branch. Included in this report are a discussion of the survey method and a summation of the major findings.

The substantive suggestions that staff and managers had for revisions and improvements to the 3-Tier process tended to cluster around two common concerns: first, ensuring the universality of any assessment tests, or in other words, requiring that all driver license renewal customers be subject to any new driver competency assessments. Secondly, staff made suggestions regarding how to increase the speed of customer processing times, and so to (ideally) reduce customer wait times in the field offices. Respondents to this survey tended to report either a neutral or mixed (both positive and negative) impact of 3-Tier on customer service. Negative impacts on customer service were generally identified with either (a) the increase in average customer processing time associated with 3-Tier, and consequently to potentially increased wait times in the field offices, or (b) the concern that 3-Tier did not apply universally to all customers. Specific concerns raised in relation to this latter point included the view that the 3-Tier process applied only (or disproportionately) to senior citizens and, secondly, did not apply to customers renewing their licenses through the use of a language other than English. Positive impacts on customer service were generally identified with either (a) an increase in the amount of individual attention given to customers, or (b) the identification of 3-Tier with an improvement in traffic safety. While there was substantial criticism regarding training, it was also clear that training was largely effective: Respondents reported that it took a relatively short

period of time (less than a month in most cases) to become comfortable with 3-Tier procedures. Moreover, they reported having relatively few questions about the process, post-training. That said, the qualitative comments reported here do reveal isolated instances of deviation from standard procedures, as presented both in training, and in post-training quality control.

Methods

Sampling Procedure

The author and three colleagues in R&D distributed the surveys to all six of the 3-Tier Pilot field offices and to the Sacramento Driver Safety office (DSO) during the Wednesday morning staff meeting on 10/31/07 and 11/7/07. On the day the surveys were distributed, the staff of each office was presented with an engraved, framed certificate in recognition of that office's participation in the program. These certificates were signed by the deputy directors of the two DMV divisions involved with the pilot: Field Operations and Licensing Operations. In addition, R&D provided those present with food and drink (muffins, bagels, and juice) as a small token of appreciation. The surveys were anonymous and participation in the survey was voluntary. See Sub-Appendix A for the actual text of the survey.

Although the sampling procedure was technically of a convenient nature, essentially the entire population of interest was surveyed. Almost no respondents declined to participate, though a few persons were absent, due either to vacation or illness. The final *N* was 130. See Table M1.1 for summary statistics of the sample population.

Analysis Techniques Used

Approximately half of the questions involved closed-ended Likert-type scales, with a range of responses from which the respondent had to pick one. The data produced from these questions are implicitly quantitative; hence much of the analysis involves simple descriptive statistics and cross-tabulations. Given the size and nature of the sample, however, in most cases statistical tests of significance are of only marginal utility.

Other questions asked for open-ended, more qualitative answers; many of these were follow-ups to a first, forced-choice half of the question. Qualitative responses were analyzed through the use of open coding procedures. On most questions, there was substantial fall-off in response rate between the quantitative and qualitative halves of questions. While nearly all respondents answered every forced-choice question, between one-third and one-half of respondents declined to fill out any given open-ended question. In the analyses presented in this paper, all figures are presented with the question-specific response rate; thus, percentages are of those respondents who responded to that question (or, in the case of two-part questions, the relevant half).

Table M1.1: Descriptive Statistics of Respondents to the 3-Tier Staff Survey

Office	*N* (%)	Job category	*N* (%)	Reported daily # of customers	*N* (%)
Carmichael	29 (21.5%)	MVFR	69 (53.1%)	1-2 customers per day	45 (35.7%)
Fairfield	10 (7.4%)	SMVT	8 (6.2%)	3 or more per day	81 (64.3%)
Folsom	17 (12.6%)	LRE	21 (16.2%)		
Sacramento -Broadway	30 (22.2%)	3-Tier Manager I[a]	6 (4.8%)		
Sacramento - South	16 (11.9%)	Manager (other)	14 (11%)		
Vacaville	16 (11.9%)	Hearing Officer	10 (7.7%)		
DSO	11 (8.1%)	Other/decline to state[b]	2 (1.5%)		
Unknown	1 (<1%)				
Total	130		130		126

a Two respondents who claimed to be 3-Tier Manager Is were re-coded as Managers (other). The number of customers these respondents reported seeing was abnormally low: 1-2 per day versus a mean of 8 customers per day for other 3-Tier Manager Is. Because of the small sample size, this affected some of the reported cross-tabulations, notably in TableM1.4B (Staff Views of 3-Tier's Impact on Customer Service, by Job Category).

b One of these was a Control Cashier.

Respondents who participated in this survey varied substantially in the nature and depth of their participation in the 3-Tier Pilot. Those holding different job categories, for instance, had very different duties when it came to processing 3-Tier customers. A Motor Vehicle Field Representative (MVFR) or a Senior Motor Vehicle Technician (SMVT) was responsible for administering the Tier 1 assessment tests (a simple memory recall exercise, a contrast sensitivity vision chart, and a structured observation of the customer's potential physical limitations). By contrast, the 3-Tier Manager I was responsible for administering some of the Tier 2 elements of the process (in particular, the educational intervention given to some customers). Office Managers, however,

might only rarely have direct contact with 3-Tier customers. Thus, in the analyses presented below, we differentiate between six job categories: MVFRs/SMVTs (whose responsibilities vis-à-vis 3-Tier were substantially the same), LREs, Driver Safety Hearing Officers, 3-Tier Manager Is, and Managers (other). This last category includes both Office Managers and Administrative Managers.

Participation also varied within job category. For instance, the majority of the MVFRs who participated in the pilot were "cross-trained." They might process a basic driver license renewal with one customer, followed by a vehicle registration with the next customer, following which they might assist a customer seeking to obtain a handicapped sticker for their vehicle (and so on, as contingent on the specific types of transactions in which the individual MVFR happens to have training). Depending on a number of factors, any given MVFR might see very few, or a great many, 3-Tier customers during the course of their work. Similarly, while only one of the LREs in any given office was responsible for administering drive tests for 3-Tier Pilot customers, other LREs in the office might, as part of their suite of duties, intermittently work at a window, and thus process customers in a manner similar to an MVFR.

This variation in participation is a key component to understanding the results of this survey for at least two reasons. First, participation in different components of the pilot resulted in different kinds of insights, suggestions, and critiques. Respondents typically wrote about those aspects of the process with which they were most familiar. Secondly, however, the depth of one's participation in the pilot, as measured by the frequency with which one had to implement 3-Tier procedures during a customer transaction varied substantially. This variation also serves as a marker of participatory familiarity with the process, as distinct from observations of how the implementation of the 3-Tier Pilot affected the work of others. Given this variation in participation, in some of the cross-tabulations presented below the results are shown grouped according to the self-reported average number of customers seen over the course of a typical day. For this purpose, responses to question #2 were collapsed from five categories to two: one to two customers per day (n = 45, 34.6% of the total sample), and three or more customers per day (n = 81, or 62.3% of the total sample). In tabulations not shown here (but available from the author upon request) additional analyses were run which subdivided responses into three categories instead of two (one to two customers per day, three to six customers per day, and seven or more customers per day): no substantial differences

from the results presented here were noted. While this grouping of answers by self-reported customer load overlaps other variables—in particular job category[1]—the overlap is only partial, and this particular measure illuminates certain patterns in the data which would otherwise remain obscure.

Limitations of These Data

These data are, of necessity, limited in scope and utility. Certain caveats should thus be kept in mind regarding interpretation. Some of the questions ask respondents to report on matters of which they had only partial knowledge. For instance, on question #5 (regarding customer service), a substantial number of respondents interpreted this question in reference to what they perceived to be the 3-Tier Pilot's impact on customer wait times. As noted in the main body of the project process analysis, the project's actual, as opposed to perceived, impact of 3-Tier on office wait times is difficult to estimate precisely. That said, the *perceived* impact of the pilot on customer wait times is critical to the success of implementation of any components of the pilot in the future. California DMV is committed to reducing, as much as possible, the amount of time customers spend in field offices, and any new procedures incorporated into basic field office practice must take that priority into account.

In addition to acknowledging the difficulties in interpreting respondent perceptions of 3-Tier, there is the additional difficulty of interpreting respondent reports of second-hand, hearsay, knowledge. Question #6, for instance, asked the respondent to report any comments they heard from customers regarding the pilot. In answering Question #7—on how "fair" the respondent thought the process to be—a substantial number of staff replied in their qualitative comments referencing what they had heard *from customers*. In module #3 of this appendix, the author analyzes customer responses to a short survey mailed out at the end of the project. That module thus directly reports the nature and distribution of customer views on the 3-Tier Pilot. The results presented here may partly tap some portion of customer concerns regarding the 3-Tier Pilot. However, the author would rather emphasize the degree to which these comments reveal the kinds of questions staff may expect to handle, and thus the kinds of customer service

[1] For instance, none of the Hearing Officers reported seeing more than two 3-Tier customers per day. By contrast, only 15% of the MVFRs reported processing fewer than three customers per day.

dilemmas likely to arise should any portion of the 3-Tier process be implemented in the future.

Finally, a word regarding the representativeness and generalizability of the survey sample. Essentially everyone of interest (all staff and managers that participated in the implementation of the 3-Tier Pilot) was sampled, and participated in the survey. Hence, we have a relatively complete picture of the views of those involved in the project, at least as measured by the questions used here. However, if the 3-Tier process is implemented statewide, we cannot assume that the staff of the seven (six field and one Driver Safety) offices included in the pilot are representative of the staff of DMV as a whole. DMV field offices are categorized into size grades, of which there are five, largely based on average customer load. No offices of Grade I or II (the two smallest categories) were included in the 3-Tier Pilot. It also bears emphasizing that, inasmuch as the staff understood that they were participating in a pilot project, as opposed to adopting new and permanent additions to office procedures, the results reported here must be taken in context. Given that staff and managers knew that 3-Tier might or might not be implemented in the future, their feedback during the pilot may differ in unknown ways from feedback given regarding a formally adopted change to DMV policies and procedures.

Results

Suggested Revisions to Forms

Approximately half (69/130) of respondents had suggestions for revising the forms. Among these, there were essentially three categories of responses: suggestions regarding the 3-Tier Driving Information survey (12 responses), suggestions regarding the 3-Tier Tracking Sheet and/or Tier 1 Score Sheet (35 responses), and "no suggestions" (i.e., the respondent answered that they thought the forms were fine, or otherwise needed no changes; this included 18 responses). A few answers (5) had to be re-coded as suggestions regarding the process, as they did not refer to the paperwork per se. Those are analyzed in a subsequent section, and are not included here. By the same token, one answer to Question #4 had to be re-coded as referring to forms and paperwork, and is included here. See Table M1.2 for a schematic summary of the

qualitative responses to this question. In response to this question, there appears to be little, if any, variation by self-reported customer load, nor by job category.

Table M1.2: Suggestions for Revisions to Forms

Type of suggestion	*N* of respondents (by customer load)		Examples of comments[a]
	1-2/ day	>2/ day	
"None"	9	9	"No — forms were very basic and to the point." "Good." "No – I think forms were easy to complete."
Survey-related	2	10[b]	"Make the questions on the survey clearer or understandable for the customer." "Survey questions were phrased in a manner which was confusing to customers." "Explain on the customer survey that all questions are not graded."
Tracking sheet- or score sheet-related	13	22	"If it can all somehow be just one sheet it will be easier to process." "Some redundancy. You have to say that the customer accepted even when you have all their stuff written down." "Eliminate the multiple response types and use only check-off boxes. No circles, no ones and zeros, no confusion."
Misc.	2	2	"Throw them away!" "Explain on forms that this was not targeting any certain group -- (seniors)."
Total	26	43	

[a] Question wording: "Think for a moment about the various forms and other paperwork that you may have used to collect data on and to process 3-Tier customers (for instance, the Score Sheet or the Tracking Sheet). Is there anything specific that you would suggest for how to improve these forms? Please be as specific as you can in your suggestions."

[b] This includes one respondent's answer to Question #4, re-coded here.

The 3-Tier Driving Information Survey, which was given to all customers, served two purposes. First, to gather information on those drivers who report restricting their own driving (i.e., not driving at night, or on freeways, etc.). These data will be used to supplement future analyses of traffic safety outcomes as regards crashes and violations. Secondly, filling out the survey at the counter provided an opportunity for the MVFR to observe the customer's ability to move their upper body. This second purpose was intended to be part of the structured physical observation that formed a key component of the 3-Tier process. No respondent raised any concerns regarding the usefulness of the survey for the latter purpose. Instead, those who raised concerns about the survey did

so largely around the clarity of the question wording. No specific questions were identified, but the survey questions were referred to as being "tricky" and "misleading." As 8 of the 10 respondents who raised this issue were MVFRs, it appears that the concerns stemmed from customers reporting trouble answering the survey, and so asking the technician for clarification.

In addition, one MVFR noted that (some) customers were guarded in their answers because of fears that the information provided might have implications for further testing. Inasmuch as the document in question was partly a research tool (re: self-restricting drivers), MVFRs may not have received training in how to answer customer questions as to the purpose behind the survey. That (some) MVFRs reported difficulty in answering customer queries regarding the survey probably indicates, if nothing else, that *if* some sort of writing exercise is included as part of the 3-Tier process, that it be of a nature such that technicians can easily address customer feedback. Currently, for instance, customers are encouraged to fill out any paperwork while they are waiting for their queue number to be called (this reduces, somewhat, the amount of time any given customer spends with a technician at a window, and so speeds up processing time). Perhaps some portion of the driver license application (DMV Form DL 44) might be "reserved" for completion at the counter; currently this includes only the customer's signature.

While the customer had to fill out the Driver Information Survey, the technician was responsible for filling out two closely related pieces of paperwork: the Tier 1 Score Sheet and the top portion of the first page of the 3-Tier Tracking Sheet. About half of those who made suggestions regarding the forms (35/69) directed their comments at one or both of these forms. These suggestions came in four types: in addition to concerns about clarity, length, and redundancy, a few respondents gave concrete and specific suggestions that essentially amounted to editing for clarity and ease of use. In general, the suggestions that were made had to do with eliminating redundant questions and reducing the number of pages to (ideally) one, to eliminate excess paperwork. Notably, only one person suggested getting rid of these forms entirely[2], and only two suggested incorporating them into the computer-based application retrieval and input system (the DMVA) used to process most types of basic customer transactions.

[2] This respondent was a Hearing Officer. This response was coded as "miscellaneous" rather than as a suggestion regarding the tracking sheet or score sheet.

Certainly any paperwork that can be made easier to complete, while still capturing necessary and vital information, should be. All forms used for future implementation should be thoroughly vetted for clarity of instructions, simplicity of presentation, and overall straightforwardness. However, the points raised by staff also raise a secondary problem. The process of using paper forms for the 3-Tier Pilot was, to some extent, deliberately archaic. Filling out forms by hand (instead of keying data into a computer) slowed down the renewal process partly in order to facilitate the structured physical observation of the customer by the technician.[3] This archaism introduced a slightly contradictory set of demands on the technicians who worked at the front-line windows. In the first instance, technicians are encouraged by various means to process customers as fast as possible, to reduce wait times and increase office efficiency and productivity. At the same time, however, the 3-Tier process encouraged the technician to carefully observe customers, through the use of simple assessment tools, for potential physical and cognitive limitations that might impact the customer's ability to drive safely. The suggestions made by staff regarding the tracking sheet in general appear to be made with any eye toward increasing the efficiency and productivity of customer processing. This includes filling out one form instead of two, reducing or eliminating redundancies, changing the question format to a series of check-off boxes rather than blanks that need to be filled in with numerical information. This will no doubt, at the margins, save customer processing time, especially if the questions on the Tier 1 Score Sheet are incorporated into the DMVA computer system. At the same time, altering the method of making structured physical observations for simplicity and speed will also, perhaps, tend toward cursoriness of observation.

Suggested Revisions to Process

Nearly three-fifths of respondents (77/130) made suggestions for revisions to the 3-Tier process. These covered nearly all aspects of the pilot program. In 11 cases, respondents said explicitly that they had no suggestions to make.[4] See Table M1.3 for a schematic summary of the qualitative responses to this question. The responses are grouped by

[3] In addition, of course, it would have been prohibitively difficult to re-program temporarily the DMVA terminals in the pilot offices.

[4] Of these 11 respondents, 8 also said explicitly that they had no suggestions to make regarding the paperwork.

self-reported customer load. There appears to be substantial variation between those who saw 3-Tier customers regularly and those who saw them rarely or not at all.

Ten respondents—all of them MVFRs or SMVTs and all of whom saw at least three customers per day—had critiques regarding the taking of a customer's picture for their license.[5] According to the training protocols, each technician was expected to walk to the camera area with their customer and take their picture. By doing so, the technician had an additional opportunity to observe the customer's gait and carriage for potential physical limitations. This was different from normal procedures, where one technician typically works the camera. The brunt of the comments in this area focused on the extra time this added to the process. If the structured physical observation is incorporated into the DMVA system, retaining this portion of the observation would be somewhat awkward (among other things, it would require the technician to keep their transaction "logged on" for the duration of the time they walked the customer to the camera). An alternative would be to require that the technician working the camera separately observe each customer for potential physical limitations.

Eight respondents (six of whom were MVFRs) reported concerns about the Pelli-Robson contrast sensitivity test (the so-called "fog chart"). The respondents who raised these concerns came from four of the six pilot field offices;[6] all of them saw at least three customers per day. These comments had to do with placement of the charts within the offices, especially with potential glare from windows, shadows, or different levels of ambient brightness. This raises a potential concern about the universality of this assessment, namely whether some customers were more likely to pass or fail depending on where in the office they were seen, or even what time of day they came into the office (i.e., if the sun was shining through a window onto the chart, as suggested by one staff member). This concern conflicts with evidence published in the academic literature (Zhang, Pelli, & Robson, 1989), which suggests that there should be no significant variation for outcomes on this assessment by levels of luminance. Given this, the issue raised here by staff has three implications.

[5] These concerns came from the three large offices (Carmichael, Sacramento-Broadway, and South Sacramento) and the Folsom office. The latter had a 3-Tier customer load more similar to the larger offices than to Vacaville or Fairfield.

[6] Carmichael, Folsom, Sacramento-Broadway, and Vacaville.

Table M1.3: Suggestions for Revisions to Process

Type of suggestion	*N* of respondents (by customer load)			Examples of comments[b]
	1-2/ day	>2/day	DTS[a]	
"None"	6	5	0	"None." "No changes."
Camera-related	0	10	0	"Walking them over to the camera…[it] really takes time…" "Taking the picture. We already see our customers walking up to us when we call them on the Q[ueue]."
Contrast sensitivity	0	8	0	"Placement of the fog charts…some customers could not pass, but when asked to use a better lighted chart, had no problems reading it."
Language	2	2	0	"Test all customers, not just English-speaking."
Perceptual response test	3	14	0	"[The PRT]…it seemed to cause the most problems with the customers and getting them to understand it." "…some of the customers do not use computers or have a hard time moving fast to respond." "Customers thought they were looking for the car or truck in the 'fuzzy screen'."
Memory recall exercise	1	6	0	"Stating SSN – the number is overheard by others (information security)." "…customers didn't want to disclose that information in front of other customers."
Suggestion	6	10	1	See text.
Misc.	0	2	1	"The weakest link is the technician (not sure how to change this part of the process)… not walking applicant to camera, additional observation, [not] completing forms properly, etc."
Total	18	57	2	

[a] DTS: These respondents did not answer the question regarding customer load.
[b] Question wording: "Aside from the forms and paperwork, if you could pick one part of the **process** that you think should be changed somehow, what would it be, and why?"

First, these concerns probably reflect some amount of customer reaction to the "newness" of 3-Tier procedures. In other words, staff may have faced questions about the Pelli-Robson chart, including perhaps questions regarding the effects of ambient lighting on a given customer's ability to read the chart. Staff may or may not have been prepared to answer these questions. At the very least, this suggests some potential revisions to training materials, to better prepare staff to handle inquiries from customers on this subject. Secondly, on a more practical note, it bears noting that current field office procedures call for allowing customers the opportunity to "switch stations" in the event of being unable to pass the standard Snellen visual acuity test.

This practice was adopted for the use of the Pelli-Robson chart during the 3-Tier Pilot. Therefore, to some degree this concern about the universality of the test, as dependent on chart location, was resolved through normal field office procedures.

However, this latter practice raises a third implication. If in fact a customer's outcome on this particular test *did* depend in part on ambient lighting, this may in turn have been a result of some underlying vision condition (e.g., cataracts, macular degeneration) which affected their contrast sensitivity ability. Which is to say, the *assessment was working as it ought to*. To draw a parallel, say DMV were to give multiple versions of the written law test, with varying levels of difficulty: an easy test, a medium test, and a hard test. If a person with poor knowledge of the law is given the "hard" test, is the problem that they don't know the rules of the road, or that they happened to get a test that was "too hard?" In the case of the actual written law test that the DMV currently uses, all versions are of equal difficulty, of course. In the case of the Pelli-Robson contrast sensitivity chart, the staff are here raising a question as to whether a person's outcome is a product purely of their own individual ability (their vision health) or the "difficulty" of the conditions under which they are tested. In order to address these concerns, the author undertook a formal examination of the passage rates at different fog chart locations (see Module #4 of this Appendix).

Four respondents raised the issue of language. Because of funding constraints for the pilot project, customers participating in 3-Tier were limited to those renewing their license in person, who were required to take the written test (typically because of a past record of traffic violations or because they were 70 years of age or older), and who chose to take the written test in English. This resulted in the exclusion from the pilot population of those taking the written test in a language other than English. Given the demographics of the field sites, it is likely that this included some immigrants from the former Soviet Union, South Asia, East and South-East Asia, and Latin America.[7] There is little reason to suspect that this will have any effect on the outcome analysis (i.e., the predictive validity of different components of the 3-Tier process for traffic safety). There is good reason, however, to believe that this had some affect on the size of the customer

[7] As of the writing of this module (March 2010), CA DMV makes the California Driver Handbook (the standard study literature for preparing for either the written law exam or the on-the-road drive test) available in 8 languages: Armenian, Mandarin Chinese, Korean, Punjabi, Russian, Spanish, Tagalog, and Vietnamese. The law test is available in audio format in 4 additional languages (Hindi, Hmong, Japanese, and Portuguese), and the written test is available in those plus 19 other languages.

load involved in 3-Tier processing (see the main body of the process analysis report). For the present purposes, it is enough to note that those staff raising this suggestion for amending the process appear to have been primarily concerned with the universality of the process, or the degree to which it applied to all customers renewing their license at a DMV office. As we shall see, this theme arises in answers to other questions as well. This theme likely derives from a key imperative expected of all DMV employees: to provide everyone with excellent customer service. The emphasis here appears to fall on the idea of giving *everyone* the same excellent service for which all DMV employees strive. Obviously, in the event of statewide implementation, the 3-Tier process will certainly be altered in accordance with procedures for testing in languages other than English.

Seven respondents raised concerns regarding the memory recall exercise. These respondents came from a range of job categories at three of the larger offices, and most saw three or more customers per day. This cognitive assessment tool involved two parts. At the moment a customer received their queue number, they were told that they would be asked to recall their social security number from memory (or their zip code, if they did not possess an SSN) when called to the window. When their number was called, they were then asked to write their SSN down, typically on the back of their license renewal application (DMV Form DL 1RN or DMV Form DL44). What the customer wrote down was then checked by the attending technician for accuracy against the number on the customer's driver record. Once the customer's transaction was complete, any paper materials containing sensitive personal information (such as a social security number) were disposed of securely. Of those who mentioned this tool, three cited concerns over privacy or information security. The four others suggested changing this part of the process, but did not explain for what reason, or how it ought to be changed. The comments given by respondents also reveal some significant deviations from training protocols: at least two staff members implied that they asked the customer to verbally state their SSN. This should not have occurred. Customers ought only to have been asked to write this number down on a paper application which was then, according to standard office procedures, shredded to protect the customer's privacy. Given these quite legitimate concerns regarding customer privacy and security, if a memory recall exercise proves useful as a combinatory element of an index of driver competency (cf. Hennessy & Janke, 2005), it may be possible to alter the format such that it does not require the use of sensitive personal information at all. For instance, if the written law test is automated (i.e., computerized, instead of the paper-and-pencil

format currently used), it may be possible to add a simple short-term memory recall exercise component using a randomly generated number.

There were 18 respondents who mentioned the Perceptual Response Test (PRT). This assessment tool involved the customer watching an alternating series of schematic images of cars and trucks flash on a computer screen for variable lengths of time (between 17 and 500 milliseconds). The customer was then asked to use the touch-screen monitor to pick, from two choices, which image had appeared. At a very basic level, this assessment measures for potential signs of dementia-related cognitive limitations (Janke, 2001; Owsley, Ball, Sloane, Roenker, & Bruni, 1991). In a few cases (5), staff cited purely mechanical concerns regarding improperly set volumes (i.e., too low for people to hear). In other cases, staff had a more generalized set of concerns. These appeared to result largely from dealing with customer questions, confusion, or outright fear of taking a computer-based test. In some cases, the comments revealed a misunderstanding—typically on the part of the customer, but sometimes also on the part of the staff—regarding what the test assessed. The most common concern cited was the so-called "fuzzy screen." This appeared after the flashed image but before the customer had to make a choice; this "snow" eliminated any retinal after-image left from the previous flash. Evidently many customers, not having seen the initial flashed image, looked for it in the "snow." They then asked the technician why they couldn't see the schematic car or truck. That the technician then reported this as a *problem with the test* (as opposed to a sign that the customer had a potential limitation in their cognition/perception) suggests that they (the technician) faced questions from customers which they may not have had the time or training to answer. In several cases (3), respondents also cited customer fear of taking a test of *any sort* on a computer. One of these latter respondents then noted that because customers might be unfamiliar with using a computer, they had "a hard time moving fast to respond." This comment is revealing inasmuch as the PRT is not, in fact, a reaction-time test. However, at least some staff, and an unknown number of customers, thought that it was.[8] In general

[8] The PRT primarily measures the speed of processing in the fovea of the retina; thus, if an image flashes "too fast" the person being tested *does not see the flashed image*. This is why the length of time that images are flashed on the screen varies, from a minimum of 17 milliseconds to a maximum of half a second, and a person's "score" on the test reflects the minimum time at which they can reliably see a flashed image. The amount of time that it takes to *decide* what one has seen is irrelevant, except insofar as the program was preset to "time out" if someone took more than 2 minutes to make a choice.

what these comments have in common is a concern with how to provide good customer service by responding to customer questions, confusion, and anxiety.

In the original design of the 3-Tier Pilot, this concern was anticipated, and measures taken to specialize the administration of the PRT. This included training particular staff in its use and, consequently, designating the 3-Tier Manager Is and their backup Administrative Managers as the appropriate persons to oversee customers taking the test. This pilot design element came up in one of the substantive suggestions:

> "Designate certain employees to administer the PRT…[who] know your customer[s,] especially the ones who are already nervous about using computers. Need an employee who can communicate effectively yet have patience and understanding."
>
> 3-Tier Manager I

In brief, this manager suggested that the PRT, if it proves useful as a driver competency assessment tool, ought to be administered in a manner similar to the way in which the pilot was originally designed: overseen by one or more specifically designated staff people, who can answer questions and respond to customer concerns from a somewhat more specialized perspective. As revealed by the survey, however, implementation played out somewhat differently in the various pilot offices. Of those who raised concerns about the PRT approximately half (8/17) were MVFRs. While these staff may have received on-the-job training in how to administer the PRT, none of them were included in the formal training sessions on its use. That said, two-thirds of the 3-Tier Manager Is—all of whom received formal training—also raised concerns about the PRT. The concerns raised by the 3-Tier Manager Is were collectively quite similar to those raised by respondents from other job categories. At the very least, the critiques raised regarding the PRT indicate how the implementation of an unfamiliar assessment tool in an agency setting may produce customer anxiety. This is turn may lead to frustration on the part of staff in managing and responding to customer questions and concerns. This anxiety and frustration may arise specifically with assessment tools that are unlike others with which staff and customers are familiar.

There were also a number of unique substantive suggestions regarding other elements of the 3-Tier process. These included (a) providing a designated area, or even a separate line, for customers with physical disabilities; (b) having to wait some period of time

after a written test failure before taking it a second time (respondent suggested a week), (c) automating/computerizing the written test as well as the Tier 1 Score Sheet, (d) the disposition of duties between Hearing Officers and field office staff,[9] and (e) having two different observers rate each customer on potential physical limitations (specifically, the technician at the Start Here window, and the technician who actually processes the customer's application).

Views on 3-Tier's Potential Impact on Customer Service

The survey question on customer service had two parts: an ordinal Likert-type scale ("very positive" through "neutral" to "very negative" impact) and a follow-up open-ended "comments" section. For this and for subsequent ordinal-type questions, the answer was converted to a number, depending on the choices available (one through five, or very positive through very negative, in this case). This allowed the calculation of means, modes, and cross-tabulations by office and job category. In these tables, the categories "very negative" and "negative" were combined, to protect respondent anonymity.

Generally, most respondents reported a positive or neutral impact on customer service (see Table M1.4A). Of those answering this question, less than a fifth (23/126) reported either a negative or very negative impact. By contrast, approximately 30% (39/126) reported a positive or very positive impact, while the remainder (64/126, or about half) reported a neutral impact. A substantial number of those reporting a neutral impact added in their comments that this really meant a "mixed" set of effects, with both positive and negative elements (to wit: increased individual attention to customers versus increased wait times, respectively). The variation across offices in views on 3-Tier's potential impact on customer service ranged within a fairly narrow band: from a converted mean of 2.5 (between positive and neutral) at the Fairfield office to 3.1 (approximately neutral) at the Sacramento-Broadway office. There appears to be a possible mild trend toward a more positive view at the smaller versus the larger field offices.

[9] Specifically, the suggestions made here had to do with (a) the method by which Driver Safety Office-mandated drive tests were scheduled (the preferred method being online, as opposed to by phone, as was done in the pilot; two comments), and with who had the authority to schedule a second drive test for a Driver Safety Office-referred customer in case of a first-time failure. The latter suggestions (two of them, both originating from DSO), favored reserving this authority to the Hearing Officer.

Table M1.4A: Staff and Management Views of 3-Tier's Impact on Customer Service, by Office and Office Size

Office	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a]				Converted Mean score (*SD*)
	Very positive (1)	Positive (2)	Neutral (3)	Negative (4) or very negative (5)	
Carmichael	3	**9**	**9**	8	2.76 (.99)
Fairfield	2	**3**	**3**	2	2.50 (1.08)
Folsom	0	**8**	6	3	2.71 (.77)
Sacramento – Broadway	0	3	**21**	6	3.10 (.55)
Sacramento – South	1	2	**9**	2	2.93 (.92)
Vacaville	1	5	**7**	1	2.64 (.93)
DSO	0	1	**8**	1	3.00 (.45)
Larger field offices[b]	4	14	**39**	16	2.93 (.82)
Smaller field offices[c]	3	16	**17**	6	2.71 (.83)
Total[d]	8	31	64	23	2.83 (.84)

[a] Question wording: "In your experience, what impact has this pilot had on customer service? Very positive, positive, neutral, negative, or very negative?"
[b] Larger field offices included Carmichael, Sacramento-Broadway, and South Sacramento.
[c] Smaller field offices include Fairfield, Folsom, and Vacaville. Neither category includes Driver Safety Office.
[d] Total includes one respondent who could not be located with a particular office; this person reported a "very positive" impact on customer service.

There was somewhat more variation across job classification (see Table M1.4B): from a converted mean of 2.30 for LREs to a high of 3.50 for Managers (other). It is worth noting as well that the number of LREs who reported a positive impact (6) almost equaled the number who reported a neutral impact (7). Moreover, about a quarter of the LREs (5/21) reported a very positive impact on customer service. The other job category reporting a largely positive impact was the 3-Tier Manager I, with a converted mean of 2.50. Together, these two job categories had responsibility for the most direct, extensive contact with customers. This was especially true for customers in Tiers 2 and 3 of the process. To the degree that the more advanced stages of assessment required more personalized interaction, this appears to have reflected positively in these respondents' assessment of 3-Tier's potential impact on customer service.

Table M1.4B: Staff and Management Views of 3-Tier's Impact on Customer Service, by Job Category

Job category	Answer counts, with modal answers in **bold** (converted numeric value in parentheses)[a]				Converted mean score *(SD)*
	Very positive (1)	Positive (2)	Neutral (3)	Negative (4) or very negative (5)	
MVFR/SMVT	3	17	**42**	12	2.85 (0.73)
LRE	5	6	**7**	2	2.30 (0.98)
Hearing Officer	0	1	**8**	1	3.00 (0.47)
3-Tier Manager I	0	**4**	1	1	2.50 (0.84)
Manager (other)	0	2	5	**7**	3.50 (0.94)
Total[b]	8	31	64	23	2.83 (0.84)

[a] Question wording: see table M1.4A
[b] Total includes 2 other/decline to state.

By contrast, the modal view of the 3-Tier process among other types of Managers (Office and Administrative) was negative. Indeed, the only two respondents in the sample who reported a "very negative" impact on customer service both fell into this job classification. Managers, who have ultimate responsibility for keeping down wait times in their offices, appear to have been more likely to equate any potential increase in wait times (such as might occur with 3-Tier) with a decline in the quality of customer service.

Both of these themes—customer service as personalized attention versus customer service as potentially increased wait time—appeared in the qualitative comments on this particular question, for which 65 respondents provided answers. A number of respondents noted explicitly that they thought the process had a "mixed" (positive and negative) impact. Because many respondents said more than one thing, the schematic summary (see Table M1.5) reports number of comments rather than respondents.

By far the most common theme (brought up by 20 respondents) had to do with 3-Tier's perceived impact on increasing wait times in the field offices. This wait time was attributed to the lengthier per-customer processing time associated with 3-tier

Table M1.5: Staff and Management Views on Potential Impact of 3-Tier on Customer Service

Qualitative comment *(N)*[a]	Modal answer (quantitative) [b] *(N)*	Converted mean (*SD*)	Examples of comments[c]
Time (20)	Neutral (8); negative (7)	3.40 (0.88)	"Negative in the sense of wait time." "Once the customer was told about how and why they were okay with it. Region needs to allow more time with customers."
Discrimination (12)	Neutral (6); negative (6)	3.50 (0.52)	Seniors (10): "Some have complained that it is another tool to get our elderly drivers off the road." "They think that it's targeting old people no matter what we tell them." General (2): "Others felt singled out because they noted not all drivers participated in 3-Tier." "I feel that some customers felt as if they were picked on or discriminated against."
Miscellaneous (6)	Neutral (3)	3.17 (0.75)	"I did not tell them they were different from anyone else." "No one ever opted to not participate."
Mixed (12)	Neutral (9)	2.92 (0.52)	"It was very mixed; some were very negative and others very positive." "Some customers felt this was a positive procedure, others felt they were singled out."
Test fear or anxiety (11)	Neutral (6)	3.09 (.070)	"They got scared; especially because of the drive test and sometimes because of vision test."
Personal attention (6)	Positive (4)	2.00 (0.63)	"A greater degree of personal attention." "A lot of customers thought the fog chart was good once [it was] explained to them what it was all about."
Traffic safety (9)	Positive (6)	1.89 (0.60)	"Most customers realize DMV is trying to improve driver safety for all drivers and screening of some drivers and limiting drivers of concern is absolutely necessary." "When people understand what we are doing and preventing, they are happy to participate and pleased that we are taking these steps to keep roadways safe."

[a] Total *N* of respondents: 65

[b] These are overlapping codes; some responses were coded as belonging in more than one category.

[c] Question wording: "In your experience, what impact has this pilot had on customer service?"

procedures. In most cases it was taken as given that any increase in wait times was equivalent to, and synonymous with, poor customer service. As one 3-Tier Manager I put it, "First, the overall wait times for all customers went up due to the lengthy processing time for each 3-Tier renewal. But that negative impact I feel was outweighed…" This was echoed by an MVFR, who stated: "We are 'on the clock': when a 5 min. transaction takes 16-20 min, it destroys my performance." Respondents who saw 3-Tier's impact on customer service in terms of time tended to see that impact as being negative. This was especially true among managers: of those who answered this question (8), half raised the issue of 3-Tier's impact on wait times. All managers who raised the issue of wait times viewed 3-Tier's impact on customer service as being either negative or very negative.

A substantial number of respondents (12) cited an impact (largely negative) on customer service that they equated with discrimination of one sort or another. About half (7) linked this concern regarding discrimination with the treatment of senior citizens or the elderly. In some of these cases the respondent cited customer complaints in this area; in other cases the respondent reported that they themselves thought the process discriminated against senior citizens somehow. The remaining 5 mentioned discrimination, but did not specify against whom.[10]

An additional 11 respondents cited what they saw as an increase in customer anxiety or fear. Typically this was linked to "additional testing" (e.g., taking a drive test, but also the contrast sensitivity chart). However, those who cited an increase in customer anxiety or fear did not necessarily see this in entirely negative terms; in their response to the closed-ended half of the question, many cited that 3-Tier's impact was "neutral." This perception of a neutral (or mixed) impact was echoed by many in the respondent pool. Of those surveyed, 12 noted that customer response was varied: "some were very negative and others very positive" (an MVFR).

Two themes came up often among those who cited a positive impact on customer service: personal attention, and improved traffic safety. To the degree that respondents saw 3-Tier's impact on customer service in either of these two frames, they saw the impact as producing positive customer response. This appears to be particularly true in

[10] In one of these cases, the respondent cited elsewhere in the survey a concern over the fact that 3-Tier included only those customers taking the written law test in English.

the case of those respondents who noted that positive customer service outcomes depended to some extent on their own ability (or the time available) to explain what the program was about. Hence, as noted by a 3-Tier Manager I "A lot of customers thought the fog chart was good *once [it was] explained to them what it was all about*" (emphasis by the author). Or, as noted by an MVFR, "Being able to spend a little more with customers one-on-one was a good thing — helping them to feel more at ease." In both cases, improvements to traffic safety, and reduction of customer anxiety depended, at least in part, on communication between staff and the customer. This may be the logical flip-side, of course, of the previously-cited comments regarding how the 3-Tier process took longer than equivalent transactions: Personalized communication requires both extra time and extra training in how to answer what may be complicated questions.

Staff Reports on Customer Feedback on 3-Tier

Closely related to staff views on customer service were staff reports of what customers had to say about the pilot. This is, of course, different from customers' views on 3-Tier (see Module #3 of this Appendix) which were analyzed with a separate survey. In many cases the responses repeated the answers provided to the previous question (on customer service), though there were also instances of respondents answering one question but not the other. The results reported here should be taken as, at best, a rough and indirect indicator of what customers in general thought about 3-Tier. However, they also provide a window onto the kinds of concerns and questions staff had to address on a daily basis. Thus these data illustrate the kinds of problems that are likely to arise when providing good customer service in a public agency setting. See Table M1.6 for a schematic summary of the open-ended responses to this question. As with the previous table, the coding scheme here is overlapping; a few comments received more than one code, and so the counts refer to responses, not respondents.

In a number of cases the respondent reported either negative or positive feedback which they linked to a specific concern or praise regarding the program. However, in other cases (12), negative and positive feedback on the program was left unspecified. That said, approximately a quarter (20/72 respondents) reported negative customer

Table M1.6: Staff and Management Reports of Customer Feedback on 3-Tier

Type of comment *(N)*[a,b]	Examples of comments[c]
Discrimination (18)	Seniors: "The older customers said it was discrimination." "The elderly customers all asked if we were targeting them." General: "Some felt as if they were a target." "They think that it's a target tactic."
Questions (8)	"About the survey — and the referral to the ophthalmologist." "Mostly questions about it."
Testing (22)	Neutral: "Most did not know about the program but several asked about the fog chart." Negative: "Most were upset regarding tests, both visual and written." Positive: "Most comments were positive. There were a lot of 'thanks for taking these steps' and 'that fog chart is great' comments."
Traffic safety (6)	"Others say [it] is good because it will keep roads safe." "Many drivers appreciated suggestions to improve their driving and the effort made by DMV (time and travel) to give them an opportunity to have limited driving."
Time (4)	"The whole process took too long for their busy schedules." "The customers appreciated the 'extra time' we spent helping to get their license."
Negative (20)	Discrimination: "Negative feedback. Customers felt singled out…" Testing: "All the people who failed the fog chart, written test, or PRT were against 3-Tier." Seniors: "Mostly elderly customers were asking why we have to this, it seemed like they were not too happy about it."
Positive (11)	Traffic Safety: "Some good, stating 'good trying to make our drivers are safe on the road'."
Miscellaneous/ unspecified (20)	"No feedback at all. Respondents were guarded with responses to survey." "Have been told this is 'stupid' to this is a 'great' tool." "Some customers though it was a good idea."

[a] Total # of respondents: 72
[b] This table reflects the use of overlapping codes.
[c] Question wording: "Did you receive feedback (positive or negative) from customers regarding the 3-Tier Pilot?"

feedback of some kind. Of these, 10 stated that it came from customers who had to take additional tests, or were upset about having failed to pass certain assessments. Among those reporting negative feedback, 5 stated that it came from seniors. This was usually linked to either perceived discrimination or anxiety regarding the various assessment tests; this code overlaps to some extent with the previous one. An additional 8 respondents reported unspecified negative comments. Of those reporting positive feedback, 8 out of 11 reported good customer reactions to the testing procedures

(especially the contrast sensitivity test, but also to the more general concept of improving traffic safety).

Noteworthy for their relative rarity were comments about the wait time that 3-Tier may have added; only 4 respondents noted that customers gave feedback about wait times, and in one of those cases the feedback was positive (i.e., appreciation for more personalized attention).

In 18 cases, the respondent reported customer complaints regarding discrimination, targeting, or being "singled out" and "picked on." In 12 of these, the respondent stated explicitly that this had to do with discrimination against seniors, while in the other 6 cases the population facing discrimination remained unspecified. In general there appears to be no relationship between a respondent's reports of customer complaints on this issue, and their own views on the impact on customer service. An additional 3 respondents reported questions or feedback from seniors; while they did not cite complaints of discrimination, they did report negative feedback, as in "Mostly elderly customers were asking why we have to this, it seemed like they were not too happy about it" (an MVFR).

In 22 cases, the respondent reported customer questions regarding the testing, and more specifically with the justification for the project as a whole (e.g., as reported by one SMVT, "Some customers asked why they had to go through all that process."), or with specific elements (e.g., the Perceptual Response Test [PRT], the Pelli-Robson chart). To the degree that respondents reported customer questions regarding the PRT, those comments tended to be negative. On the other hand, to the degree that respondents reported customer questions regarding the Pelli-Robson contrast sensitivity chart, those comments tended to be neutral or positive. Comments on the various testing elements were approximately evenly split between reported negative customer reactions (8 responses), and neutral or mixed reactions (9 responses).

Staff Views on the Fairness of 3-Tier

In a format similar to the phrasing used for customer service, the survey used a two-part question about staff views on the fairness of the 3-Tier process. The ordinal forced-choice portion had a range from one ("very fair") to four ("not very fair"), after which the respondent was invited to give qualitative comments. See Tables M1.7A and M1.7B

for a summary of findings on the quantitative portion of this question by office, office size, and by job category. Because very few respondents replied that the process was "not very fair" these responses were combined with the next category ("somewhat fair") to protect anonymity.

Table M1.7A: Staff and Management Views of the Fairness of the 3-Tier Process, by Office and Office Size

Office	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a]			Converted mean score (*SD*)
	Very fair (1)	Fairly fair (2)	Somewhat fair (3) or not very fair (4)	
Carmichael	7	**11**	**11**	2.21 (.92)
Fairfield	3	2	**4**	2.44 (1.33)
Folsom	4	**7**	5	2.06 (0.77)
Sacramento – Broadway	9	**10**	**10**	2.07 (0.88)
Sacramento – South	4	**8**	4	2.13 (0.96)
Vacaville	1	**8**	7	2.44 (0.73)
DSO	**4**	3	2	1.89 (1.05)
Large field offices[b]	20	**29**	21	2.12 (0.88)
Smaller field offices[b]	8	**17**	16	2.29 (0.90)
Total[c]	32	49	44	2.18 (0.91)

[a] Question wording: "How fair do you think the 3-Tier process was? Very fair, fairly fair, somewhat fair, or not very fair?"
[b] For definition of larger and smaller field offices, see Table M1.4A
[c] Total includes one respondent who could not be located with a particular office — this person reported that they thought the 3-Tier process was "somewhat fair."

There does appear to be a substantial degree of concern among staff regarding the fairness of the 3-Tier process. The modal answer overall, and for most offices, is "fairly fair." However, more respondents believe the process is "somewhat fair" or "not very fair" than believe that it is "very fair." There is substantial variation across job categories: while LREs and Hearing Officers are most likely to regard the process as "very fair," MVFRs and SMVTs typically regard it as "fairly fair," while managers of any stripe (including 3-Tier Manager Is) are most likely to regard it as either "somewhat fair" or "not very fair." This variation does not appear to be driven by either self-reported customer load or whether or not a respondent received formal training for the 3-Tier Pilot (cross-tabulations not shown, available upon request from the author).

Table M1.7B: Staff and Management Views of the Fairness of the 3-Tier Process, by Job Category

Job category	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a]			Converted mean score (SD)
	Very fair (1)	Fairly fair (2)	Somewhat fair (3) or not very fair (4)	
MVFR/SMVT	13	**35**	27	2.25 (0.82)
LRE	**9**	5	7	1.95 (0.97)
Hearing Officer	**4**	3	1	1.75 (1.04)
3-Tier Manager I	1	2	**3**	2.50 (1.05)
Manager (other)	4	3	**6**	2.23 (1.01)
Total[b]	32	49	44	2.17 (0.90)

[a] Question wording: see Table M1.7A

[b] Total includes two other/decline to state.

Those with the most direct responsibility for making licensing decisions (LREs and Hearing Officers) appear to have been most likely to regard the process as "very fair." By contrast, those most removed from the processes of collecting assessment information and using that information to make licensing decisions (i.e., Managers) were most likely to view the process as unfair in some respect. One potential explanation for these patterns is that those with the most day-to-day experience in making decisions regarding fitness to drive safely are both (a) most familiar with the training protocols which undergird these decisions, and (b) most comfortable with the responsibility for making such decisions. However, given the nature of the data in this survey, any speculation on this pattern must remain tentative.

As a follow-up, respondents were invited to explain their answers, especially if they had concerns about the fairness of the program. Given the question wording, the answers were (perhaps not surprisingly) skewed towards those who thought that 3-Tier was unfair in some respect or other. There were 52 respondents who gave an open-ended comment; this was the second-lowest response rate (40%) of any of the open-ended questions.[11] Concerns over the fairness of the 3-Tier Pilot came in three categories: language, discrimination, and testing elements (ten responses were assigned

[11] It is probably best not to read too much into the low response rate to this question, which may flow from a number of causes: relative satisfaction with the fairness of the pilot, question difficulty, or even question fatigue.

to a rump category of "misc./none"). See Table M1.8 for a schematic summary. Since all responses contained a single theme, the *N* of responses equals the *N* of respondents in this table.

There were 11 respondents who raised the issue of 3-Tier Pilot enrolling only those customers who took their renewal test in English. As stated earlier, in the event of statewide implementation of the 3-Tier process (or any of its constituent elements), it will presumably be made available in any of the languages for which DMV provides educational and testing materials. Of the 9 respondents who reported that the process was "not very fair," 4 raised concerns about it being available in all languages (3 raised concerns about discrimination, and 2 did not answer the qualitative follow-up). An additional 14 respondents raised a concern about perceived discrimination in the process; of these, 11 linked their concerns explicitly to treatment of seniors (3 also mentioned those with physical disabilities as well as seniors). In combination, this suggests that staff concerns regarding any potential unfairness to the 3-Tier Pilot can be addressed in a relatively straightforward manner, as they are largely limited to these two issues (language, and differential impact on senior citizens).

There were 17 respondents who critiqued various assessment tests of the 3-Tier process. These were spread out over all elements. However, the bulk of comments were directed toward the contrast sensitivity charts and the memory recall exercise. In many cases the concerns raised here paralleled statements made by respondents elsewhere in the survey (i.e., in answer to question #4, regarding suggested revisions to the process). For instance, 7 respondents raised here their concerns that differential lighting on the Pelli-Robson contrast sensitivity charts was unfair to some customers. There were 4 respondents who mentioned the memory recall exercise; here the comments tended to reveal some confusion regarding the purpose of the test, which may have been related to the perceived unfairness of this assessment tool.

As a more general matter, the concerns raised here point up two ways in which staff saw "fairness" (or universality) as an issue. The first has to do with how customers view the process as they experience it. At least according to some of those surveyed, some (unknown) number of customers told DMV staff that they thought the 3-Tier process was unfair. This appears to have been largely confined to senior citizens. To the degree that customers observe the processing and treatment of other patrons in comparison to

their own experience, they may be sensitive to any perceived differences in what is required of any given customer. This may partly explain why the PRT—as opposed to the contrast sensitivity chart—became a focus of customer questions and complaints. While all 3-Tier customers had to take the contrast sensitivity test (which was part of Tier 1), a much smaller number had to take the PRT (which was part of Tier 2).

Table M1.8: Staff and Management Concerns Regarding the Fairness of the 3-Tier Process

Qualitative comment *(N)*[a]	Modal answer (quantitative) *(N)*	Examples of comments[b]
Language (11)	Somewhat fair (4) not very fair (4)	"Process should cover all languages, out of state original applicants." "All drivers should have been included, not just English speaking."
Discrimination (15)	Somewhat fair (6)	"Screen every customer so that everyone knows they have been screened for 3-Tier, then no chance of feeling elderly being targeted." "Maybe come up with a better response to give to the elderly, because they felt they were being targeted." "I think it targets the elderly."
Testing (16)	Somewhat fair (9)	Memory Recall (4 comments): "I don't understand the purpose of the memory test." "The main problem I saw as far as fairness was the customer having to memorize their SSN. If you never had to memorize [it] in 40 to 70 years, why should they be penalized[?]" Contrast sensitivity (7 comments): "The design of offices needs to be taken into consideration. Large windows cause extra glare on fog chart." "Each fog chart had different ambient lighting characteristics." Educational Intervention: "It was noticed that we did not give the same education info to all customers." PRT: "…the PRT, a computer test…is rather frightening to most of the elderly. They are afraid and not sure what to do."
Misc./none (10)	Fairly fair (5)	"FYI: I am so happy because the people who did not too good are the ones that really need to stop driving." "This was a thought out project by someone." "None." "Not sure."

[a] Total *N* of respondents: 52

[b] Question wording: see Table M1.7A

From the staff perspective, however, there may exist some variation in understanding the nature and purpose of assessment testing. Although only a few respondents bluntly said, as did one SMVT, "I think it targets the elderly," it appears that this sentiment was shared by a not-insubstantial minority of staff. This issue of "targeting," or discrimination—especially as regards senior citizens—appears to be linked to two assessment tools in particular: the memory recall exercise, and the PRT. Both of these are designed to assess, at a basic level, two potential symptoms of dementia: short-term memory loss, and perceptual speed. Inasmuch as risk of dementia is correlated with age, and to the degree that these assessments were administered properly, it should not be surprising that those customers who were flagged by these assessments for further testing were disproportionately senior citizens. That some staff then perceived this as "discriminatory"—rather than the assessment tools accurately working to flag those with potential cognitive limitations—suggests a potential gap in understanding the evidentiary basis for traffic safety screening tests. This may be something that can be addressed in training, with an additional focus on the rationale for, and traffic safety implications of, various assessment tools.

Staff Feedback on Training

There were three questions related to training: a general query regarding its usefulness, a follow-up about the speed with which staff and managers became familiar with the 3-Tier process (i.e., the "learning curve"), and a third question regarding the number and type of questions that came up in the post-training period of the pilot. See Tables M1.9A and M1.9B for the tabulated results of answers to the question regarding usefulness of training. Note: approximately 10% of the surveyed population did not attend formal training for 3-Tier — their answers are excluded from Tables M1.9A and M1.9B.

Because of anonymity concerns, the original four categories available for this question were collapsed to two: "extremely" or "very useful", and "somewhat" or "of limited usefulness." This collapsing of categories produces the appearance of more variation than exists in the data: only 11 people each answered that the training was either "very useful" or "of limited usefulness." In other words, the overwhelming majority of respondents chose the middle two categories. This resulted in relatively little variation

Table M1.9A: Staff and Management Views of the Usefulness of Training, by Office and Office Size

Office	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a]		Converted mean score (*SD*)
	Extremely (1) or very useful (2)	Somewhat (3) or of limited usefulness (4)	
Carmichael	12	**13**	2.56 (0.82)
Fairfield	**6**	4	2.20 (0.79)
Folsom	**11**	4	2.13 (0.64)
Sacramento – Broadway	8	**15**	2.70 (0.70)
Sacramento – South	**7**	6	2.38 (0.87)
Vacaville	6	**7**	2.46 (0.88)
DSO	0	**9**	3.44 (0.53)
Large field offices [b]	27	**34**	2.57 (0.78)
Smaller field offices [b]	23	**24**	2.49 (0.86)
Total[c]	54	61	2.53 (0.82)

[a] Question wording: "How useful did you find the formal training, knowing what you do now about the process? Extremely useful, very useful, somewhat useful, or of limited usefulness?"
[b] For definition of larger and smaller field offices, see Table M1.4A
[c] Total includes one person that could not be located with an office. They reported training to be "very useful."

Table M1.9B: Staff and Management Views of the Usefulness of Training, by Job Category

Job category	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a]		Converted mean score (*SD*)
	Extremely (1) or very useful (2)	Somewhat (3) or of limited usefulness (4)	
MVFR/SMVT	**32**	31	2.46 (0.75)
LRE	**9**	7	2.13 (0.83)
Hearing Officer	0	**8**	3.50 (0.54)
3-Tier Manager I	3	**4**	2.71 (0.76)
Manager (other)	**6**	4	2.50 (0.97)
Total[b]	51	58	2.53 (0.82)

[a] Question wording: see Table M1.9A.
[b] Total includes 2 Other/Decline to State, one of whom reported that training was "very useful" and one of whom reported that it was "somewhat useful."

across offices or office size. However, it is clear that the Hearing Officers in the Driver Safety Branch had a markedly more negative assessment of the usefulness of training.

No Hearing Officer viewed training as "extremely" or "very useful." In part this may reflect the fact that the staff in that office were trained first, before any other office. As the training protocols were later changed, some of the information that had been given to the Hearing Officers became obsolete. This led to some confusion during implementation of the project.

Table M1.10: Staff and Management Concerns Regarding the Usefulness of Training

Type of comment	Number of responses[a]	Examples of comments[b]
Confusion/ changed	23	"The original instructions were vague and got changed as implementation progressed." "It was useful but needed to be a little more set before the training." "Too many unanswered questions during training class."
Good	5	"A well put together class." "Trainers were great." "It made everything pretty clear and let us know what to expect."
On-the-job training and/or role-playing	11/16[c]	"Nothing like on the job experience." "The training was great for questions but you never learn everything in training. Doing the work in the field is the best training." "I learned as I went along with help from everyone in the office" "The training explained the tracking and score sheets. But not much on the actual field training on how to handle field problems or scenarios."
Misc.	2	"I was in the first class." "Received formal training for [my job category]."

[a] Total *N* of respondents: 48
[b] Question wording: see Table M1.9A
[c] Note: This includes five respondents who did not attend formal training. See text.

This particular critique was made explicit in the comments (see Table M1.10) given by respondents. Of the 48 staff who gave comments, nearly half (22) remarked that training was confusing, changed halfway through, or that the information communicated by the trainers was inconsistent.[12] Given the feedback in this module and in other modules of

[12] As with the previous section, it is best not to read too much into the low response rate (37%) to this question, which may result merely from question fatigue.

the 3-Tier Process Report, these concerns will be addressed in the event of statewide implementation.

A substantial number of respondents (16) remarked on the necessity of on-the-job experience; approximately a third of these (5) had not attended formal training. Of these respondents, only one explicitly suggested doing training during the regular Wednesday morning staff meeting. The rest noted, in various ways, how on-the-job experience complemented, or (in the case of those who did not attend training) substituted for formal training. Closely related to this, a handful (3) of staff suggested that training incorporate role-playing or "walk-throughs" regarding how to handle customer scenarios. This method is already used by CA DMV Training Branch, and given the nature of staff concerns raised (above) regarding how to handle customer questions and complaints, the suggestion regarding role-playing might well be a straightforward way to cover these areas.

In terms of how quickly staff reported that they became familiar with the 3-Tier process, the learning curve was relatively steep (see Table M1.11A). Due to concerns over anonymity, in this table I have collapsed the original five answer options to three: "comfortable right from the start," "more than a week but less than a month," and "more than a month." These tables control for customer load by removing from the analysis those who rarely, if ever, saw 3-Tier customers.[13] This has the effect of reducing the number of respondents in the third category ("more than a month") and, therefore, of depressing the reported converted mean. That said, the modal answer for all offices was "more than a week but less than a month." No more than one or two people at any given office reported being uncomfortable with the 3-Tier process after 4 weeks of practice. The same was true by job category, as can be seen in Table M1.11B.

It is clear that the number of questions the staff had on the process post-training was quite minimal. Out of a choice of four options (ranging from "several times per day" to "hardly ever") the modal answer was "hardly ever" (see Tables M1.12A and M1.12B).

[13] This also removes all of the Driver Safety Office respondents. Of the 10 DSO staff, 3 reported becoming comfortable with 3-Tier after one week, and 3 reported that it took between one and two months. The converted mean score was 2.90, with an *SD* of 1.20.

Table M1.11A: Staff and Management Self-Reported Speed of Learning 3-Tier Process, by Office and Office Size

Office	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a,b]			Converted mean score (*SD*)
	"Comfortable from start" (1)	> 1 Week but < 1 month (2)	> 1 Month (3); > 2 months (4); or never comfortable (5)	
Carmichael	5	**15**	1	1.86 (0.66)
Fairfield	1	**4**	0	1.80 (0.45)
Folsom	2	**8**	2	2.00 (0.60)
Sacramento – Broadway	6	**12**	1	1.84 (0.90)
Sacramento – South	1	**7**	3	2.18 (0.60)
Vacaville	4	**6**	0	1.60 (0.52)
Larger field offices[c]	12	**34**	5	1.92 (0.74)
Smaller field offices[c]	7	**18**	2	1.81 (0.56)
Total[d]	20	52	7	1.87 (0.69)

[a] Question wording: "Once the pilot was implemented, about how long did it take for you to get used to 3-Tier processes and procedures?"
[b] Reported tabulations control for self-reported customer load (3 or more customers per day).
[c] For definitions of larger and smaller field offices, see Table M1.4A
[d] Total includes one respondent who could not be located with a particular office — this person reported being "comfortable with 3-Tier procedures right from the start."

Table M1.11B: Staff and Management Self-Reported Speed of Learning 3-Tier Process, by Job Category

Job category	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a,b]			Converted mean score (*SD*)
	Comfortable from start (1)	> 1 Week but < 1 month (2)	> 1 Month (3); > 2 months (4); or never comfortable (5)	
MVFR/SMVT	15	**35**	5	1.87 (0.68)
LRE	2	**7**	0	1.78 (0.44)
3-Tier Manager I	1	**3**	1	2.20 (1.10)
Manager (other)	2	**2**	1	1.80 (0.94)
Total[c]	20	52	7	1.87 (0.69)

[a] Question wording: see Table M1.11A
[b] Reported tabulations control for self-reported customer load (3 or more customers per day). For information on Hearing Officers, see footnote 13.
[c] Total includes other/decline to state. This respondent reported that it took them more than a week but less than a month to get used to 3-Tier procedures.

Table M1.12A: Staff and Management Frequency of Questions About 3-Tier, by Office and Office Size

Office	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a,b]			Converted mean score (*SD*)
	"Several times / day" (1); "About once / day" (2)	"About once per week" (3)	"Hardly ever" (4)	
Carmichael	7	**11**	3	2.67 (0.91)
Fairfield	0	**2**	3	3.60 (0.55)
Folsom	0	**7**	6	3.46 (0.52)
Sacramento - Broadway	2	5	**11**	3.39 (0.98)
Sacramento - South	1	**5**	**5**	3.36 (0.67)
Vacaville	1	3	**6**	3.50 (0.71)
Larger field offices[c]	10	**21**	19	3.08 (0.94)
Smaller field offices[b]	1	12	**15**	3.50 (0.58)
Total[d]	11	33	35	3.24 (0.85)

[a] Question wording: "Think about a 'typical' week over the course of the last month or two (so, sometime in September, for example). How often did you find yourself asking someone (a co-worker, a manager, someone from R&D) for advice or help about some aspect of 3-Tier?"

[b] Reported tabulations control for self-reported customer load (3 or more customers per day). For information regarding Driver Safety Office, see footnote 13.

[c] For definitions of larger and smaller field offices, see Table M1.4A

[d] Total includes one respondent who could not be located with a particular office. This person reported that they had questions about 3-Tier "hardly ever."

The original four categories are here collapsed to three, to protect respondent anonymity. The resulting categories represent, ultimately, being able to work independently of post-training guidance (had questions "hardly ever"), reporting a need for minimal post-training guidance (having questions "about once per week"), or reporting a substantial need for some degree of post-training follow-up and refresher training (had questions at least once per day). Although these results control for self-reported customer load, the distribution would not change substantially if they were included (cross-tabulation not shown, available upon request from the author).[14]

[14] The modal answer for DSO was "hardly ever", with 7 of 8 respondents marking that answer. The converted mean was 3.86, with an *SD* of 0.38.

Table M1.12B: Staff and Management Frequency of Questions About 3-Tier, by Job Category

Job category	Answer counts, with modal answer in **bold** (converted numeric value in parentheses)[a,b]			Converted mean score (*SD*)
	"Several times/day" (1); "about once/day" (2)	"About once per week" (3)	"Hardly ever" (4)	
MVFR/SMVT	8	**27**	25	3.23 (0.80)
LRE	0	2	**6**	3.75 (0.46)
3-Tier Manager I	1	1	**3**	3.40 (0.89)
Manager (other)	2	**3**	1	2.50 (1.23)
Total	11	33	35	3.24 (0.85)

[a] Question wording: see Table M1.12A

[b] Reported tabulations control for self-reported customer load (3 or more customers per day). For information regarding Hearing Officers, see footnote 13

Conclusions

The staff and managers of the California DMV operate under at least three bureaucratic directives: ensuring traffic safety for all drivers on the road; efficiently processing applications, renewals, and drive tests in the field offices; efficiently processing hearings in Driver Safety offices; and providing excellent customer service.[15] At the margins, reaching these three goals simultaneously can generate some tension, and it is that goal tension that appears to undergird most of the comments, suggestions, and concerns reported in this survey.

So, for instance, when it came to discussing 3-Tier's impact on customer service, staff and management's concerns stemmed from potentially contradictory demands: efficiency vs. personal service and universality of treatment vs. discriminant assessment. More personal attention for customers (a definitively positive customer service outcome) tends necessarily to entail slower processing time (a potentially quite negative customer service outcome). Conversely, ensuring traffic safety requires the use of discriminant testing to identify driver with potential limitations in their ability to drive safely. This necessarily means that not all customers have the same experience

[15] See the DMV's mission statement (http://www.dmv.ca.gov/about/profile/mission.htm) which clearly emphases the first and second of these goals, or the department's most recent strategic plan (CA DMV, 2009; available at http://www.dmv.ca.gov/pubs/strat_plan-09.pdf), which lays out in some detail all three of these goals, plus a fourth regarding consumer protection from deceptive business practices.

while in the office. Since universal treatment is a core principle of good customer service in a public bureaucracy, this generates tension among staff, who must be "fair" while at the same time identifying potentially hazardous drivers. It may also generate tension among customers, who tend to expect the same treatment as is given others, probably know that DMV's core function is to ensure traffic safety, and certainly value as short an office visit as possible.

These tensions also underlie the suggestions made by staff regarding the paperwork and various process elements. Certainly the suggestions made by staff to reduce and eliminate excess paperwork stems from genuine concerns with keeping customers moving through the lines as quickly as possible. This may, however, mean spending less time with any one customer, which may reduce the amount of individualized attention. Secondarily, increases in processing efficiency may also shorten the time in which staff may make structured physical observations. This may in turn impinge upon the quality of information gathered regarding customers' abilities to drive safely.

The critiques raised regarding the memory recall exercise appear also to have stemmed from the two goals of efficient customer processing and of providing universally good customer service. The memory recall exercise (quite apart from concerns regarding personal security) appears to have been administered by technicians with varying levels of precision.[16] This may have had something to do with the fact that it was relatively easy to "cut corners" on this assessment tool to cut down processing time. Also, inasmuch as this was a tool designed to flag a potential sign of dementia, it probably tended to "catch" those most at risk for this disorder: to wit, the elderly. Thus some staff may have felt uneasy with what they perceived as "differential" treatment, despite the fact that customers of all ages had to undergo this test. In other words, even though all 3-Tier customers, a population that was in fact quite age diverse, and the majority of whom were younger than 65, had to undergo this simple test, those applying the assessment *perceived* the outcome of the assessment to "unfairly" differentiate between seniors and non-seniors.

This particular concern regarding the memory recall exercise on the part of staff can be addressed by potential changes to the 3-Tier process. If this kind of short-term memory

[16] This observation comes not from data reported in this survey, but rather from quality control done in the pilot offices by the author and other R&D staff during the project.

test proves useful as a combinatory element of a larger driver competency assessment system, and assuming automation/computerization of the written law test, it should be possible to incorporate a simple and basic test of memory to the testing interface. A customer might, for instance, be given a randomly generated number to memorize and then shortly thereafter be asked to recall the number and enter it into the computer. Not only would this shorten customer processing time at the windows, it would circumvent staff concerns about administering discriminant assessment tools. Those staff in charge of overseeing the testing center might also be specifically trained in how to answer questions regarding this test.

A similar issue regarding discriminant assessment tools also arose regarding the PRT. Here the tension appears to have stemmed from the marginal tension between two goals. On the one hand, there is the goal of ensuring traffic safety by, among other means, using the 3-Tier system to identify drivers with certain kinds of functional limitations related to their ability to drive safely. To the extent that the PRT is useful as a tool for identifying certain kinds of driving-relevant functional limitations, it will, technically speaking, "discriminate" between those possessing those limitations and those who do not possess those limitations. On the other hand, there is the goal of providing excellent customer service, as defined by universal treatment. But since *not everyone* takes the PRT, staff and customers see the test as "unfairly" being applied only to certain populations. Staff may also have directed their concerns toward the PRT because it was so unlike other, already familiar, assessment tools, such as the Snellen charts used to assess visual acuity. If this assessment tool is adopted for statewide use, familiarity (among both customers and staff) will likely come with time. However, the underlying goal tension associated with the PRT may remain: traffic safety vs. excellent customer service qua universal treatment. To address this dilemma, two potential revisions to the 3-Tier process present themselves. The first alternative would involve requiring all renewing customers take the PRT, perhaps as an additional component to the (hopefully by then computerized) written law test. This would be different from the protocols developed in the pilot project, where only a subset of Tier 2 customers took the PRT. This would, of course, mean a great many more customers using the computers on which the test is given; in the pilot, approximately 20% of 3-Tier customers took the PRT. However, requiring all customers to take this assessment would also mean that the test was seen, by both customers and staff alike, to be given to everyone. The second solution would be to have one or more designated staff members

administer this test to any customers that required it. This was part of the original pilot design, though significantly altered in some field offices during the pilot implementation period. Specializing the task of giving the PRT, much like the specialization of the drive test, would mean that any questions from customers regarding the test (why it's being given, why it's only given to some customers and not others, what it's actually testing, etc.) could be answered by someone with a greater level of experience and expertise. Both proposals implicitly assume, of course, that each field office would have enough properly trained personnel available to provide coverage in case of illnesses, vacations, lunches, break times, etc.

As regards training more generally, it is worth emphasizing that the preparation for 3-Tier appears to have provided staff and managers with the tools and background necessary to implement the project in a relatively short period of time. Although there were a number of critiques raised, these critiques were directed largely at the exigencies of a brand-new, first-time project. That said, the staff suggestion to have "walk-throughs" or role-playing exercises would likely prove a useful addition (or rather emphasis, as this technique is already used by Training Branch) especially as regards the kinds of questions that may arise from customers. In addition, given staff concerns about the utility of the memory recall exercise and the PRT, it may be necessary to develop additional materials that address specifically the link between dementia and traffic safety. Clearly, to the degree that 3-Tier components are explicitly and obviously tied to improvements in traffic safety, this increases the confidence of both staff and (by report) customers that the process is a valuable addition to DMV procedures.

MODULE #2: ANALYSIS OF THE STAFF INTERVIEWS

ANALYSIS OF THE STAFF INTERVIEWS

Introduction

Upon completion of the field office portion of the 3-Tier Pilot, the Research and Development Branch (R&D) of California's Department of Motor Vehicles (CA DMV) interviewed a stratified non-random sample of staff and managers who participated in the project. The collection and analysis of the data gathered from these interviews provided a key opportunity to (i) assess staff understanding of the pilot goals, (ii) document variation in the implementation of various 3-Tier process elements, (iii) gather information regarding the sources of deviations from project protocols, and (iv) determine the nature and extent of 3-Tier's potential effect on alterations to office workflows and inter-branch/inter-division cooperation. These findings thus supplement and expand upon the findings presented in Module #1 ("The Staff Survey"). In particular the findings discussed here echo the serious critiques raised in the survey responses regarding the training provided for the pilot. They also flesh out staff and management's views on 3-Tier's impact on customer service, variously defined. As with the surveys, the collection of these interviews (and the subsequent distribution of this report) is intended to improve communication flows between R&D, Field Operation Division (FOD) of CA DMV, and the Driver Safety Branch of Licensing Operations Division (LOD). Included in this report are a discussion of the methods and major findings of this portion of the 3-Tier Process Report.

Staff and managers both reported—in relatively consistent and homogenous language—that 3-Tier was designed in an overall sense to improve traffic safety through the field testing of new driver competency assessment tools. There exists a strong commitment to the achievement of this goal on the part of essentially everyone interviewed. Though only a few respondents articulated this explicitly, there also appears to exist an undercurrent of excitement at being part of an organization dedicated to pro-actively improving traffic safety and reducing the number of crashes and deaths on California's roads. In terms of the component assessment tools of the 3-Tier process, there appears to have been relatively consistent application of the physical observation protocol, the Pelli-Robson contrast sensitivity chart, and the Supplementary Driver Performance Evaluation (SDPE). While very few Area Driver Performance Evaluation (ADPE) drive tests were administered, this appears to have been consistent

with current practices within both Driver Safety and the field offices. There was much less consistency in the implementation of the memory recall exercise, in understanding what the Perceptual Response Test (PRT) actually measured, and in the administration of the quasi-experimental design of the educational intervention videos developed for the Perceptual Response and contrast sensitivity assessment tests. It appears that these deviations from project protocols derived partly from training, the quality of which was roundly criticized by most respondents. However, it also appears that a significant amount of the variation in the application of project protocols stemmed from goal conflict, which is to say from the exigencies of navigating competing demands placed upon (especially) field office front-line staff. These competing goals include the pressure to reduce wait times, the provision of consistent/universal service to all customers, and assuring that drivers are skilled and knowledgeable. These conflicts appear most clearly in discussion of 3-Tier's impact on customer service, a point raised in the analysis of the staff surveys (see Module #1). Somewhat surprisingly, the respondents noted few substantial changes to intra-office workflows or inter-office communication as a result of 3-Tier. The main exception to this point involved a substantial improvement in communication between R&D and the field offices that occurred mid-project, largely as a result of the hiring of liaison staff. The conclusions contain a discussion of the implications of these findings for analysis of the outcome data of the 3-Tier Pilot, the implications for potential statewide implementation of the 3-Tier system, and the implications for any future R&D projects that involve changes to Field Office or Driver Safety procedures.

Method

Sampling Frame

For this study, the author constructed a purposive (non-probability) stratified sample of all DMV staff and managers who worked on the 3-Tier Pilot. Potential respondents were identified first through the staff survey, which included a cover sheet asking for interview volunteers (this sheet was immediately detached to assure anonymity of the survey responses). In order to ensure the participation of multiple respondents from each of the six pilot offices, the author recruited additional participants via nomination by the six Office Managers. The author also recruited key informants who participated in the planning and implementation of the pilot. These key informants came from FOD

Region III, FOD Staff Services (at DMV Headquarters), and the Driver Safety Branch of LOD. The final sample (*n*=49) included:

- 13 Motor Vehicle Field Representatives (MVFRs) and Senior Motor Vehicle Technicians (SMVTs)
- 6 Licensing Registration Examiners (LREs)
- 4 Hearing Officers
- 6 3-Tier Manager Is
- 4 Administrative/Back-up Managers
- The Office Managers of the 6 pilot offices, plus an interim Office Manager who held that responsibility during a medical leave
- 9 Managers working either in the FOD Region III office, the Sacramento office of the Driver Safety Branch, or at DMV Headquarters.

Because the study design incorporated a non-probability sampling frame, no claims can be made that this sample is strictly generalizable to the universe of staff and managers that worked on the pilot. To partly compensate for this, the sample includes multiple representatives from all job categories with direct participation in the pilot. In some cases the sample incorporated all possible representatives of a particular job category: this included Office Managers, 3-Tier Manager Is, and the key informant coordinating personnel working at either FOD Region III or DMV Headquarters.

Analysis Techniques Used

In consultation with other members of R&D, as well as with two managers from FOD Staff Services, the author developed a flexible interview protocol (see Sub-Appendix B). This protocol included some questions common across multiple job categories—such as the query regarding the project's overall purpose—and some questions that were unique to specific positions. Position-specific questions included probes regarding the respondent's understanding and implementation of the various assessment tools. These were typically under the responsibility of one or another job category: for instance, only front-line MVFRs and SMVTs administered the memory recall test, while 3-Tier Managers administered the educational intervention. Each interview lasted between 20 and 60 minutes, and was conducted at the respondent's office, typically in a break

room, a meeting room, or an unused manager's office.[17] Each interview was tape-recorded, then transcribed either by the author or a staff member of the Driver Safety Branch. The vast majority of the interviews were conducted by the author; a small number were conducted by two members of the R&D team, both of whom were trained by the author in interviewing. One respondent declined to be taped and instead submitted answers to the protocol questions in written form. All interviews were analyzed with NVivo (ver. 8) qualitative analysis software. In order to preserve the confidentiality of the respondents, all quotes are identified by the job classification of the author of the quote. Because the managers working at the field office sites and at DMV Headquarters participated in both the planning and implementation stages of the project, this group is referred to collectively as "upper management/Headquarters coordinating personnel." Each job title is also accompanied by a number—this refers to the order in which the interviews were transcribed and does not correspond to the order in which they were conducted.

In developing a coding scheme for the interview transcripts, the author first used simple descriptive codes to group together answers to questions regarding (i) the goals of the 3-Tier Pilot, (ii) each element of the pilot process (memory recall exercise, physical observation, etc.), (iii) training, (iv) intra-office workflows, and (v) inter-branch communication. The author also assigned descriptive codes corresponding to job class of the respondent, the size of the office (if they worked in one of the six pilot field offices), and office location. Secondly, the author developed analytical codes based in part on the results of the staff survey, discussed in Module #1. These included the various definitions of customer service as well as certain related concerns such as perceived impact on wait times, perceived discrimination (based on either age or language), the importance of personalized attention in customer transactions, and potential improvements to traffic safety.

Finally, a second set of analytical codes were inductively developed from the interview material itself. These last codes had to do with disparate views regarding the purpose of driver competency assessment tools. As analysis of the interviews progressed, it became clear that there exist two divergent, though partially overlapping, conceptions

[17] In general the interviews with front-line staff took 20 minutes. Interviews with managers typically took 30-45 minutes. Only a few interviews—all with upper management/Headquarters coordinating personnel—took longer than 45 minutes.

regarding what testing is *for*, which is to say what function it serves. On the one hand, certain staff regard driver competency assessment tools as mechanisms for identifying potential driving-relevant limitations (e.g., vision disorders, physical impairments). This view—which I term "testing for competency"—was applied more commonly in the context of some project components as opposed to other components. Similarly, respondents in particular job categories were somewhat more likely to evince this view than other respondents. The second view—which I term "testing as hurdle"—encompasses a view of driver licensing requirements as a series of bars or hurdles that everyone could—indeed, *should*—overcome. It bears emphasizing that, for many respondents, these views are not mutually exclusive; most of those with whom I spoke used both perspectives at different points during an interview. Indeed, it is the variation in application of these two views (across process element especially) that illuminates a probable source for these disparate paradigms; it appears that, to some degree, "testing as hurdle" was applied to those 3-Tier process elements which staff either did not understand, or for which they were skeptical of the traffic-safety relevance and value. The simple fact that there exist two implicit understandings of the nature of driver assessment has important implications for how staff understand their role in administering these assessment tools, for understanding the variation in implementation of various project components, and ultimately for preparing appropriate training of future cohorts of DMV employees should 3-Tier be implemented statewide.

Limitations of These Data

Qualitative interviews provide a richness of detail that cannot be obtained by standard quantitative survey procedures. This allows for, among other things, the in-depth exploration of research hypotheses beyond simple correlations, and in some cases, the inductive generation of new ideas for understanding the problems at hand. Interviews are also necessarily more expensive to conduct than other, quantitative, techniques—and hence tend to comprise smaller sample sizes. Partly as a result, interview-based research usually incorporates non-probability sampling techniques of various kinds. Both of these constraints upon the method tend to limit claims to generalizability. Interviews are also, by their very nature, irreproducible. This latter point takes on added emphasis in an instance such as the current case; for a number of reasons, the interviews for this project were time sensitive, and had to be conducted as soon as possible after the completion of the pilot. Not only was it likely that individual

respondents' memories would fade regarding key points, but in many cases participants have since moved on to other positions within the DMV hierarchy. Therefore, all interviews were conducted within three months of the final date for the enrollment of new 3-Tier customers through the field offices (10/31/07).

Because respondents volunteered to be interviewed, the results reported here may be subject to various sources of bias. This is most particularly true for front-line staff—respondents who held MVFR, SMVT, or LRE positions—whose participation in the interview process was subject to the consent of their supervisor(s).[18] At the most basic level, it is likely that those front-line staff interviewed for this study had a relatively higher degree of participation in the pilot, compared to other field office employees (as measured by, for instance, the number of 3-Tier customers processed on a daily basis). This probably influenced their knowledge of overall project goals, their understanding of the purpose of specific project components, and perhaps even the ways in which they implemented various tasks. Because the study design could not incorporate a direct "control group" in the specific sense of experimental methods, the magnitude of these potential biases is difficult, if not impossible, to estimate. Respondents may also have been biased in the normative views they expressed towards the project; it is possible, for instance, that only those front-line staff with relatively positive views towards 3-Tier volunteered to be interviewed. As to that, however, all respondents, from whatever job category, likely shaped their responses regarding the normative worth of the project. Respondents were, after all, interviewed by R&D employees—and hence they may have been somewhat likely to mute direct criticism of R&D, since 3-Tier was perceived within DMV as an R&D project. Where appropriate, the data discussed here are compared with the findings presented in Module #1, which included far broader participation by front-line staff. While an imperfect baseline of comparison—the questions on the survey and the interviews are substantially different in content and wording, and the sample populations do not perfectly overlap[19]—such a comparison can illuminate, for instance, where particular pockets of concern about the project may

[18] Note: in only one case did a manager recommend against interviewing a specific volunteer; in the manager's words this person "did not participate in the pilot." In a second case, a manager recommended in favor of interviewing someone who had not originally volunteered. These were the only two cases of explicit censoring of participation in the interview portion of the pilot.

[19] Specifically, many of the respondents categorized here as "Upper Management/Headquarters Coordinating Personnel" were not included in the staff survey.

(roughly) be located. This in turn may provide potential direction for the creation of training protocols, or even revisions to the 3-Tier process, in the event of statewide implementation.

Results

This section describes, first, the variation in respondents' collective understanding of, and commitment to, the goals of the 3-Tier Pilot. Secondly, this section describes the variation in the articulated understanding, and implementation, of each of the elements of the 3-Tier process: the memory recall exercise, the physical observation protocol, the Pelli-Robson contrast sensitivity test, the PRT, the educational intervention materials, and the on-the-road drive tests. Thirdly, the author analyzes the common elements that link variation in understanding and implementation across project elements. These common elements can be grouped together as a series of overlapping organizational goals; it is the tension between these differing goals which can point towards an explanation regarding the sources of deviations from project protocols. The section ends with a discussion of 3-Tier's effects on intra-office workflows and inter-division cooperation.

Project Goals

The overwhelming majority of staff and managers who worked on the 3-Tier Pilot articulated a common view of the project's goals. When asked straightforwardly "in your own words, what is the goal or goals of 3-Tier?" most respondents spoke of the pilot in quite common terms:

> "To take drivers off the road that [can] no longer operate a motor vehicle."
>
> Administrative/Back-up Manager #1
>
> "To evaluate all renewal customers…as to their physical and mental ability to operate a vehicle."
>
> LRE #2
>
> "To try to help find positive ways to…reduce accidents in the state of California."
>
> Administrative/Back-up Manager #3

> "To try to improve our testing policy, [and] help make drivers safe for California roads."
>
> MVFR #5

Across nearly all job categories, 3-Tier's *overall* purpose was clear: to improve traffic safety and to do so by developing tests to better evaluate driver competency and skill. As this was one of the major elements of the training curriculum provided to staff, it is not necessarily surprising to see that reflected in their understanding of the project purpose.

The only real exception to this pattern occurred among Hearing Officers, whose responses to this question tended to be rather vague:

> "To determine whether or not there are better ways for us to handle the renewal process and how to handle customers in the Field Office."
>
> Hearing Officer #3
>
> "A long term program that would provide us with statistics."
>
> Hearing Officer #2
>
> "To assess people that kind of drop off in the middle."
>
> Hearing Officer #4

The vagueness of the answers here—which focused less on improving traffic safety and more on "collecting statistics" and improving the renewal process somehow—likely stem, at least in part, from training. Hearing Officers were among the first to go through training for the pilot, and most of these respondents prefaced their replies to the question regarding project goals by noting the inadequacy of the preparation they had received:

> "I kind of had to learn on my own, reading memos."
>
> Hearing Officer #3

One respondent put it even more explicitly:

> "I don't know, really, whether everything was well thought out or not."
>
> Hearing Officer #2

This criticism of training (which was widespread among all types of respondents, more on which below) appeared nowhere more explicitly than among Hearing Officers and appears to have been a major source of confusion for this class of employees over the pilot's purpose and goals.

Some smaller sub-themes also appeared. These were not as common, but they do illustrate (especially) some of the positive ways in which staff understood the pilot. One theme had to do with driver mobility—or in other words, with: "Trying to keep everyone as mobile as long as we can" (LRE #3) as well as seeing "if we can help [customers] improve their driving and mainly to keep them on the road, not off the road" (3-Tier Manager I #5) and finding "ways that we can help people to stay on the road longer, be able to maintain their license privileges longer, that type of thing" (Hearing Officer #1). The respondents most likely to voice this theme came from job categories specifically charged with counseling and advising drivers, such as LREs, 3-Tier Manager Is, and Hearing Officers.

In a similar manner, a few respondents saw 3-Tier as providing a kind of health-related service to customers, in the identification of early-stage degenerative health disorders. As two staff members phrased it:

> "3-Tier was trying to project any situations that may occur in the future [regarding] physical abilities to drive a car. To check their vision ahead of time, and see if maybe something was coming down to where it could affect their vision. [By] using that fog chart, it could tell us ahead of time if somebody was starting to get some kind of…condition that maybe they need some corrective lenses."
>
> SMVT #1
>
> "I would say that the goal of 3-Tier is to…help bring up problems. Such as like cataracts…I find that we're finding that with the fog chart. Once we point out 'Hey, you're having a problem with the vision. How about you try to go to the doctor?' we see a lot of people coming back, and it has to do with cataracts or other vision impairments."
>
> MVFR #1

This prevention-related function was most clearly tied to the Pelli-Robson contrast sensitivity chart, especially as it led to referrals to ophthalmologists/optometrists.

Among those who identified this as a goal (and only a few respondents did so explicitly), the referral for, and treatment of, potentially dangerous health conditions was a key positive aspect of the project.

Clearly these three ideas are linked: improvements to traffic safety through the identification of unsafe drivers, preserving mobility through the improvement of driver skill, and the identification of potentially progressive disorders. It was also clear that respondents generally regarded these goals in a positive light, primarily on their own merits—i.e., saving lives is already a key mission for the department, extending safe driving years for individuals is seen as a positive good, and providing a flag for previously unrecognized health concerns provides a new kind of service to the customer.

A few respondents—spread across job categories—added a fourth theme. This view emphasized CA DMV as being an "active" or progressive public agency. Here, respondents saw 3-Tier, by virtue of its first-order goals, as achieving a second-order positive good. This secondary goal involved changing both the public's perception of CA DMV as well as the perspectives of the department's employees regarding their jobs. Hence, in the words of one respondent:

> "You know, I was involved a couple of times when the *LA Times* or *The [Sacramento] Bee*, or different reporters were out here, wanting to get a handle on what 3-Tier was all about. And I think one of the real positive aspects of this was that the State of California Department of Motor Vehicles is trying to make a difference in today's day and age."
>
> Upper Management/Headquarters Coordinating Personnel #1

This respondent, who participated in the planning and coordination of the project from its earliest stages right through implementation, saw a distinct advantage to CA DMV in the media coverage of the project. Other respondents in managerial positions—both in field offices and at DMV headquarters—shared this view, but emphasized the potential effects for the morale of employees:

> "I think it gave our employees a real good vision of the Department as looking at education. It's looking at safety. It's looking at each person's individual needs

> and knowing it's their livelihood that you're looking at, whether they're 25 or 90. So I think it really gave our people [the idea] that the Department is looking to be more thorough, and not just slapping people through like cattle."
>
> Upper Management/Headquarters Coordinating Personnel #8

Even employees on the front line saw the project in this light:

> "I mean, I'm trying to look at it honestly and objectively…what I know—if I can say this—[is] with good intention, or a good DMV heart, with the new [DMV Director] George Valverde. And I'm happy and impressed with the direction, and I hope he keeps going to change the image of DMV. It's going to be a tough call to weed all this out and to streamline it, but if it's for the safety of the public, you just have to take a look….If they really want to change the image [of the Department], when it comes down to it, they really need to approve it."
>
> MVFR #5

Or, as stated more simply by another field office employee:

> "I think it would only make our Department look better. [It] would show that we care more, and it would be a positive influence"
>
> LRE #6

According to these respondents 3-Tier thus presents the possibility for improving the public's perception of CA DMV, shifting those views to an image of the department as "making a difference," but with a "good DMV heart." Also, to the degree that a change in the public image of the department is tied to "the safety of the public," staff believed that would increase their own morale and job commitment.

With the notable exception of Hearing Officers, all of those who participated in the pilot clearly articulated a common vision of the pilot's goals: to improve traffic safety, and to do so by finding new methods of assessing driver competency and skill. These goals are already a fundamental part of DMV's core mission (www.dmv.ca.gov/about/profile/mission.htm), and hence, 3-Tier made sense in concept to most field office and headquarters employees. One additional goal—the extension of safe driving years for individual customers—is also clearly a logical extension of departmental goals. Staff also identified two new ideas, however: the flagging of perhaps previously

undiagnosed health problems, and changing "the image of DMV" to "make a difference in today's day and age." These latter two points may constitute untapped selling points for this or other departmental projects.

3-Tier Process Elements: Memory Recall Exercise

VARIATION IN IMPLEMENTATION

The 3-Tier process comprised a series of tiered assessment tools and interventions designed to raise specific flags regarding potential driving-relevant limitations. Among the first-tier assessments that a customer experienced was a memory recall test. This exercise was intended as a very basic flag for cognitive limitations such as dementia; it required a customer to write down, from memory, their social security number (SSN) in two different places (their driver license renewal application, and a separate piece of paper). The staff-person processing the application then compared the two numbers for accuracy with that which the department had on record. If a customer did not possess an SSN, the staff-person was instructed to ask the person for their zip code. Whatever number was used, if a customer could not recall the number accurately (or could not recall it without extensive prompting), this constituted a trigger point for further assessment in Tier 2.

Part of the purpose of the interview protocol was to gather information regarding the degree of variation in the implementation of the various assessments. In this particular case, it appears from the interviews that the majority of respondents adhered to training protocols by asking respondents to write their SSN down on a piece of paper which was later shredded to protect the customer's security. However, it also appears that a few—what percentage cannot really be established from these data—asked customers to *verbally* state their SSN rather than writing it down, per training protocols:

> "The problem is we live in an age where everybody is worried about identity theft and a lot of people that I was testing or saw fit that category that they were over a certain age and they talked louder, and they can't hear as well. So I was uncomfortable when somebody would have to say their Social Security Number so loudly knowing that there are people in the office—not the employees, but people in the office—who may or not be listening and using that information for

> their gain. And that bothered me that there wasn't a better way to do that than having people repeat it. Perhaps it would have been better if you asked the person to write it again in my opinion, because that was uncomfortable."
>
> MVFR #9
>
> "But they would verbally say it to us…a lot of people have a problem with verbally saying their social security number out loud."
>
> MVFR #4

Both of these staff members (who worked in different field offices) reported in the most explicit terms possible that they asked customers to verbally state their social security number. It bears emphasizing that the second respondent's concerns are entirely valid—and already incorporated into standard field office procedures. According to the department's security protocols (which were re-emphasized during pilot training), any documents with sensitive personal information are shredded at the end of each business day. These comments indicate that there existed some degree of variation in how this particular assessment tool was administered. This introduces an element of doubt regarding the consistency—and therefore the utility—of the data produced by this particular test.

In other ways, there exists evidence of substantial variation in implementation of this project component. A number of respondents indicated that they never (or hardly ever) had a customer fail the exercise:

> "Of the applicants I had, I remember only having one that had to look back at their documents to remember what their social security number. That was the only one I got."
>
> MVFR #4
>
> "Oh, the Social Security exercise? I will tell you that nobody failed that."
>
> Administrative/Back-up Manager #6

On the other hand, a number of respondents indicated that at least some segments of the DMV customer base had significant difficulty with this exercise:

> "I'm sure you've heard this before, but some customers really don't remember their social security number. Or they never thought to use it that way."
>
> MVFR #6

> "Most people have their [SSN] memorized but I'd say about 10% or so don't."
> MVFR #11

Both of these latter respondents worked at the same two offices as the respondents cited above (i.e., MVFR #6 and #4 worked almost side-by-side; while MVFR #11 and Administrative Manager #6 worked together). Thus, while it is possible that they saw different customer populations, with consequently different gross propensities to pass this particular assessment, this is unlikely given the randomization inherent in CA DMV's computerized queuing system. Instead, it would seem more likely that staff administered the test with some degree of variation—with some technicians seeing as many as 10% of their 3-Tier customers fail this test, while others had only one failure during the entire pilot period. This variation was described first-hand by one respondent, as reported in the following exchange with an SMVT:

> RESPONDENT: A lot of the technicians, I noticed, would go 'Oh, if you don't know your social write your zip code.' But my interpretation of the policies was, if they don't know it, that's a point against them. Only use the zip code if they don't qualify for the social security number.
> INVESTIGATOR: So, part of your concern was that there was inconsistency between how some technicians implemented it versus others?
> RESPONDENT: Yes.
> INVESTIGATOR: And so the customers weren't getting the same tests?
> RESPONDENT: Right. That's how I felt, yes.

It therefore appears that the implementation of this particular assessment tool varied substantially by technician—and likely produced, in particular, some unknown number of type 2 (false negative) errors. As part of the outcome analysis, it may be possible to document this variation quantitatively—though doing so would not necessarily establish a population baseline.

VARIATION IN UNDERSTANDING

In their discussion of the memory recall test, a great many front-line employees (as well as a number of managerial staff) raised concerns about this exercise. These concerns varied in content and tone, with some (as above) sharing their anxieties over

document/identity security. Others, however, worried that some customers couldn't pass the test, or more generally that some would have a harder time remembering their social security number than others. As several respondents put it,

> "I think we may have assumed that everybody knew their social security number. But I found with a lot of customers—[both] 3-Tier and regular renewals—many don't know their social security number. And that was…I think that may have made it a little bit inaccurate when we were doing the memory assessment. Because I find that a lot of customers don't know their social security number. So when we were basing their memory recall on that, many had to, you know, 'I don't know it. I don't know it.'"
>
> MVFR #1
>
> "I had very few people who didn't know their Social Security Number by heart and in those cases I think it was the fact that they don't know it by heart, not that they couldn't remember it. They never knew it to begin with."
>
> MVFR #12

This sentiment appeared in various guises among approximately half of the MVFRs and SMVTs interviewed for this paper. In a basic way, it appears that these employees—and some unknown number of others—questioned the fundamental validity of the memory recall exercise. In brief, their critiques stem from a view that the memory recall exercise did not constitute a basic flag for cognitive function, but rather a test of long-term habit. Or, to put the matter differently, it appears that many of those interviewed saw the memory recall test as (only) measuring *whether or not somebody knew their social security number by heart*.

Those who raised concerns about this test typically identified one group in particular—senior citizens—as having "trouble" with the memory recall test. More specifically, many of those interviewed viewed it as inappropriately—more specifically, *unfairly*—flagging those who "didn't know by heart" their social security number and who consequently "couldn't remember" when asked to write it down twice. Advanced chronological age is an epidemiological risk indicator for dementia-type disorders. However, the apparent correlation between age and the ability to complete a relatively simple memorization task was not necessarily interpreted by (some) staff as a potential flag for undiagnosed, or early-stage, cognitive limitations. Instead, the cognitive health implication of the memory recall assessment was either misunderstood or disregarded

by many staff. Some respondents were quite explicit in calling into question the face validity of the test. As noted by two managers:

> "But it would be really interesting, I think, if someone can't figure out, you know, 'I don't know my Social [Security Number], I don't know my phone number,' whatever, but can they read Stop? Do they know what that means? Can they read One Way, Do Not Enter? Or Merge/Yield, what that means? So I can see the benefit of using this as a tool to maybe make our tests a little bit more specific."
>
> Upper Management/Headquarters Coordinating Personnel #5
>
> "I'll be interested in those [statistics], if they say, for instance, that the majority of people who could not remember their Social Security are also in this group that had accidents and speeding tickets and stuff. That's fine. That's a great study. What are you going to do with that study?"
>
> Administrative/Back-up Manager #6

Implicitly, many of those working on the project viewed failure on the memory recall test as indicating type 1 (false positive) errors. Thus, whatever this particular assessment may have tested, any link between memory recall and driver competency was regarded as being at best unclear, likely tangential to driver competency—though see Hennessy and Janke (2005) for empirical evidence on this question—and at worst discriminatory against the elderly.

Even when respondents accepted the face validity of the memory recall test as a potential cognitive flag, they still expressed strong reservations against assessing customers for their cognitive health:

> "I only had one person challenge me on their Social Security Number and that was mostly because they wanted to know why I wanted it. And I think it's difficult to say 'Well, I'm really checking your memory' because I'm not a physician. I can't say something like that. That's uncomfortable. I have to say 'We're checking to see if you remember it.'"
>
> MVFR #9

Though rarely stated this explicitly, it was clear that many (perhaps most) respondents simply did not feel that they possessed the training or expertise necessary to assess—even at the most basic level—a customer's cognitive health. Of course, this test (like all of the Tier 1 and Tier 2 assessments) is only used as a flag to indicate the need for further scrutiny by someone with specialized professional training—such as a licensed doctor in the case of vision disorders, or an LRE in the case of driver competency. Indeed, according to the study design the memory recall test never *by itself* triggered an on-road driving test (Tier 3 of the 3-Tier process), but instead only *in combination* with identified physical or visual limitations. Thus, even though this assessment tool had very narrow implications for further assessment, most of those interviewed expressed a reluctance to administer it even when they were not outright skeptical of its utility.

These widely-shared sentiments have a number of implications. First, it suggests that if some version of the memory recall exercise is retained as part of the 3-Tier process, that DMV staff would benefit from a great deal more training in the purpose of this assessment flag, both to enhance understanding and, hopefully, to improve the uniformity of implementation. Secondly, the format of the test could be changed to remove, or reduce, the potential for human error. If, for instance, the written law test is automated by the time of statewide implementation—it is at the time of writing (March, 2010) still administered in paper-and-pencil format—it would be relatively easy to incorporate a simple short-term memory recall test as part of the interface. This would, at the very least, standardize the implementation of this assessment tool. Thirdly, this critique of the memory recall exercise fits with a broader view of the nature and purpose of the license renewal process. Expressed in one form here, this view holds that (a) all assessments should fall equally upon all renewing customers, or in other words (b) if some people fail the assessment, the problem is *with the test*, not with the skill, knowledge, or health of the customer. Thus, if some people fail the memory recall exercise, it must be (according to this view) because the test is "harder" for them than it is for other customers—and not because of variation in individual cognitive functioning. This critique appears in similar form with regards to some of the other elements of the 3-Tier process—particularly the PRT.

3-Tier Process Elements: Physical Observation

The physical observation element of Tier 1 of the 3-Tier process involved a simple checklist, filled out by the front-line employee, of potential driving-relevant physical

limitations they observed of their customer. This list incorporated observation of both the upper and lower body, and was used to identify gross (as opposed to subtle) physical limitations. This included the loss of the use of a limb, obvious shaking or stiffness, and the inability to walk unaided. Depending on how many physical limitations were observed, a customer might be flagged for further assessment in Tiers 2 or 3.

VARIATION IN IMPLEMENTATION

There appears to have been very little variation, according to these data, in the implementation of this assessment tool. Front-line staff uniformly expressed familiarity with the observation protocols. While some respondents voiced frustration with the hassle of paperwork—and more particularly with the amount of time that the physical observation protocol added to the renewal process—by and large those interviewed saw this part of the 3-Tier process as straightforward and procedurally unproblematic.

VARIATION IN UNDERSTANDING

In the judgment of those interviewed, the physical observation protocol had a substantial, significant, and potentially long-term positive impact (even after the pilot ended) on field office procedures. This view tended to come from higher-level employees, but appears across job categories:

> "I think it actually helped the MVFRs do a better job in observing customers and coming to me with questions. Whereas in the past, they didn't question any customer or any problem they had. Well now it kind of opened up their eyes and the observations started getting better and better and better. And I could see that and I could train them on better ways to talk with a customer about their limitations."
>
> Administrative/Back-up Manager #3
>
> "As an examiner, I know from the training from before 3-Tier that techs were supposed to watch customers and be observant. But they weren't. So the good thing about 3-Tier was it forced them to be observant. And that I thought was a real positive element."
>
> LRE #2

> "Oh! I think there's a benefit that our counter personnel received from getting trained in what to look for…the training these individuals had is far and above what everybody else working the counter has. And so these folks will learn—even after 3-Tier—they'll know if [a customer] needs assistance walking to a window… at least to get an examiner, or somebody else, involved in this."
>
> Upper Management/Headquarters Coordinating Personnel #7

Even some MVFRs commented on this, especially as it encouraged them to provide assistance to disabled customers:

> "I thought [the physical observation] worked very well. Actually I think a lot of that has translated over to now: I think we're more aware, and we try to make compensations for customers that have physical limitations that we might not have noticed before. We'll make sure that we're paying attention more. So if they're having trouble standing, we'll get them a chair. I think it's made everybody a little more sensitive to that aspect of it."
>
> MVFR #2

Whether or not the 3-Tier Pilot has this kind of long-term effect on the observational habits of front-line staff is beyond the scope of this analysis. However, it is clear that a substantial number of project participants—especially among management, but also among LREs and even some MVFRs—saw a positive good in finding a mechanism for encouraging front-line employees to observe, assess, and assist customers with potential physical limitations.

Some of the front-line employees did express a subtle tension between having an objective checklist of things to mark, but at the same time feeling that their own non-professional observation was inherently subjective. This came up specifically for temporary disabilities:

> "Typically, it worked well. The only time I would say it wouldn't work well is if there was a temporary limitation, like if someone had a broken arm or broken leg and then they would have to go through the 3-Tier steps…. And it seemed like it should have left a little bit more provisions for if it was a temporary problem or if it was you know something that was debilitating. If you asked them and they said this had been going on for a few years, it would nice to have a little part

where you could write a note for that or something. Or like 'I just broke my arm, it should be out of the cast in like a couple of weeks,' something like that, that would have been nice to maybe not mark them, or maybe, you know, have a half a point?"

MVFR #8

This same tension was also noted by an Administrative/Back-up Manager:

RESPONDENT: I had several people here for awhile that were trying to get very objective. And it's really kind of subjective when somebody limps. I limp, you know, a little bit. My knees get tired at the end of the day. My foot hurts.

INVESTIGATOR: So having the checklist there is good because it does get the [MVFR] to do [the physical observation]. But there's still some subjectivity to it, in that some [MVFR's] just kind of put in zeros and don't even really pay attention?

RESPONDENT: Right…it can go either way with the tech.

INVESTIGATOR: Is there any way to fix that?

RESPONDENT: I don't know if you would fix it. I don't know if you should say "Let's be more definitive", you know?

This tension between objectivity and subjectivity came up even for the permanently disabled:

"It worked well that we would note everything down. But where it got confusing or hard for me is when I had a totally sharp, confident person pass everything. And because they had a stroke, or because maybe their arm was bad and had a cane, or a wheelchair, but they knew current topics, sharper than I was, they still had to be thrown into the 3-Tier [program]. And I'm referring to just a leg or just an arm only, other than that they were right on the ball, they got thrown into it. And it was kind of confusing because they passed everything else. With flying colors. I mean some would see better than me without glasses, they passed the fog, they passed the written, but because they had a cane or a permanently injured arm or a leg or a hand, you had to check off on the sheet…"

MVFR #5

To some extent this kind of unease can be addressed in training. The driver license curriculum for new field office employees could potentially emphasize even more than it does now that observation on the part of DMV leads in the first instance only to additional assessment by a more highly trained employee, such as an LRE—and not to automatic license restrictions or revocation. That said, it is also worth noting that respondents did not voice a parallel unease with other tests—specifically vision—that result in referral to a non-departmental professional health experts. In other words, while (some) respondents voiced unease with flagging customers for further assessment by other DMV employees, none of those interviewed were uncomfortable flagging customers for further assessment by non-departmental health professionals.

3-Tier Process Elements: Snellen Visual Acuity and Pelli-Robson Contrast Sensitivity Tests

During the first tier of the 3-Tier process, customers had to take two vision tests. The first involved CA DMV's standard Snellen chart. If a customer demonstrated, according to this test, a visual acuity of at least 20/40, they were then asked to read from a Pelli-Robson contrast sensitivity chart. Colloquially known as the "fog chart," the Pelli-Robson chart measures contrast sensitivity with letters that fade progressively into the background. The inability to see light letters on a light background can signal the existence of certain kinds of vision disorders more effectively than a visual acuity test. This includes cataracts (Elliott 1998), glaucoma (Hawkins, Szlyk, Ardickas, Alexander, & Wilensky, 2003), and age-related diabetic retinopathy (Stavrou & Wood 2003). As administered during the 3-Tier Pilot, the contrast sensitivity test admitted of three possible outcomes: if a customer demonstrated the ability to read all the letters on lines 1, 4, and 5 of the chart, they passed. If they could not read letters on the first or fourth lines they were referred to a vision specialist for professional examination. If they could not read letters on the fifth line (which were lower contrast, and thus more difficult to discern), they were flagged for assessment in Tier 2.

VARIATION IN IMPLEMENTATION

As with the physical observation, it appears from these data that there was relatively little variation in the administration of these two portions of the 3-Tier process. The implementation of the Snellen visual acuity test required no changes to standard renewal procedures; as a result, this was regarded as unproblematic by all of those interviewed. This was also true of the Pelli-Robson contrast sensitivity test: in response

to the question "how well did it work," front-line employees almost uniformly spoke not about any difficulties administering the test—as appeared when asked about the memory recall exercise—but rather about customer reaction. When asked directly about potential difficulties they may have found in *giving* the test, respondents tended to answer that it was procedurally straightforward:

> "If they were to implement 3-Tier statewide or all DMVs were to do it, I personally feel that the fog chart would be a positive. In other words, that now is just part of the normal routine. You've got the eye chart and the fog chart. Period."
>
> MVFR #5
>
> "I think in reality the fog chart made common sense. I personally think that everybody should take a fog chart."
>
> MVFR #9

Some respondents even noted that it was in fact quite similar to what they were already used to doing with the visual acuity chart:

> "Honestly, it fit right in. Because it was back to back with the visual acuity it was really a smooth transition. You go from one eye chart to the next. Sometimes customer did need to adjust their eyes. But you give them a second, talk to them for a minute and then say 'Okay, we're going to go ahead and do this,' and they would move right along."
>
> MVFR #10

Both the inherent procedural ease and its similarity to the Snellen test evidently made this one of the simpler aspects of the 3-Tier process to implement.

To the degree that respondents noted any variation in implementation, it had to do not with what technicians did, but rather with variation between testing stations in the amount of ambient light, the depth of shadows, or the existence of glare from nearby windows. Six respondents, at five different field offices, raised this concern. Each noted that differential lighting may have flagged some customers for further assessment not because of variation in their visual health, but rather on the basis of the counter window

at which they happened to be standing.[20] This was partly addressed by in-office procedural improvisation—staff mentioned both walking their customers to different charts to get "better lighting," as well as "angling" (or bending) the charts to improve the customer's chances of passing the test. This latter practice is not so different from current procedures with the visual acuity charts, where technicians routinely walk with customers across the office to different counter windows, or encourage them to stand at whatever angle they happen to feel comfortable viewing the Snellen chart. See Module #4 of this Appendix for a formal analysis of the robustness of the Pelli-Robson charts by location and other variables.

VARIATION IN UNDERSTANDING

There appears to have been relatively little variation in how respondents viewed the purpose of either the visual acuity or the contrast sensitivity tests. In both cases, the relationship between outcomes on the assessments and driver competency was apparently obvious. Therefore both tests were seen, almost uniformly, as integral to the driver competency assessment repertoire of CA DMV. There was, however, some variation in what respondents perceived *customer* reaction to be to the new contrast sensitivity test. In some sense, the Pelli-Robson chart appears to have been the "public face" of the 3-Tier Pilot; nearly all the field office staff, and even some of the upper management at DMV Headquarters, noted the public's reaction to this assessment tool. There were a few who anticipated customer fears or anxieties about instituting a new test:

> "I thought it was great. I didn't have any problems with it. [But]…the customers were afraid of it. Because the media had gotten ahold of it, and blown it out of proportion. And [so] they were afraid of it when they walked up. And they were asking about it before we even started going into the process."
>
> LRE #1

However, in most cases this anxiety simply provided opportunities to answer customer questions, which then resulted in quite positive interactions:

[20] No-one at the Fairfield office raised concerns about lighting, glare, or shadows on the Pelli-Robson chart. There exist any number of possibilities to explain this exception, not the least of which is that the ambient lighting may simply have been more even there—this was the only office that wasn't in the Central Valley, where the summertime sun is particularly strong.

"[laughter] It was funny. I mean, certain people were like 'Are you kidding me? What is this?' To me it was kind of fun, because people were open to it."

MVFR #3

"Well, people were thankful that we were doing something like this…because some people didn't know what the fog chart was for. And once we explained that to them then they were like 'okay!' It was only probably a handful that I had that were negative about the 3-Tier."

MVFR #7

"Let me give you example: [the] fog chart. I'd explain to 'em that this is to [test] the contrast. I would tell them…'We lose our contrast just like we lose our vision. So we're trying to determine, because contrast plays a big part in our lives and in driving. Like whether you're looking down the road and you see a dark car versus a light car at the side of the road. Which one are you more apt to see?' So I went into that, explaining to them little situations and stuff. And they would say, 'Oh! You know I did not know that. I didn't know anything about that. That's really good—I would have never thought of that! That's a really good thing to have, then. I think they should have that in all offices!'"

3-Tier Manager I #1

"Most of the customers enjoyed the fog chart. They said 'Wow! This is cool!' And in fact our non- customers were going 'What is that?' So we let them practice and it was great."

Upper Management/Headquarters Coordinating Personnel #12

These comments suggest that the institution of a new assessment tool can, in some cases, provide the opportunity to enhance customer service. The very newness of the chart instigated customer questions, which staff then answered—evidently in ways that then typically produced positive customer reaction. These comments regarding the Pelli-Robson chart also provide indirect evidence for the point discussed above regarding the "active department" purpose of 3-Tier. According to those interviewed, at least some portion of the positive customer comments came in the form of implicitly commending CA DMV for proactively adopting new strategies to ensure traffic safety.

The positive views of staff and managers toward the Pelli-Robson chart appears to have rested largely on what those interviewed saw as the obvious utility of the test for

identifying driving-relevant visual impairments. This was emphasized by respondents across all job categories:

> "The fog chart for instance, the new tool that we got? It was able to identify vision problems that we wouldn't normally identify. So that was an excellent tool."
>
> Upper Management/Headquarters Coordinating Personnel #14
>
> "A lot of [the customers] were amazed at the vision. They thought the fog chart was an excellent choice, because we were getting a lot of people that had vision problems just through that fog chart. Once they came back with DL62s[21] from their doctor, we were seeing lots of problems with their vision."
>
> LRE #3

There appear to be two key aspects to the points raised by these respondents. First, customer outcomes on the Pelli-Robson chart were usually affirmed by the judgment of a licensed vision professional. Secondly, outcomes on this test were routinely confirmed as stemming from driving-relevant health problems—in this case vision disorders. For both reasons, managers and staff found the purpose of the test obvious, and perhaps as a result, easily explainable to customers. While some staff expressed discomfort regarding the physical observation protocol—which is not necessarily confirmed by a non-departmental health professional—no-one expressed the same reservations regarding either the Snellen test or the Pelli-Robson chart. Moreover, no-one raised in regards to visual acuity or contrast sensitivity chart the kind of face-validity critiques reported for the memory recall test.

3-Tier Process Elements: The Perceptual Response Test (PRT)

There were thus four assessment tools that constituted Tier 1 of the 3-Tier process: the memory recall test, the physical observation protocol, and the two vision tests for visual acuity and contrast sensitivity. If a customer passed all of these assessment tests with no identified flags, they took their written law test and (if they passed that) had their

[21] A DL62 is the name of the DMV form given to a customer who fails the department's vision standards. It requires that a customer visit a licensed ophthalmologist or optometrist for a professional examination of their visual health. The ophthalmologist/optometrist then fills out the form (which is rather detailed) with their diagnosis of any identified health problems, information about any prescribed corrective actions, or (if this is the case) their prognosis regarding progressive disorders.

license renewed. If, however, a customer's performance on any of the Tier 1 tests raised flags regarding potential cognitive, visual, or physical limitations, they were then further assessed in Tier 2 of the process.[22]

The primary component of the second tier was a computer-based tool called the Perceptual Response Test, also known as sub-test 1 of the Useful Field of View test (UFOV) (Janke 2001; Owsley et al., 1991). In brief, this test required that a customer use a computer touch-screen to correctly identify schematic images of cars and trucks. An image would flash on the screen for a variable amount of time (between half a second and 17 milliseconds). This image was then followed up by a "snow" screen of randomized white and black "bits" which removed any after-image on the test-taker's retina. The customer then had to select (from a choice of two) which image had flashed on the screen. The computer program would vary the amount of time the images would flash on the screen in order to establish the minimum amount of time the test-taker required to reliably (at least 75% of the time) identify the schematic car/truck images correctly. The "score" on this test thus varied from 17 to 500 milliseconds, and constituted an indirect measure of the processing speed of the visual system. In other words, the score on this test measured the amount of time someone needed to see something of which they had only a brief glance. The amount of time a customer took to make their choice and press the screen—in other words their reaction time—had no bearing on their final score, though the program was pre-set to abort if the test was not completed within two minutes. In the traffic safety and public health literature, this test has been found in multiples studies to be reliably associated with the early stages of dementia-type cognitive disorders (Owsley et al., 1991). More specifically, performance on this test has also been found to predict crash risk (Clay et al., 2005; Hennessy, 1995).

VARIATION IN IMPLEMENTATION

Because this was a computer-based assessment, there was relatively little room for employee-induced variation in the administration of the test. In some cases respondents reported giving the test multiple times to the same customer; however, according to

[22] If a customer failed the written test twice, they were also flagged for assessment with the PRT. This trigger was instituted in order to test the hypothesis that multiple failures might indicate a limitation in cognitive function (see Hennessy & Janke 2005).

training protocols, this was within the discretion of the staff person administering the test.

VARIATION IN UNDERSTANDING

There was a great deal of variation in what staff and managers thought of the PRT. This variation occurred across three dimensions. The first had to do with normative judgment of the worth of the test as an assessment tool: while some respondents saw value in it, many more critiqued the test in various ways, including most particularly the face-validity connection to driver competency and traffic safety. In a subsidiary and related critique, respondents reported a great deal of customer confusion and anxiety about this test—especially among senior citizens—much more than as compared with other elements of the 3-Tier system. Many of these same concerns echo those raised for the memory recall exercise. On a second dimension, there appeared to exist a substantial degree of variation in understanding what the test measured. Many of those interviewed thought that the PRT was a reaction-time test, which was not in fact the case. Thirdly, in a manner similar to that evinced with the memory recall test, respondents discussed the PRT with one of two views regarding the nature of testing: as a flag for some kind of limitation or, alternatively, as one of a series of hurdles to be overcome for licensure.

In terms of normative evaluation, a few respondents offered relatively positive views of this element of the 3-Tier program:

> "The PRT seems to be a good process. I like the PRT… [it's] going to tell exactly what they can grasp."
>
> Upper Management/Headquarters Coordinating Personnel #3
>
> "I do like the PRT machine. There's nothing wrong with it. I think it's an eye opener."
>
> LRE #6
>
> "I even had customers tell me, especially when they got to the PRT, 'You know, I believe this is a really good thing, this program that you're doing.'"
>
> 3-Tier Manager #1

Others, however, simply did not like the PRT. This normative assessment was

sometimes expressed in explicit ways; one 3-Tier Manager I even went so far as to say "I hate that machine." More usually, however, respondents voiced their views in less explicit, but still clearly critical ways:

> "The PRT was a big logjam."
>
> Upper Management/Headquarters Coordinating Personnel #12
>
> "I'd say that just from overhearing and seeing how the techs did, that the PRT was an effort. It took a lot of effort, I think, on everybody's part."
>
> LRE #4
>
> "The only part of [3-Tier] that I thought [customers] had difficulty with—I didn't, but I was a little uncomfortable for them—was when it came to setting up the PRT."
>
> Administrative/Back-up Manager #1
>
> "People that were taking it? I got a lot of negative feedback from [them]. Typically there was negative feedback from people who didn't quite understand it, or who didn't do well on it, or who couldn't see it right when the car flashed…"
>
> MVFR #8

In general, criticisms of the PRT outnumbered positive evaluations. Given the nature of the sample, however, it would be difficult to estimate with any precision the percentage of staff who held negative versus positive views of this element of the 3-Tier process.

Concerns and critiques of the PRT generally focused on customer confusion, especially among senior citizens:

> "I thought the PRT was an excellent source of testing. We did have confusion with it…I would go in and just go through the process with them because a lot of the elderly people, the minute they see a computer, they freak: *I don't do computers!* And so they had this mindblock that they couldn't do it. But once I would go over there and explain the test to them, then they would understand it. And we usually wouldn't have a problem."
>
> LRE #3

These concerns repeat, in a slightly different form, the points raised above with the memory recall test, regarding how tests should apply equally upon all customers. Here again, the population of concern—the customers for whom this test was especially frightening or confusing, and who were evidently more likely to fail—consisted of senior citizens. That seniors had difficulty with the PRT was seen by many respondents as a problem with the test—and not as a potential sign of cognitive difficulties.

Much of this criticism of the PRT centered specifically on the instruction module which introduced the test:

> "The PRT was explained totally differently from how it operated….okay, let me see how I can [explain] this: I sit you down, and we do the test—I mean we listen to the instructions and then we do the practice test. At no time is it explained that as you go along the car and the truck will appear faster. So you take the practice test, okay, you get them correct. But then you start the test and you're going fine and the next thing you know, they're going really, really fast…"
>
> Upper Management/Headquarters Coordinating Personnel #9
>
> "Maybe a little slower on the PRT when giving the instructions to the customer on how to use it? A lot of them wanted to read it three or four times. And with the practice, a lot of them were concerned that they only got three tries to practice and they weren't getting it because they didn't understand it. They thought they were supposed to be looking at those "magic screens" where you stare at the static and, you know, a picture pops out or something like that. Really, we did get that [laughs]."
>
> Upper Management/Headquarters Coordinating Personnel #12

In some cases, respondents reported that the PRT presented opportunities to provide extra levels of communication and personal attention to their customers—though they found this enhanced customer service necessary precisely because of the anxiety produced by the test:

> "[It's] about learning how to talk to people, and making them feel at ease. Especially with the PRT machine. [If] you sit an older person down to a computer screen, they're thinking 'It's a computer!' They don't realize how simple it is until they get started, and they're afraid right away…"
>
> LRE #1

"The customers thought it was unique. And a lot of the older customers—when we'd have to do the PRT tests?—they thought we were taking extra time to help them, you know? So it was a positive interaction with them, I think."

MVFR #11

Regardless of whether staff thought the PRT was "a good process" or they "hated it," administering the test required staff time to calm customer anxieties and to answer questions. Much of this time involved human communication to supplement that which was provided in the introductory module to the test.

To the degree that staff had to interact with customers and explain the PRT, how to take the test, and what it measured, clearly the employee must have accurate information at hand. Unfortunately, it appears that many staff and managers, at all job classification levels, had disparate and often quite incorrect information about the PRT. When asked "what is the PRT designed to measure?" many respondents remarked that it measured reaction time:

"I believe the PRT was designed to measure their reflexes, and their ability to act in a timely manner?"

Administrative/Back-up Manager #2

"The PRT, what their reaction time was; whether that person sat there for five minutes as the thing slipped right through and [they] didn't make any attempt to even read anything."

Upper Management/Headquarters Coordinating Personnel #8

"Their perception, and to see how…when something is flashed in front of them real quickly, how long it takes them to respond. So basically like if they're out driving how would they be able to react if something happened like split second."

MVFR #7

"The response time, you know? If someone were to see something, their response time. How they would react."

MVFR #1

"It measures your reaction time between seeing an object and reacting to it on the screen. How long it takes you to react."

3-Tier Manager I #3

"Your ability to recognize an object and have your brain send a little electrical signal to your finger to hit that little object real fast."

Administrative/Back-up Manager #6

Even when respondents explained the PRT as a test of perceptual speed, rather than reaction time, their explanations sometimes conflated the two. Indeed, some staff suggested that the PRT measured both properties at once:

"It was designed to measure how quickly somebody can observe an object, know exactly what it is, and respond to that object. I noticed with some of the older people that they had trouble actually trouble observing the difference immediately. That's why I think we had to repeat the test on some people. It wasn't that they were unresponsive timewise, it was that they weren't picking up the difference between the car and the truck initially."

LRE #2

"The PRT was designed to measure how quickly you recognize and how fast you act. It was measuring your reaction time and [your] visual time."

3-Tier Manager I #4

Others, however, noted that the test had to do with seeing things that were moving quickly, rather than making quick choices:

"Their perception of what they're seeing, and how quickly they're seeing. And if they can identify what they're seeing."

LRE #3

"It's designed to measure your ability to perceive and recognize objects. Such as if you were driving down the road and you just glance very quickly to your right, are you actually able to recognize that there is a car in your blind spot? That's just kind of a practical application, but the computer was basically just seeing if they could recognize objects at a fast rate. Perceive what they were, recognize them, and process that information."

Administrative/Back-up Manager #1

"Well, I assume it had to do with being able to see things, you know? Like whether it be a child versus a ball or a this or a that… particularly if it's moving quickly."

3-Tier Manager I #1

This last (correct) understanding of what the PRT measured—perceptual speed—was rarer than the (incorrect) understanding that it measured reaction time. Moreover, the reaction-time definition was common across job categories and across field offices—suggesting that the source of this misunderstanding was general, and likely lay in the materials presented during staff training. That said, there does appear to be some relationship between the amount of training in the PRT, and the likelihood of articulating a correct understanding of the test's underlying purpose. Front-line employees (MVFRs) were most likely to express an incorrect understanding of what the PRT measured; these employees received the least amount of training on this particular process element, as compared to managerial-level employees. In fact, the bulk of the training given to MVFRs in this process element was done in the field offices, by their respective 3-Tier or Administrative/Back-up Managers. The latter employees—who were somewhat more likely to articulate a correct understanding of the purpose of the PRT—were trained at DMV Headquarters, by departmental Training Branch or R&D staff.

But there also appears in some of the interviews a second-order issue regarding how staff understood the test. Many staff regarded the instructions as unclear and confusing, and others found that many customers experienced significant anxiety when told they would be using a computer. In both cases, the implicit argument was that the problem was *with the test*, rather than with (for instance) the customer's cognitive abilities. After a second set of (hopefully clearer) instructions, given either by the same staff person or by someone else in the office, customers often ended up taking the PRT multiple times:

> "It seemed like when people came for renewals the ones who seemed to be failing [the PRT] the most were of a non-computer era. And my understanding was that they didn't understand the PRT, so I told [the MVFRs], if you get aborts, I need to see them. And I would talk to them and I realized that they didn't understand what the technician told them. Even though they were concise and clear…So I said 'if you get an abort, bring it to me…'"
>
> 3-Tier Manager #5
>
> "Oftentimes we had to go through a whole test, and realize that they still didn't get it even after they did the pre-test. And so we would set them up a second time and, you know, oversee them a little more closely."
>
> LRE #2

Taking the PRT multiple times was certainly within the pilot protocols. There is even some evidence from the academic literature that training and practice sessions improve an individual's perceptual and speed-of-processing abilities, both in the short- and long-term (Ball, Beard, Roenker, Miller, & Griggs, 1988; Roenker, Cissell, Ball, Wadley, & Edwards, 2003). However, the "practice" sessions described by the respondents in this study were *ad hoc* and unstandardized. It is thus unclear from these data whether the improvements in customer performance described by staff are a result of improvements in customer understanding of how to take the test (the general view of the staff) or of improvements in perceptual speed resulting from improvised practice and training. More to the point, the very idea that practice and better instructions are at the root of the "problem" of PRT failure suggests a substantial misreading by some staff regarding the nature and purpose of assessment testing. As articulated explicitly by one LRE:

> "The problem that I saw, being an examiner for so many years, is that the [MVFRs] that were putting them on the machines were trying to make [the customers] pass, not realizing that they're not an examiner. Because I could identify that some of these people that weren't passing, it was not because they were confused. It was because there was an issue. But the MVFRs, because they haven't been an examiner, they didn't understand that. They just immediately assumed that "they're just confused with the test." I'd say probably 80% maybe didn't understand, and they would pass on the second time. The other 20%? They had issues, that's why they weren't passing. But the MVFRs, the technicians, they were just giving it over and over so many times…"
>
> LRE #3

The LRE viewed customer confusion as not necessarily stemming from poor instructions. Rather, customer difficulty on the test was in their view more likely a sign of some cognitive difficulty. This was a distinctly rare understanding of what the PRT was for, and appears (if at all) among LREs and 3-Tier Manager Is. In other words, those staff members most highly trained in judging driver competency—and who received more formal training in this particular assessment tool—were also more likely to view the PRT as a potentially useful tool in this endeavor. That others—particularly MVFRs—were "just giving it over and over" suggests in the first instance that many

staff were profoundly skeptical of the worth of this assessment tool. But it also suggests that at least some front-line staff saw the PRT as simply one in a series of hurdles that all customers had to overcome in order to receive their license. In a word, the point of the PRT was seen—like the written test on the rules of the road—as one stage (among several) that had to be "passed" in the licensing process.

3-Tier Process Elements: The Educational Intervention

For some of the test elements of the 3-Tier process, a customer either "failed" or "passed," where failure constituted a flag for higher-level assessment. In the case of two test elements—the contrast sensitivity chart and the PRT—customers could also have mid-range outcomes, referred to in project jargon as "somewhat fails." This category was created on the basis of work by Hennessy and Janke (2005), which indicated that drivers at the beginning/early stages of progressive disorders are more at risk for crashes than those with more advanced-stage disorders. The authors hypothesized that this was due to the fact that those at the early stages of (for instance) a vision problem were less likely to be aware of their limitations, and hence less likely to take the kinds of precautions necessary to ensure continued safe driving. As a further implication of that work, the designers of the 3-Tier Pilot developed two educational videos (each in a 4-minute DVD-based format), one for contrast sensitivity and one for the PRT. Both videos discussed the nature of the flagged limitation, and the relevant driving strategies which a driver could use to compensate.

In order to assess scientifically the impact of these videos on driver performance, the project incorporated a randomizing experimental design whereby some customers—those who had "somewhat failed" either the contrast sensitivity chart or the PRT, and who also had driver license numbers ending in an odd digit—would be shown the videos. In most cases a customer would see only one video, though if they had somewhat failed on both tests (and had an odd driver license number) the protocol called for them to view both. All other customers—those who had fully, or "extremely," failed the contrast sensitivity chart or the PRT, as well as those with even driver license numbers—were not supposed to be shown the videos. However, they may have received other drive test preparation materials if that was appropriate for their particular case. In any event, the 3-Tier Manager I at each office typically distributed

these educational materials, though other staff may have taken over these duties due to vacations, health leaves, etc.

VARIATION IN IMPLEMENTATION

There were two ways in which the implementation of the educational intervention varied. The first had to do with variation across videos: those interviewed reported that they administered very few educational interventions for the PRT. The second source of variation had to do with adherence to the randomizing experimental protocol. Here it appears that there was substantial variation across offices, with some staff administering the videos quite freely, and not simply to those that had somewhat failed, or who had odd-numbered driver licenses.

Those who administered the educational intervention reported that they gave out very few educational materials related to the PRT:

> "I did a few on the PRT, but not as much as I did on the fog chart."
> 3-Tier Manager #1

This evidently had largely to do with the fact that the Pelli-Robson chart is a somewhat more sensitive test; a higher proportion of customers "somewhat failed" (and so were eligible for educational intervention) than on the PRT. Thus, as reported in one exchange with another 3-Tier Manager I:

> INVESTIGATOR: How many times did you give [the educational intervention for the PRT], by the way?
> RESPONDENT: uh….[long pause]
> INVESTIGATOR: Often? Not often?
> RESPONDENT: Not too often. And the reason for that was because if they aborted, we went back in and we re-did the PRT. So those that would have been caught in that, we were able to walk them back through it, and they understood it. So I didn't have to give that as often.

Other respondents (see above) also reported encouraging customers to take the PRT multiple times. However, it is not clear—at least from these data—whether this effected the distribution of the educational videos. Where staff reported that they had customers

attempt the PRT a second (or third, etc.) time, this appears to have been primarily in cases where the customer "aborted" (or, in pilot jargon, "extremely failed")—i.e., did not complete the test within the 2-minute time limit. Staff did *not* report administering multiple tests in cases where the customer "somewhat failed" the PRT, i.e., had a score between 24 and 40 milliseconds.

The interviews also indicate that some respondents did not follow the protocol indicating randomizing experimental distribution of the educational intervention, where only those who "somewhat failed" the tests, and who also had odd driver license numbers viewed the material:

> "They were very good tools, and I know [this office] leaned on them heavily. Our 3-Tier Manager I told me 'If you ever have to, make sure you show these before you schedule the drive test, because they are very good educational videos, and they do give out a lot of information to our drivers that have to take the drive test for 3-Tier.'"
>
> Administrative/Back-up Manager #5

Thus, in at least this case, the educational videos were given to anyone that was scheduled for a drive test—regardless of their driver license number. Similar practices occurred at other offices as well, as indicated by this LRE:

> INVESTIGATOR: What did customers think of the educational videos?
>
> RESPONDENT: …They felt that that was good education for them. And the ones I would put on there that I didn't need to, I felt that maybe it would better explain to them. And they understood it. They took the educational materials and were pleased to have them…
>
> INVESTIGATOR: So the customers who benefitted from the educational intervention, were those folks who were going to go out for a drive? Or were they folks that maybe got a [somewhat fail] on the fog chart? Or people that took the PRT and did fine and so weren't going to go out for a drive?
>
> RESPONDENT: All of those people were put on it. Because I, as an examiner, when I was doing the interviewing, if I felt that my customer didn't

> understand why they had to take a road test that week, I put them [in the educational intervention] for whatever reason.

Many staff found the videos (especially the one developed to explain contrast sensitivity) quite useful, and this appears to have driven the deviation from pilot protocols. They were useful not simply for the subject at hand—understanding how to compensate for limited contrast sensitivity—but also for preparing for the drive test more generally. As such, they were a valuable tool for providing personalized customer service.

VARIATION IN UNDERSTANDING

Those interviewed for the project had very positive things to say about the educational intervention videos, though this was more true for the video regarding contrast sensitivity than for the PRT video:

> "[Customers] didn't get as much out of [the PRT video]. It was kind of redundant in what it said. We didn't have to give it all that often, either. Which was a good thing."
>
> 3-Tier Manager I #3

In particular, staff found the videos useful as customer service tools, e.g., for calming customer anxiety:

> "I think that they were very helpful. Most people, when put in that situation…were somewhat defensive. 'I've never had an accident. I've never had a ticket. I've driven for 40 years. Why am I having to go through this?' But at the end, they said 'Thank you very much—you made this very easy. And I appreciated this information.' So when the process was finished, they seemed to be appreciative. And some even expressed that they learned things that they hadn't thought about in a long time."
>
> Administrative/Back-up Manager #1
>
> "Well, they didn't understand why they were taking the test, so I would explain to them… Most of them ended up having to complete the educational intervention. They loved it—they were like 'Wow! That was great!', but even

> when they failed they still asked me 'Are you going to be doing the drive? Because you're a really a nice individual…'"
>
> 3-Tier Manager I #5

When faced with heightened levels of scrutiny regarding their driving competency—especially if this involved taking an on-the-road drive test—many customers expressed frustration or confusion as to why they were being assessed. Those charged with the task of explaining this—3-Tier Manager Is and Administrative/Back-up Managers in particular, who coordinated the scheduling of 3-Tier drive tests—found the educational videos especially helpful for supplementing any verbal explanations they provided to the customers regarding the pilot program or the new assessment tools.

Staff also used the educational intervention videos as preparatory materials for the on-the-road drive test. As indicated previously (pp. 73-74), some staff provided the videos to all (or nearly all) of those customers they scheduled for drives. The educational videos were useful for their pedagogical content, regarding how to drive safely:

> "That was part of my job also. Because of the fog chart, or the PRT, once we explained to them why [they had to take a drive test], then…it was okay. Because they knew that this was a safety issue. …And the videos were a very big help, too. Those were very good, [very] explicit, very well done."
>
> 3-Tier Manager I #1

They were also useful for calming drive test-related customer anxiety:

> "I think one of the most important components of the pilot was the education for people who need to go out on the road. It seemed to reassure them. It made it less intimidating…it seemed to be valuable with the customers that I [worked with]."
>
> Administrative/Back-up Manager #1
>
> "Once again, that's what those videos are for. It's to relieve their cares about the drive test and to reassure them that this is in no way going to take away their privilege. We just need to reevaluate their driving skills because from our information they might be a candidate for a serious incident, and we want to make sure that doesn't happen. So we want to educate them as much as we can."

Administrative/Back-up Manager #5

The usefulness of the educational intervention for the drive test was indirectly confirmed by at least one of the LREs during their evaluations of driver skill:

> "I want to bring up [something] that I noticed personally on the drive tests—because I love to gather information—the customers who actually sat down with the [3-Tier] manager, and viewed the videos, did much better than our vision [referrals] who never viewed them, who didn't go through any training, but just came in, saw their eye doctor, and we took them on a drive test. I thought [the videos] were fabulous. I thought it brought [our customers] more awareness, and more education. More like 'Wow! Maybe this is something I need to look into or Maybe this is something I need to practice or Maybe this is why this happening.' I think education is always important, and I would like to see that happen for all our vision [referral] customers."
>
> LRE #6

In sum, the educational intervention—especially the video on contrast sensitivity—was seen as tremendously useful by 3-Tier staff inasmuch as it helped calm customer anxieties and reiterated information provided verbally by the 3-Tier Manager I regarding the new assessment tests. The educational videos also, according to those interviewed, materially improved drive test results. Neither of these outcomes, however, constituted the explicit purpose of the educational intervention—which was to test a hypothesis regarding long-term driver behavior as influenced by knowledge of incipient or progressive limitations. Instead, respondents understood the utility and purpose of the video in much more immediate terms: the provision of customer service, and the preparation of customers for passage on their drive test.

3-Tier Process Elements: Drive Tests

The on-the-road drive test—CA DMV's ultimate measure of driver competency and skill—constituted the third tier of the 3-Tier process. Very few customers—less than 10% of all customers enrolled during the pilot—ended up having to take a drive test. Those that did had two options. The first—called the Supplementary Driver Performance Evaluation (or SDPE)—is nearly identical, with a few exceptions, to the test used for original (new license) applicants, the Driver Performance Evaluation

(DPE). These exceptions are generally intended to assess cognitive function. The first involves a multiple-directions task that requires a customer to complete a series of maneuvers (say, turning right at an intersection, and then changing lanes) in sequential order without prompting; this assesses the customer's short-term memory. Partly because of this multiple-directions task, the SDPE incorporates a slightly greater number of lane changes than the DPE. Thirdly, the SDPE requires that the examiner chat or converse with the applicant while the test is in progress; this evaluates the customer's ability to concentrate on driving while their attention is divided. The fourth difference typically occurs at the end of the test. Here, the examiner leads the applicant on a short route (typically involving 2-3 turns) and then asks them to return to their starting point (the field office) using the same route. This tests both short-term and spatial memory. Finally, the SDPE is somewhat longer than the DPE; approximately 30-40 minutes, as opposed to 20 minutes.

The second drive test option—the Area Driving Performance Evaluation (or ADPE)—involves a specified set of routes or destinations worked out between the examiner and the customer. These routes typically start at an applicant's home and proceed to destinations to which they regularly travel, such as the doctor's office, their church, retail stores, or the homes of relatives. The chosen routes determine the extent of route restrictions (which may take the form of a bounded area, rather than specific routes) which the examiner assigns, assuming a satisfactory test result. By taking an ADPE, a customer is restricted to driving within the assigned boundary (or on the assigned routes) within which the test had been conducted.

VARIATION IN IMPLEMENTATION

3-Tier involved few changes of any substance to drive test procedures per se. As a result, there appears to have been no substantial variation in the implementation of this 3-Tier process element—either by office, by employee, or even by job class (e.g., between Hearing Officers and LREs). Although 3-Tier did not involve any substantial changes to drive test procedures, nevertheless the pilot design team anticipated an increase in the use of three procedures: the assignment of "Special Instruction Permits" (SIPs) by Hearing Officers, the granting of 60- or 90-day temporary restricted licenses by

Field Office LREs, and the administration of the ADPE.[23] R&D expected these increases primarily as second-order results of the *general* increase in the number of drive tests expected (and seen) to result from the pilot. In particular, it was expected that a substantial number of those flagged for higher scrutiny because of potential physical, visual, or cognitive limitations might fail their first drive test. As a result, it was expected that those who failed their first driving test would then prepare for a second (or third) on-road test by either (a) taking specialized instruction, (b) practicing their driving with another licensed driver, or (c) electing to take the ADPE (and its accompanying restrictions) as part of a move toward more limited driving. While it is true that some customers had to take multiple drive tests, this does not appear to have led to a wider use of SIPs, to an increase in the assignment of either type of temporary restricted license, or to a rise in the use of the ADPE. The relative rarity of these procedures was true of all six field sites, as well as of Driver Safety referrals to the field offices.

VARIATION IN UNDERSTANDING

Many of those interviewed noted that customers were often anxious or even frightened at the possibility of having to take a drive test. However, after having been led through a process of assessment testing focused on driving-relevant potential limitations, customers also understood—at least according to staff—why they had to demonstrate to CA DMV their ability to drive safely:

> "They were prepared, because of this process, which I thought was good. Because it made the customer feel that the department understood them. But they had to meet these requirements. So, for instance, they understood the importance of the law test; that if you don't [pass] it that doesn't mean you can't get your license, but now this is the next step…"
>
> LRE #4

[23] SIPs are typically given by Hearing Officers to suspended or revoked drivers who are trying to reinstate their driving privilege. Temporary restricted licenses (60- or 90-day) may be issued to licensed drivers seeking renewal. In both cases, they are typically issued to customers who, in the judgment of the LRE or Hearing Officer, must practice their driving with another licensed adult driver or driving instructor before taking an SDPE or ADPE to renew their license.

Staff also noted that their own explanations supplemented and expanded upon what customers already understood from the process of having been tested:

> "That was another part of my job. And because of the fog chart, or the perceptual response test, it was explained to them why [they had to take a drive test]. Then it was okay. Because they knew this was a safety issue, you know?"
>
> 3-Tier Manager I #1
>
> "Yes, it was good. Because the process… I mean we're giving them the literature already, knowing that they're coming into our office to test for the 3-Tier Pilot. And when they get to the corrections [counter], we give them that advisory statement again. That's where we emphasize them to study a little bit more, before they test again. So then by the time they get to me, they're already aware of it. Everybody that I came across? They were fine with it. Glad to take a drive. I mean they were happy to see us again. Because I usually ask the customers, when they came back from their drive, I would say 'How was the drive?' And they would say 'Wonderful. The examiner was nice, courteous.' So that was good feedback."
>
> 3-Tier Manager I #2
>
> "Initially a lot of them were shocked, you know? They were shocked and a little upset. But once [the 3-Tier Manager I] or myself sat down and talked with them, and showed them the videos, and had them start filling out the questionnaire about their driving habits, then they got the sense that it's not just because we said they had to. They got the sense that we were really trying to evaluate their skills and understand where they were coming from."
>
> Administrative/Back-up Manager #5

Thus, to the degree that taking a drive was framed as "a safety issue"—as opposed to an arbitrary or age-based form of discriminatory action—staff observed that (most) customers saw the need for their drive test, and prepared themselves accordingly.

One of the few minor changes to drive test procedures during the 3-Tier Pilot involved a scripted set of questions that DMV staff would ask the customer regarding their scope of driving (see Sub-Appendix C). These questions were intended to elucidate information regarding which drive test—the SDPE or the ADPE—was most appropriate for the customer, depending on their particular circumstances and driving needs. In

terms of their general utility for enhancing customer-staff interaction, respondents generally found these questions quite helpful:

> "I liked being able to have the freedom to take that extra time with the customer. I didn't feel rushed…because I was reassured by management that I could take the time to spend with the customer, to explain what to expect. And a lot of times we have a better result when we take that time, ultimately. Because they tend to relax when we…explain to them what to expect, what we're looking for, and just to reassure them to relax. It makes a difference, I think."
>
> LRE #1
>
> "[3-Tier] required me to speak more in-depth to the customers, and to try to make them less fearful of the process of going out on a drive, what it was about. To encourage them that despite whatever disabilities they might have, that didn't mean that they couldn't be a good driver. They could still be a good driver, with some disabilities, and/or that they could improve their driving, that they should start paying more attention to certain things."
>
> LRE #2

These questions—and in fact, the entire process of assessment testing more generally—not only calmed customer anxiety, but encouraged them to prepare, in various ways, for their on-the-road exam. In terms of the specific wording, however, staff were somewhat more qualified in their judgment, suggesting that a yes/no question format did not necessarily provide the right kind of information they needed to work with a customer regarding their drive test. In that respect, the questionnaire appears to have functioned less as a definitive set of criteria on which to make recommendations regarding the choice of SDPE versus ADPE, and more as a set of prompts regarding areas to cover in conversation, using perhaps alternate phrasing.

As an extension of this view, some of those interviewed highlighted a secondary purpose of the drive test. Not only did it serve as an opportunity for the department to determine driver competency and skill, the test itself served as an opportunity for the driver to improve their driving habits. This came up in several interviews:

> "The pilot was implemented in a way that we would actually be able to give people more information while we were collecting information from them, so it was a two-way benefit. It was a good relationship. We got something from our

customers, but the customers who needed it got something from us. They got either information, education, or were possibly given a drive test (if they needed one) to assess their skill."

Upper Management/Headquarters Coordinating Personnel #10

"I think the majority of the applicants that went through 3-Tier…it changed their habits of driving. Because most of those that undergo the SDPE, most of them have been licensed for more than 20 years. Some for 40 or 50 years. When they found out that these are the new rules of the road, and their driving skills and habits…they've been telling me 'Oh, I didn't know that. I need to do this and do that? It's much safer for me, and much safer for the public?' So it affected them, really. Because some of them—especially those that failed the first SDPE drive, when they came back, they came having undergone some kind of driver educational training to update their skills. So it changed the habits that they were doing before."

LRE #5

A key part of this preparation—one emphasized in the questions developed for the 3-Tier Pilot, involved private instruction:

"I would always question [the customer] on the way out to the car and ask 'Since you knew you were going to take this drive test, have you had any professional driving lessons? Or Did you drive with a friend or family member or do anything else that would prepare you for this drive test?' That was always a very important question. I always found that extremely helpful."

LRE #6

"To help them, hopefully to better prepare them to make them[selves] more aware of some shortcomings [and] to help them improve their driving. One of the really good things I've noticed [was] among the people that answered in the affirmative one of the questions in 3-Tier: 'Since you found out you had to take a drive have you had any kind of instruction?' Most people answered no, but quite a few people—I don't think it was half—said 'Yes,' particularly if they had failed the first drive."

Administrative/Back-up Manager #6

"I've had [customers] who have gone and gotten additional training because that's what we suggested to them. We encouraged it because of how poorly they

> did on their SDPE, so we suggested that they get some professional driving school. They came back and they went 'I never realized how poorly I drove.' They got the instruction, they came in, and they passed the test. So overall, giving the additional SDPEs is wonderful. Yes it means we're doing more drives in the field office but I think it's more of a service to all customers, to all citizens…"
>
> 3-Tier Manager I #3

Anecdotally at least, LREs and 3-Tier Manager Is in particular observed a distinct improvement in drive test outcomes among those that elected to take private instruction in preparation for their on-the-road exam. Thus, many of the staff interviewed for this project observed that when customers had to take a drive test, this improved their driving skill for two reasons. First, as a direct effect of the test itself, inasmuch as preparing for the exam served as an impetus to concentrate on driving habits. Improvement in driving skill also occurred through receipt of preparatory private instruction. However, as noted above, very few staff reported that it was necessary to *require* private instruction (via a Special Instruction Permit or 90-day Temporary Restricted License); instead customers sought out instruction voluntarily.

The training customers took in preparation for their drive test(s) constitutes one of three reasons why, according to staff, so few people elected to take an ADPE. In the first instance, many saw the ADPE as unnecessary for most customers simply because they could pass on an SDPE drive test with sufficient study and practice:

> "We encouraged [customers] to undergo that specialized professional training first. Because there are three chances that they're going to get, and we give them all the chances first. If we find that they're missing these kinds of skills, to overcome them we suggest 'Maybe you could drive with a member of the family. Or perhaps it's best if you get professional driving instruction?' When they get that, they're really prepared. So the second or third [drive test], they pass. So we don't go to the ADPE."
>
> LRE #5

By encouraging customers to enhance their driving skills, staff also—explicitly and implicitly—encouraged customers to seek unrestricted licenses:

"I like the fact that we tried the SDPE. We try to get them through it before we do the area drive. We limit what we can so that we don't have to do that. And I don't think a lot of customers want to be boxed in like that."

3-Tier Manager I #3

"I would tell them 'This is a formality. We're doing this to gather information. If at any point you feel that you're not comfortable, you do have at least three tries. We can always do an area drive, but let's don't go there. Let's be positive. Let's see how you do.'"

LRE #6

This dovetailed with the second reason for the rarity of ADPEs, which is that staff and (by report) customers saw the area restriction as a "last resort"—a test that a customer took because they had few other choices, given their inability to pass an SDPE.

This second view was shared by the many of those interviewed, though often expressed in slightly different ways:

"I would explain to them, go over their drive with them on the errors that they made, [saying] 'Well, because of what you did here, which were basically critical errors, we have no choice but to do this area drive. This is your option right here. Either that or be revoked.' And of course none of them wanted to take it."

3-Tier Manager #5

"Once you do an area drive, it was one opportunity at an area, and that's it. You don't make it, you're done. So knowing that, people don't want to put themselves in that kind of position. And some people don't want to be restricted to an area. Some older people have a tendency that if you're taking away their independence and [if] you're trying to dictate to them what they're capable of doing and [what they] can no longer do, then they're not agreeable to being restricted. And they don't want any part of that."

Hearing Officer #2

"In any case the ADPE is always a last resort for us. Regardless if it's 3-Tier or if it's requested through Driver Safety, [with] the ADPE they only get once chance, that's it. If they cannot pass in such a restricted environment…we send it up to Driver Safety for further evaluation."

Administrative/Back-up Manager #5

The ADPE was seen by most of those interviewed as the last stop on the road to revocation. It was therefore a less-than-attractive option; partly because of the nature of the area/boundary restrictions, but more because of what it represented symbolically in terms of loss of freedom and the (likely) impending loss of the driving privilege.

There was yet a third reason why ADPEs may have been administered only rarely during the 3-Tier Pilot; most staff saw them as an inappropriate option for drivers living in urbanized or even suburban areas. This view was shared by both field office and Driver Safety staff:

> "For 18 years I gave drive tests. And the rule of thumb back then was [ADPEs were given] only if they lived in a rural area. That is no longer the case."
>
> Hearing Officer #2
>
> "In four years I think I did five Area drive tests [when] I was in Yuba City. So we were in a more rural area—well, not more rural, but you know it's a little less populated. And that's really what we saw, so I would expect [fewer] in the Sacramento area."
>
> Hearing Officer #3
>
> "For the most part, area drives are not realistic. A lot of people will tell you they don't drive much, but when you get out to their house they drive all over everywhere. And so area drives just aren't a realistic thing in this area. It's very crowded, population-wise."
>
> 3-Tier Manager I #3
>
> "The original ADPE was for people that lived very rurally. And drove the one-lane country roads."
>
> LRE #1

This came up in a different form among other staff, who regarded an area drive as inappropriate if a customer's chosen routes required that they demonstrate skills that would otherwise qualify them for an unrestricted license. Hence:

> "I'm obviously not an official spokesman or anything, and not a policymaker, in the area of area drives. [But] if you live in downtown Sacramento, why am I going to give them an ADPE? What if somebody only lives two blocks from this office? I had one person where he was doing his ADPE, and it covered most of our [SDPE] route, so what's the point?"

Administrative/Back-up Manager #6

"I had a gentleman that went to the Elks Lodge every day. And that was his area, which was fine because it was just a few miles, a straight shot and back. But then when I get back with him, he wants to go down to [two major thoroughfares in the greater Sacramento area]….and I had to explain to him, trying to be as fair with him as I could, that that doesn't constitute [an area drive]. If you can drive in…high-traffic areas, then you should be able to pass our [SDPE] test at the office."

LRE #3

Given that all six field offices were located in cities with at least 90,000 residents (excepting Folsom, which is effectually a suburb of Sacramento), most of those interviewed regarded the ADPE as inappropriate for the overwhelming majority of drivers—even if their skills were limited.

For all of these reasons, those ADPEs that were administered were given only to customers who voluntarily requested them. Many did so precisely because, in their own view, it was the only option:

"Most of those people—the three customers I had take the ADPE—they really wanted to restrict themselves. They didn't want to go on the freeway, or anywhere else, because they were so scared that if they failed the test then their license would be revoked. So they restricted themselves, and that's the reason…This was the last option they had, and they said 'Okay. I will limit myself but I will not lose my license.'"

3-Tier Manager I #4

"There are some of them who did ADPEs right out of the gate. They said 'No, I do not drive outside of this area.' And, you know, you can tell that because they have somebody drive them here to renew their driver license… They saw that in driving in this area, they saw that they didn't have the reaction time, that they had problems with the lights and the stop signs and things like that. And in talking to the examiner who was doing the drive, they realized that being restricted to an area may be their only option. But we let them come to that decision"

3-Tier Manager I #3

> "Realistically, it would have to be a customer requesting it. They would [have to] just say it right up front. Or they would maybe go through one or two drive tests and then somebody gave them the idea. But I never had that come up. Nobody got to that point."
>
> Hearing Officer #3

It thus appears from the evidence gathered in these interviews that some number of the ADPEs administered during the pilot period came from customers whose driving skills had diminished substantially. Despite the fact that an area restriction was an unattractive option, according to the staff there were still some customers who voluntarily elected this drive test option when faced with the evidence of their own limitations. This was, in fact, one of the explicit goals of the 3-Tier Pilot—to set up procedures within the renewal process whereby individual customers would explore different drive test options with their examiner.

Sources of Variation in Understanding and Implementation of 3-Tier Process Elements

PROCESS ELEMENTS

The variation (and lack thereof) described in the preceding pages appears to have derived from multiple sources. In the first instance, the overwhelming majority of respondents raised critiques regarding the training provided for the pilot. These critiques centered on the perception that the information provided during training changed over time, and thus staff received quite different instructions depending on when they happened to take the training classes. This critique appears to have arisen because of changes that were made to the forms used for data collection in the pilot. Respondents did <u>not</u> raise critiques of training that had to do with the explanations and background provided regarding the pilot's goals—as noted above, those interviewed reported a remarkably consistent understanding of what the 3-Tier Pilot was designed to accomplish. Staff and managers also made several constructive suggestions for future training protocols. Given that the brunt of criticism focused on changes that were made to the paperwork, it appears that any inconsistencies in training can explain some, but by no means all, of the variation in implementation and understanding described earlier.

The second source of variation had to do with tension between differing organizational goals. Staff—especially, but not solely, the MVFRs and others working the "front-line" counters in the field offices—operate under conditions that call for the achievement of three organizational goals: (a) increasing productive efficiency (in particular, reducing customer wait times), (b) providing good customer service (which, it turns out, has multiple definitions), and (c) improving/ensuring the safety of drivers on the road. Navigating these three goals simultaneously can produce tension at the margins. This (perhaps irreducible) tension can at least partially help us to understand why staff reported more variation with some project components than with others; those project components for which more variation was reported tend to encapsulate sites of conflicting organizational goals. Goal tension can also explain the source of the two somewhat distinct understandings reported above regarding the nature and purpose of assessment testing.

TRAINING

Nearly all of those interviewed critiqued the training provided for the pilot. These critiques had a number of aspects. First (and most common, especially among managers), those interviewed were troubled by perceived inconsistencies in the information provided during training. This was closely related to a second critique, namely that what changed was the basic paperwork used during the pilot (the Tier 1 Score Sheet and 3-Tier Tracking Sheet in particular). That changes to the paperwork were seen as necessary for simplifying and streamlining the forms for ease of use—and furthermore made largely as a result of input by those trained—did not blunt the force of this critique regarding inconsistency. Staff and managers also made two substantive suggestions for improvement of training: the first involved distributing training materials at the regular (on-site) Wednesday morning staff meetings; managers in particular saw this as a way of mitigating any disruption in personnel coverage at the field office level. Training in the field offices (as opposed to off-site, at DMV Headquarters) was also regarded as a method of ensuring consistency in procedures, at least *within* each office. A second substantive suggestion involved the incorporation of more role-playing scenarios into training; this was seen as a means of familiarizing staff quickly to the new procedures (especially the filling out the forms), of educating Driver Safety on assessment tools that they did not themselves administer, and, finally, of

enhancing the customer-service needs that many saw as intrinsic (though not unique) to the 3-Tier process.

The most common critique raised regarding training had to do with "consistency." By consistency those interviewed meant the uniformity of information presented to those trained. To the degree that those interviewed saw training fall short of the ideal, it was because they perceived that those who went through training did not complete the coursework with a common (uniform, homogenous) understanding of the duties expected of them for the pilot. Basically all Office Managers and Headquarters/Coordinating Personnel raised this point:

> "The quality was lacking in the details of the training. People from the same office would go to training on a different day and get different information, because the rules and the business flow were not settled and were never tested."
>
> Upper Management/Headquarters Coordinating Personnel #1
>
> "It's really important to me that people don't flip the instructions halfway through."
>
> Upper Management/Headquarters Coordinating Personnel #13
>
> "Well, initially when the employees went to training I guess there were a lot of questions on the procedures. If I remember correctly, the training split up where the MVFRs went at different times [to Headquarters]. And I guess at some point the training was different, or it was explained differently. Because I would heard complaints from the employees during the Wednesday morning staff meetings 'Well, we were told to do it this way' but others were told to do it a little bit different. So there was clarification that was needed once the actual program was implemented."
>
> Upper Management/Headquarters Coordinating Personnel #4
>
> "One area with the training, which I would define as having been negative: when the employees started going to training, they would come back and interact with each other on what they learned. Well, one group didn't get quite the information that another group got. So that was frustrating to me as a manager…because it was really confusing for my staff as well as the other attendees from other offices. Because we all wanted to know the same thing, to implement in the same manner. Not to be different from a neighboring office, or even from a neighborhood station with the office."
>
> Upper Management/Headquarters Coordinating Personnel #16

> "When we sent our people for training, they came [back] with five different ideas. So we had to streamline them into what we were actually doing instead of all of the different ideas that came out of training."
>
> Upper Management/Headquarters Coordinating Personnel #8

Managers, especially, were concerned about having to broker conflicting views among staff; to manage, in other words, confusion in their offices regarding what was expected of their employees. Given that training materials evidently differed from class session to class session, this meant not just that front-line staff had different information from each other. It meant also that managers (to some extent) had different information from their own staff regarding program requirements and procedures. This presented management level staff with substantial difficulties at the very beginning of the project.

Other staff had critical things to say regarding the consistency of training; however, this sentiment was not nearly as universal as among managers. Approximately a third of the MVFRs and SMVTs, but fewer still among other job categories (LREs, 3-Tier Manager Is, Administrative/Back-up Managers) noted that aspects of the 3-Tier process changed between training and implementation. When they did bring up changes or inconsistencies, it typically had to do with changes that were made to the paperwork required for each customer transaction:

> "When they set up [training], they said 'Okay, we're going to have this chart.' But they didn't have an example of the chart. Or with the forms; when the sheet finally came to our office—which was probably the fourth generation of it—it didn't match what we had in class a month prior. So from my perspective, for me and some of my co-workers, it was confusing…"
>
> MVFR#5
>
> "With training, well from the start I think you have to have the forms ready to go. That was one of the most confusing things for this project. With the forms revision, they would send a group of us to training. And then we'd come back with our knowledge. And then they'd send another group, but they'd revised the forms so that what the first group learned was now obsolete. So, first of all, before they start the training have the forms they way they're going to be."
>
> Administrative/Back-up Manager #3

Staff with direct responsibility for customer contact typically directed their critiques of training to the paperwork involved—its simplicity or complexity, how to fill out different forms, etc.:

> "Being an MVFR and wanting to do your job well, you want to really get a grasp of what's going on and because the forms initially weren't as… well, the training kind of scared me, honestly."
>
> MVFR #10
>
> "I still think that the Score Sheet was way too confusing. It could have had a better design. That would have made it easier. I left that training not having a clue other than that we were going to change the way we were doing more renewals."
>
> MVFR #9
>
> "Look, there were a bunch of errors that came up that couldn't be explained…the technicians kind of felt that the Score Sheet was redundant. Or they weren't sure what to check and what to circle. So they would get confused, and that's where I would have to come in."
>
> 3-Tier Manager I #5
>
> "I think that maybe the survey forms could go out in the notification to come in and renew their license. So that they're pre-filled out. That would be helpful, if you're going to continue the survey. Also the tracking sheet and the score sheet, to somehow be tightened up to be just one. As far as the training for the [MVFRs], it would just be helpful if they had less to do at the window."
>
> Administrative/Back-up Manager #1

The inconsistencies in training thus appear to have related primarily to the editing and the simplicity of the forms used for data collection—and not from confusion over the overall project purpose. As noted above (pp. 45-49), neither staff nor managers were in any way confused about the project's goals, and only a few respondents—all Hearing Officers—expressed any criticism of the project tied to this aspect.

Thus, to the degree that staff located their critiques in specific components of training, it had to do with filling out paperwork. These changes to the paperwork, and hence to the instructions in how to complete the forms for data collection, were largely the result of the fact that the initial training sessions were used as ad-hoc editing workshops:

"They were very, very complex and busy forms. And because of the amount of data that had to be collected, and the detail needed in collecting that data, it was hard to simplify the forms so that they were easily understandable for the user. And it took a lot of revisions. And so eventually by the time the pilot was implemented we came up with some very good forms. But it was a process where the people we were training were actually giving us input for any improvement of these forms while we were training."

Upper Management/Headquarters Coordinating Personnel #10

"Well, because it was the beginning of the process, training was….well, we were given misinformation. It wasn't training's fault; I know that when I was in training we changed the whole [process] to simplify it. So when I looked at it in the end, I thought it was pretty good; I thought the tracking sheet was pretty self explanatory for a technician or myself, or anybody really. The score sheet also; when I actually sat down and really looked at them, they were pretty self-explanatory—in the end, I mean, once we made those changes."

LRE #3

"The problem as I saw it was that they were trying to get [the pilot] out in such a short period of time. And so the training was not as concise or complete as it could have been. For instance, our work flow chart. When the Manager Is and the Backup [Managers] and the Admin[istrative manager]s went into training on what we had to do…we were the ones who went in and said 'This isn't going to work.' And we basically redid the flow chart. The problem with that was that everybody [else] had already been trained."

3-Tier Manager I #3

The revisions to the forms were regarded as absolutely necessary, especially inasmuch as simplification would reduce application processing time. The paperwork originally developed for the pilot was seen as unnecessarily complicated, redundant, and time-consuming. Consequently, those ultimately responsible for keeping the wait times down (i.e., managers) demanded, and received, substantial revisions to the paperwork.
That said, many respondents regarded the revisions to training that resulted from the editing of the forms as severely problematic. This was especially the case for management:

"I think the #1 curve in the road was that, from the training in March to the initial inception [of the program in May], so many changes had taken place in the program requirements. And we were never notified of those changes. Which was extremely frustrating; to not be notified of that. And then, when we *were* notified of the changes, it was after the fact…"

Upper Management/Headquarters Coordinating Personnel #2

"So there was a lot of confusion, especially among the Office Managers, Administrative Managers, and the Manager Is who had primary responsibility for overseeing implementation. There was a lot of disgruntlement early on from these folks. To the point that…I had a couple of Manager Is call me up and say 'Can I volunteer <u>not</u> to be on this pilot project anymore?' Because they were very concerned about being held responsible for something that they didn't understand and [about which] they didn't get any answers in training."

Upper Management/Headquarters Coordinating Personnel #1

The point was echoed most pointedly by Hearing Officers:

"When we first started with 3-Tier, the training was really bad. The training staff were not very informed, and I don't think that they were prepared to give us that training at the time. And then we went ahead and we implemented the project, and we didn't have the tools necessary to follow through as far as the booklets, the survey sheets…We were making copies from the samples that had been given to us. So in that regard I think of what could have been; that got us on a road of 'You know what? This is not something I want to do.' It was frustrating at the beginning, so it put you in that mode already where you're not happy with this program. I think if it was a little smoother it would have helped. Right from the start it would have helped everyone accept it more."

Hearing Office #2

"The first training that we received was [at DMV Headquarters]. We all just walked out of there going 'What the heck are they talking about? What are we doing?'"

Hearing Officer #1

In brief, the confusion and frustration that attended the editing of forms (and subsequent changes to program requirements, workflows, and instructions for procedures) produced a substantial amount of unease at participating in the pilot. This

unease expressed itself partly as confusion (as among front-line staff), but even more as frustration and anxiety regarding commitment to the project (among managers and Hearing Officers). This severely compromised what many managers had initially seen as a positive good: the coming together of DMV staff around a new and exciting vision for improving traffic safety (see above, pg 46).

In addition to critiques, respondents had suggestions for future projects. One of these involved moving training away from DMV Headquarters, where most newly-hire Field Office Region III employees receive the bulk of their training in subjects such as driver license (DL) renewal. Instead, many of those interviewed (especially, but not entirely, managers) believed that much of what was involved in the 3-Tier process could have been covered during training sessions on-site at individual field offices:

> "I think the training could have been done in the [field] office. The trainer could have come to the office, rather than us going downtown."
>
> MVFR #9
>
> "The training that was given…instead of pulling everyone out of field offices and into Headquarters, could have probably been done in a couple of Wednesday meetings. Our people are used to getting information; you tell them about it, they read it, and then they implement. We're sort of regimented that way. We're very used to the Wednesday morning training session."
>
> Upper Management/Headquarters Coordinating Personnel #5

In addition to holding the initial 3-Tier training on-site, a substantial number of those interviewed endorsed the idea of holding subsequent "refresher" training on 3-Tier procedures, also during Wednesday morning staff meetings:

> "If the project is six months long, maybe we should do a refresher training every two months, and get feedback from the MVFRs."
>
> 3-Tier Manager I #2
>
> "In our weekly meetings there was always a section or a time when we could get more information. Or if one person was having an issue in one area, we could come together as a group to discuss and devise a plan: 'This is how we're going to approach this.' And that made you feel more secure during the workweek, you know? 'This is the way I'm supposed to be doing it.'"

MVFR #10

Shifting the training to the field offices (rather than at CA DMV Headquarters) was seen as having two principal benefits. First, this would reduce the amount of perceived disruption to office productivity:

> "We do training of our employees constantly…we get training modules and we spend time every Wednesday morning from 8 to 9. And we have put into [production] some very complicated programs, where did major changes in, for instance, the commercial licensing [procedures]. That worked very well for Field in the past. But the concept of taking all the MVFRs into headquarters? That wasn't possible—I still needed to run my office. Because the customers waiting three hours aren't going to understand [why] 45% of my people are in a training class for something they don't want us to do anyway."
>
> Upper Management/Headquarters Coordinating Personnel #1

This idea was raised by only a few respondents—all of them with managerial responsibility for improving office production efficiency and reducing (or maintaining) customer wait times. Other respondents—again, most of them managers—raised the idea of a secondary benefit of in-office training:

> "I believe you can revamp it to train the managers of the office, and then let them train their technicians accordingly, to each office. The training itself was great; when I left the training I felt like I knew what we were doing. But for the technicians, I think we needed something more in-depth. So I believe it would have worked better if we had just used a Wednesday meeting, or whatever, to have our own training for each office. Because I'm sure every office had to adjust the training to their office."
>
> Administrative/Back-up Manager #2
>
> "Implementing follow-up training is probably one of the more important things that we could do. You know, send out a couple of different refresher training [packets] to all the offices. This is a really good way to ensure quality control, i.e., that everybody is doing things the same way, because you do have people who are taking different thoughts away from training when they're done."
>
> Upper Management/Headquarters Coordinating Personnel #10

By having the training "adjusted" to each individual office, those interviewed suggested that they could more adequately manage within-office variations in implementation. This would thus (presumably) ensure that everyone within a given office received the same set of instructions, and hence the work conducted in that office would be more uniform. Though this went unstated, it appears that having more direct control over training materials would allow managers to forestall or prevent many of the supervisorial headaches that they saw as stemming from the "inconsistent" training provided for the 3-Tier Pilot. Respondents left unstated two logical corollaries of this view: firstly, that *cross-office variation* was seen as unavoidable (at least to some extent). Secondly, respondents did not discuss how doing training at the field offices (as opposed to some centralized site) may affect variation *across job categories*, as seen in (for instance) respondents' understanding of the purpose of the PRT.

Regardless of where training took place (in the field offices or at DMV Headquarters) a substantial number of front-line staff—including half of the MVFRS and nearly all of the LREs—advocated for a particular pedagogical technique. This specific training method consisted of role-playing and hands-on practice sessions:

> "For me, the training helped immensely, to understand what we were doing, and why we were doing it. That's key. The role-play we in training helped immensely…I felt like we were prepared from the day we started."
>
> LRE #1
>
> "I remember when I was in DL training they had you practice on the orthorator.[24] You actually get to look at it. So maybe have two trainees: one be the tech and one be a 3-Tier applicant. Just to practice and fill out the form as if it was a 3-Tier customer. Just to get more comfortable with it."
>
> MVFR #4
>
> "Actually interaction with a customer would help a lot better. If there was any type of training where we could do like a mock version of a 3-Tier transaction. That would be a lot better than reading from the book or 'Here, read this and we'll explain it.' Actually seeing and interacting helps a lot more."
>
> MVFR #8

[24] This respondent was evidently referring to the Optec 1000 Vision Tester, a device for measuring a customer's visual acuity. It is a simplified version of what many ophthalmologists use in their clinical offices, adapted for use by trained non-health-care professionals working in a public agency setting.

> "This may be hard to do for a field office, perhaps. But maybe taking a few people at a time and actually simulating a 3-Tier transaction. So that you get comfortable with it."
>
> MVFR #1

This type of hands-on practice evidently worked well in the case of one new employee hired after implementation started and trained on the job:

> "Well, I sat with the Senior Motor Vehicle Technician, and she helped me. She went over each procedure and how to do each step of 3-Tier. And then the 3-Tier Manager I also helped me. It only took me a couple of days to get it really down pat. I'm a quick learner with this kind of stuff, so it wasn't all that difficult."
>
> MVFR #7

For MVFRs in particular, role-playing was seen as a way to reduce the anxiety stemming from implementation of a new, unknown, and potentially complicated set of procedures. This was especially true when it came to filling out paperwork.

Respondents from other job categories also saw role-playing and hands-on practice as a useful addition to the training pedagogy, though for slightly different reasons. Several Driver Safety staff, for instance, reported a desire to see what the new equipment looked like:

> "I would have liked to have seen what the Field Office part of 3-Tier was. I went downstairs and just kind of watched a few times as somebody went through. And I think, what do they call it, the fog chart? I think that was a great idea….but if they're talking about the PRT? I still have no idea what that is."
>
> Hearing Officer #3
>
> "I think it would have been nice too for us to have some hands-on with the equipment that they were using. That way we would know what it was that they were referring to in the field [offices]."
>
> Hearing Officer #2

Here, even though Hearing Officers were not required to (for instance) administer the Pelli-Robson chart or the PRT, they wanted to know more about them. Hearing Officers use multiple sources of information—including professional medical reports, interviews

with customers, and drive test results—in making determinations regarding licensure. Having practice with any new assessment tools would, in the view of those interviewed, better equip them to incorporate the results of those assessments into their hearing decisions.

Among some managers, but especially among LREs, role-playing served a secondary purpose beyond familiarizing employees with the mechanics of filling out forms and entering appropriate data into computer terminals. Here, role-playing, case-studies, and hands-on practice could teach new employees the various ways to provide sensitive, personalized customer service:

> "Maybe we need some additional training with the [MVFRs] in how to talk to people, to communicate, and make [the customers] feel at ease."
>
> LRE #1
>
> "Especially if this is going to be implemented in full…we need to train how to be sensitive with these people. Because some of these applicants were thinking that we were trying to discriminate against them because of their age, or their physical or mental conditions…"
>
> LRE #5
>
> "Is that driver fit to drive? That question is so fundamental, but it's also uncomfortable. And the training needs to make [the MVFRs] <u>not</u> feel uncomfortable if they have to bring it up. Because if you see someone in a scooter or a wheelchair, then you need to politely read that statement [from the application]: 'Have you had a condition? Have you had any hospital visits?'"
>
> LRE #4
>
> "I think if they started doing case studies, or role playing—I'll give you an example…You need to be sensitive as to how you would like to be treated if you forgot something or didn't quite hear right. Lowering your voice, for instance, maybe helps to calm a person. What kind of service do you want if at the bank? Or if you go into Nordstrom to turn in a pair of shoes but you really bought them at Mervyn's and they still give you full credit anyway—wouldn't you be like 'Omigod, now I'm embarrassed!' These are the kinds of things that can bring up at training."
>
> Upper Management/Headquarters Coordinating Personnel #8

Inasmuch as the provision of excellent customer service constitutes a key departmental goal, those interviewed saw a distinct need to ensure that 3-Tier—and the training necessary to institute the new pilot—furthered the achievement of that goal.

Given the critiques raised regarding training it is difficult—at least on the basis of these interviews—to explain why the implementation of some project components varied but others did not. To be sure, some pilot elements required close attention to detail in filling out the various forms: in the case of the Pelli-Robson contrast sensitivity chart, a customer might have 3 possible outcomes (pass, somewhat fail, or extremely fail), with three possible boxes for the technician to record a customer's "score." Both the size of the number (0, 1, or 2) and the placement of that number (lines 1, 4, or 5) determined whether that customer would be given a DL62 and required to obtain a professional eye exam, flagged for further assessment within the office in Tier 2 of the process, or simply asked to take the written law test without undergoing any further assessment within the 3-Tier system. On the basis of the evidence discussed here, there appears to have been little variation in the implementation or understanding of this assessment tool, despite the complexity of the paperwork associated with keeping tracking of customer outcomes.

On the other hand, some project components required very little in the way of paperwork: the memory recall test, for instance, admitted of two possible outcomes (fail or pass) with only one box in which to record the score (0 for pass, 1 for fail). Similarly, administration of the educational intervention required the checking of a single box on one form, in addition to sitting with a customer during the playing of a four-minute DVD. Per the frustration and confusion expressed by respondents regarding the uniformity of training and the complexity of the forms, we might expect that these two project components would have been implemented with relatively more uniformity. However, this was not the case. According to the evidence provided in these interviews, the complexity of paperwork (on the one hand) and the uniformity of implementation or understanding (on the other hand) do not appear to be correlated.

The implications of this point are somewhat subtle: given that frustration with particular aspects of training do not map on to those parts of the pilot where we find the most evidence of variation, it is not necessarily clear that revising training is the most straightforward answer to improving the uniformity of implementation. To be sure, the constructive suggestions made by those interviewed may improve particular

aspects of the 3-Tier process and, potentially, other DMV programs and procedures. It seems plausible, for instance, that role-playing and case studies are among a battery of pedagogical methods that can be used to improve the personalized nature of customer service in the field offices. However, neither simplification of the paperwork, nor (probably) role-playing is likely to alter the kinds of variation in staff understanding or implementation seen in the cognitive assessment tests (memory recall and PRT) or the administration of the educational intervention. Other factors, quite separate from training, appear to be at work as well.

GOAL TENSION: TIME VS. PERSONAL ATTENTION

In the surveys conducted at the end of the pilot project—see Module #1—the analysis of respondents' answers revealed a set of subtle tensions between two of CA DMV's core organizational goals: efficient processing of customer applications, and providing excellent customer service. These two goals also appeared in the staff interviews, though each was embedded in a slightly different context, and emphasized with varying weight depending on the job classification of the person interviewed. Perhaps more intriguingly, the surveys suggested that "customer service" admitted of different definitions that varied from respondent to respondent. The interviews substantiated and fleshed out this variation: for some (mostly, but not solely, managers), providing good customer service was synonymous with productive efficiency and low wait times. Thus, when asked what it would take to implement 3-Tier statewide, these respondents noted that successful implementation will require (a) shortening the 3-Tier process by incorporating it into the computerized, automated license renewal system (DMVA), and/or (b) hiring new personnel. Both of these suggestions would, in the view of those interviewed, reduce the amount of time customers spend in the field offices and so, according to this definition, improve customer service. For other respondents—especially, but not solely, MVFRs—the quality of service lay in the nature and amount of personalized attention that could be spent with any given individual customer. Staff noted significant tension between these two definitions; many stated explicitly that to the degree that 3-Tier allowed or encouraged personalized attention (a positive good), it also produced a decline in productive efficiency as measured by an increase in customer wait times (a distinctly negative outcome).

Time, as in time added to the license renewal process and (by extension) to the length of waiting required by customers on their office visits, was the single most common issue raised by those interviewed. Nearly everyone (94%, or 46/49) brought this up. All of those who did expressed substantial concern with how long the 3-Tier process took. In some cases, this concern involved defining customer service purely in terms of time spent in the field offices:

> "The concept of taking all of the MVFRs in these six offices into a headquarters setting [for training]? Well, first of all, we had to dicker back forth about how many training classes we needed in order to do that. Were we going to do all of them in 3 days? No, that wasn't possible. And there was—or there appeared to be—either a lack of understanding, or a lack of concern, [with] how this pilot took precedence over customer service. Because the customers waiting for three hours aren't going to understand why 45% of the staff are in a training class, for something [the customers] don't want us to do in the first place.
>
> Upper Management/Headquarters Coordinating Personnel #1

Here, wait times were impacted by one specific component of the pilot—the training needed to prepare field office staff in the new procedures. To the extent that training required pulling staff out of the offices and into DMV Headquarters—thus reducing staffing levels in the field offices—this was seen as negatively impacting customer service precisely because it raised the wait times. This same sentiment was echoed by other managers as well. When asked what impact 3-Tier had on customer service, they responded in terms of the impact on processing and wait times:

> "Because each transaction took so long, the customer service for the rest of the office [laughter] kind of would go down. We couldn't keep up with our driver license applications because each transaction took so long, so in my eyes our customer service was going down. And, I think, because the transaction itself was taking so long, you lost some of your customer service even during the transaction."
>
> Administrative/Back-up Manager #2
>
> "Well, one bad effect [on customer service] was the amount of time it took per customer for 3-Tier. And it was both a negative with that particular customer, and with the ones that we were waiting. So it increased our overall wait times for our driver license customers."

Administrative/Back-up Manager #3

In the view of these respondents, 3-Tier affected customer service in two ways: by lengthening each individual transaction, and secondly by increasing wait times even for customers who weren't participating in 3-Tier. In both cases, however, "good customer service" was synonymous with as short a visit to the DMV as possible. And so to the degree that 3-Tier lengthened the amount of time customers spent in the office, it meant "customer service was going down."

Managers, in particular, are charged with the ultimate responsibility for maintaining low customer wait times in the field offices. Thus, when asked what would be necessary to ensure the success of 3-Tier in the event of statewide implementation, nearly all managers spoke in the first instance about the necessity of improving production efficiency. For some, this would require making changes to facilitate faster processing:

> "I think we are definitely going to have to look at an automated solution. For instance, creating a form which could reside as possibly a web page, whereby the [MVFR] would be able to tab over to this web page through the processor at their counter, enter the information for data collection, and then submit that form. Because as the analysis goes forward it's going to be interesting to see if wait times in these six pilot offices really were affected by this additional workload of having to go through a manual process of data collection for the 3-Tier Pilot. And if it does, then it would just make it a smoother transition, and also help to indoctrinate our working staff to become even more familiar with automation and different I.T. [Information Technology] tools which we could then use for future projects as far as collecting information."
>
> Upper Management/Headquarters Coordinating Personnel #2

As piloted, the 3-Tier process involved a great deal of paperwork that needed to be filled in by hand ("manual processing"). Many respondents assumed—and, as this manager did, also emphasized—that any wider implementation should involve "automating" as many parts of the 3-Tier process as possible. Specifically, this manager envisioned incorporating new data fields into the driver license application that each MVFR enters into a computer on behalf of renewing customers. Thus an MVFR would enter additional codes into a customer's license renewal application for such Tier 1

process elements as physical observation and contrast sensitivity chart outcomes, just as they currently do for Snellen visual acuity chart outcomes. Having front-line counter staff enter data into a computer (as opposed to filling out paperwork) was regarded as an obvious, and essential, step to improving production efficiency.

Even more often than this, however, many managers noted that statewide implementation would likely require the hiring of additional staff:

> "The first challenge is obvious. It's the staffing and the time required. In my opinion, you would need staff that's dedicated to the 3-Tier mission. And while I don't know if it would have to be like a Driver Safety kind of procedure, I am saying that it was a very time-consuming process and we need more people."
>
> Upper Management/Headquarters Coordinating Personnel #14

> "If it is in fact implemented statewide, I think we really need to look at whether we have an adequate amount of staff to do the job. Because I think we calculated with…having to do all the extra paperwork, and with a limited amount of customers, we gave [our staff] a little bit of extra time. But if it's every single customer, then we really need to do a time study of, frankly, what is that 7 or 8 or 9 extra minutes going to do to [our] schedule?"
>
> Upper Management/Headquarters Coordinating Personnel #2

> "The challenge for the field offices will be trying to meet the quota for customer wait times. That would be the only problem that I see if we went for statewide implementation. You're simply not going to be able to keep the wait times at 65% at 20 minutes, if we have 3-Tier. It just takes a little bit longer to process than a regular driver license transaction. And so we have to have more personnel. We just need the personnel."
>
> Upper Management/Headquarters Coordinating Personnel #3

Those interviewed took seriously the degree to which hiring additional staff would impact DMV's budget—but this was regarded by many as vital to (a) the success of 3-Tier, and (b) the shortening of wait times that was seen as a key component of that success. Even front-line employees saw increased staffing as the most important step toward statewide implementation:

> "I really hope that the data collected will allow every office, regardless of volume, [to have] one or two more competent human beings. To allow the

> proper time to have less overtime, fewer customer complaints, fewer accidents, [etc.]. And I know there's state budget cuts, but they need to give us the funds. And I don't mind going on the record on this. Resources—give us the resources. Just like the army."
>
> MVFR #5

Inasmuch as providing good customer service is a core department goal, many of those interviewed saw the timely processing of customer applications as a (perhaps the) key component to providing good customer service. And whatever could be done to improve production efficiency—from eliminating paperwork in favor of computer-based data collection, to the hiring of new personnel—was seen as a positive step toward the improvement of customer service.

Among others—both managers and front-line staff—customer service tended to admit of a different definition, one grounded in personalized contact. Good customer service consisted, in this view, of individually-tailored assistance for a customer in accomplishing whatever they needed—whether it be a license renewed, a drive test conducted, or potential limitations evaluated. This view usually overlapped with concerns about time, and in fact personalized attention was often seen as a potential—though perhaps unavoidable—threat to perfectly efficient processing of customers. However, production efficiency was seen as theoretically separate from—even if related to—providing each customer with, individualized help:

> "You took the time with the customer, and the customer appreciates you taking the time with them and explaining what's going on. Because some of them were wondering why we were doing [this] and then you would have to explain to them what was going on. And that's…that's a positive way of giving good customer service: letting them know what you're doing and how they're doing. And just spending that time with them."
>
> MVFR #3
>
> "[Some] people were a little frustrated with us because in essence the transaction times were taking a little bit long because there were a couple of extra steps…[and] that two or three minutes could be kind of a wait for the customer…But the positive aspect, of course, is that you get more time to deal directly with the customer. You're asking them some direct questions—it's not

just 'Hello. Welcome to the DMV, how can I help you?' You've got to ask them questions, like 'Do you mind filling out this questionnaire for me while I process your transaction? And if you have any questions about the survey, please ask me.' You get to give them a little bit more information. And you find out a little bit more about your customer because you're required to pay attention to how they are writing: is their hand shaking or not?"

Administrative/Back-up Manager #5

"Of course, time is always a factor here. We try to service everybody as fast as possible, but I would say that with 3-Tier…well, I don't want to say it tested my patience, but it taught me to be a little more patient. And when you're patient you're able to see things and to look at things a little bit more closely. And because I became more patient, I was probably a little bit more polite."

MVFR #10

"I don't think the extra testing changed the customer service aspect. I think they got the same quality customer service they would have gotten without 3-Tier. It was just a little more time-consuming for them and for us."

Administrative/Back-up Manager #1

Though sensitive to the amount of time 3-Tier added to processing and customer in-office waiting, this group of respondents—which constituted a majority of both front-line staff and lower-level managers—saw 3-Tier as providing distinctly valuable opportunities to improve the quality of customer contact. This quality was measured in terms of observation, conversation, and the collection of otherwise-lost data on driver competency. Some respondents even noted that the extra time spent, and the consequent increase in the quality of customer service, produced an improvement in license renewal outcomes:

"I liked being able to have the freedom to take the extra time with the customer. So I didn't feel rushed. I took the time, because I was reassured by management that I could take the time to explain to the customer what to expect. And a lot of times we have a better result when we take that time. Because they tend to relax when they know what [examiners] are looking for. It makes a difference, I think."

LRE #1

In the observation of this examiner, spending more time with each customer tended to produce higher passage rates on drive tests, mostly because extra time meant more opportunities to answer customer questions and let them know what to expect from their on-the-road drive test.

3-Tier's impact on production efficiency was thus a common, and urgent, concern. However, it is worth noting that the amount of time that 3-Tier took does not necessarily illuminate why staff reported variation in the implementation of various process elements. To the extent that respondents identified particular sub-components as taking excessive amounts of time, they mentioned (a) observing the customer for physical limitations, especially while walking with a customer to take their picture at the video capture station, (b) manually filling out paperwork, and (c) in the case of 3-Tier Manager Is, the electronic scanning of documents for the purposes of recording-keeping. In none of these cases do we see substantial evidence of variation produced by, for instance, the cutting of procedural corners to save time. If anything, the amount of time taken to observe customers for potential physical limitations was seen as a positive good (see above, pp. 56-57), even one that respondents hoped would continue after the pilot had ended.[25] Instead, those areas that were identified as varying in implementation or understanding—the cognitive flags in particular—either took almost no time at all even when done carefully and correctly (the memory recall exercise) or had a time duration pre-set by computer program, and hence were essentially unalterable (the PRT). Furthermore, in the case of the educational intervention videos, much of the variation in implementation came in the form of giving the video *even when it wasn't required*—which of course took extra staff time.

[25] There exist other sources for any potential increase in processing time that resulted from 3-Tier. In particular, the SDPE takes, on average, substantially longer to administer than a "regular" DPE drive. This is not simply time that the LRE could spend conducting other drive tests; it is also time that they could spend processing customers "at the window." To the extent that 3-Tier resulted in an increase in the number of SDPE drive tests given in the field offices, this likely contributed to any real or perceived decline in production efficiency. However, no respondents mentioned this in the interviews as a threat to reducing customer wait times in the field offices.

GOAL TENSION: DISCRIMINATION VS. TRAFFIC SAFETY

In addition to efficient processing and personalized service, a number of respondents defined customer service in terms of *universality*—providing each customer with essentially the same experience during their visit to the DMV. This goal was revealed most often in comments regarding perceived or actual instances of discrimination (whether by age, language, or unspecified "targeting" and "profiling"). Here, respondents noted tension—sometimes explicitly, sometimes only implicitly—between on the one hand the potential for unfair discrimination and on the other hand CA DMV's organizational goal of improving traffic safety through the licensing of competent and skilled drivers. In other words, while many respondents were excited to be a part of a program that would make California's roads safer by identifying drivers with potential limitations to their ability to drive, staff were also concerned that the methods by which 3-Tier identified such drivers were possibly unfair, and perhaps fell more heavily upon particular groups. That said, at least a few also noted that they deflected customer complaints of discrimination precisely by appealing to the 3-Tier's potential for improving traffic safety.

A substantial number—approximately half—of respondents raised concerns regarding the universality of the 3-Tier. At the most basic level, staff and managers quite explicitly held the view that any driver competency assessment program, such as 3-Tier, should apply to all drivers. This appeared most forcefully in comments related to language. Because of budget and time constraints, enrollment in 3-Tier was limited to drivers electing to take the written renewal test in English.[26] Although only mentioned by a few of those interviewed (8/49), those who did so saw this limitation as grossly unfair:

> "If they continue the program, will it before everybody that renews their license, rather than just English-speaking? Because when you've got customers who have points or tickets on their record, but they want to take the [written] test in Spanish, it didn't seem very fair that [3-Tier] was just for English-speaking customers."
>
> SMVT #1

[26] Depending on the office location, this limit on language potentially excluded anywhere between 3% and 22% of customers. A more formal analysis of this limit on the estimated impact on office production efficiency is included in the main body of the process analysis.

Implicit in this critique was the view that language was unrelated to driver competency—and hence limiting enrollment in the program to one particular language meant not evaluating some drivers whose competency should have been examined. Respondents therefore viewed this limit on language as not just unfair, but also as having a negative impact on DMV's ability to assess driver competency:

> "If this goes it, it will go into all languages, right? [laughter] Then I foresee it being quite a bit more than what it was in [English]. It will be a good thing—because we want everybody safe on the road."
>
> 3-Tier Manager I #1
>
> "I feel that if we're only going to do English-speaking [customers]… we miss the opportunity to evaluate a lot of people that we may need to evaluate. So I think we should expand it to anybody taking a written test, regardless of language."
>
> Administrative/Back-up manager #5

Obviously, as required by California state law and CA DMV departmental practice, should any elements of 3-Tier be adopted in the future, these elements will be made available in whatever languages are deemed necessary at that time.[27]

Those interviewed argued not just that 3-Tier *as a whole* should apply to all customers, but also that any specific method by which CA DMV assesses driver competency should apply, in a basic sense, "equally" to all drivers. Thus, the emphasis on universality was seen as applying in two senses: all customers should have to undergo 3-Tier, but also any given process element should not "discriminate" against a particular group. This second meaning of universality came up in the context of two process elements specifically: the memory recall exercise, and the PRT. In both cases, staff and managers alike noted that, in their observation, it was senior citizens who appeared to have "trouble" with the tests:

[27] In accordance with the requirements of the Dymally-Alatorre Bilingual Services Act (CA Government Code Sections 7290-7299.8), CA DMV is fully committed to provided equal access to departmental programs and services to all persons, including those who are Limited English Proficient or non-English speaking. The department constantly assesses customer needs in this area through the regular use of periodic surveys of field office contacts in languages other than English. These surveys help shape departmental policies regarding new translations of tests and educational materials, as well as in the hiring of bilingual staff.

"There are a group of people—mainly seniors—that really do not know their Social Security Number. Think about it. When they apply for credit cards and phone services and electricity, gas, water, they never had to use their Social Security Number. They just went down and said 'I'm so and so, blah, blah,' you know? So to ask these people now to…rewrite their Social Security Number from memory, I thought was sort of unfair from the beginning. I think there was another number that we could have used that would have been a better source to check their memory."

Upper Management/Headquarters Coordinating Personnel #9

"Well, you got some senior citizens. I think some people just don't have their Social Security Number memorized."

SMVT #1

While seniors were assumed to be at an unfair disadvantage when asked to state their Social Security number from memory, staff also asserted that seniors were less willing to take, or less skilled at taking, any test that involved a computer—such as the PRT:

"I thought, gee, I want to know how many of our seniors have gotten in front of a computer before. I mean, I would think they would be intimidated by it, and that alone would…I don't know if they were on a time limit or not, but that alone would probably cause them some delays in the process, I expect. And…well, I just can't believe that they're doing that."

Hearing Officer #2

"I did quite a few people with the PRT. And that was kind of funny. Because a lot of the older generation—not all of them, but a lot of them, particularly those that don't embrace a lot of the newer technology—they had a little harder time learning to use the PRT, in figuring out what they needed to do."

LRE #2

These respondents argued that if the memory recall test and the PRT appeared to flag members of a particular group—in this case senior citizens—more often than members of other groups, then this indicated a problem *with the test*, rather than indicating the distribution of driving-relevant functional limitations within the population. In other words, at least some staff viewed these two process elements as unfairly discriminating against senior citizens.

This critique appears to stem primarily—but not entirely—from skepticism regarding the utility of these two tests. Only rarely did respondents report concerns about universality and/or discrimination with regard to non-cognitive assessment tool—vision, physical observation, or even the on-road drive test. Thus, even though the types of vision disorders most often identified by the Pelli-Robson contrast sensitivity chart—glaucoma, diabetic retinopathy, cataracts—are more common among the elderly, in only one case did a respondent note that the Pelli-Robson chart unfairly discriminated against senior citizens. Similarly, even though driving-relevant physical limitations are also more common among the elderly, staff did not mention age discrimination when discussing the physical observation protocol. This suggests that the concern that particular tests discriminated against the elderly were, at least to some extent, grounded in a more basic skepticism of the validity of cognitive health assessment tools.

Concerns regarding universality also appeared in the form of second-hand reports of age discrimination. Not only did (some) staff believe that specific assessments unfairly "picked on" or "targeted" the elderly, but they evidently had to manage those complaints as expressed by customers. This was a common sentiment, mentioned by respondents in all positions:

> "We've had older people say 'Well, you're just picking on me, and I would have to say No, we're not! We get younger people in here too, and sometimes the older folks do better than them…'"
>
> MVFR #2
>
> "We had a lot of customers that thought we were targeting the older age brackets. So this was a very sensitive position, and you had to be a very sensitive examiner to deal with this…"
>
> LRE #3
>
> "If they [went to Tier 2 or 3], then they complain, saying 'DMV is doing this to me because I'm old. I'm a senior citizen and that's why I'm in this program. If I were young, you guys would not do this to me.' They believe this was only to catch senior citizens."
>
> 3-Tier Manager I #4
>
> "The question 'Why is this need now? Is this because I'm older?' That was the main question: 'Is this because I'm older?'"
>
> Administrative/Back-up Manager #2

"Well, [there] is the public perception of being picked on in regards to seniors—that these new assessments were developed to target them."

Upper Management/Headquarters Coordinating Personnel #11

Note that these comments cannot be used as a direct measure of the scope of customer sentiment on this issue (see Module #3, which analyzes data taken from a survey of 3-Tier customers). However, they do indicate that staff had to manage some unknown quantity of customer complaints.

From other comments it appears that a few staff may have shared these views, and wondered themselves whether or not 3-Tier discriminated against the elderly:

"When you look at the people who have gone through 3-Tier all the way through to the drives? I would venture to guess—and it's just a guess on my part, having walked to people, having watched my examiner, having seen the customers come in—as much as we don't want this to be about age? It is. In my opinion. And I'm not saying we shouldn't do it. I'm absolutely not saying we shouldn't do it, but what I'm saying is maybe we need to have a different approach to how we do it. Maybe. I know that AARP were on board with this, but can't we call a chicken a chicken? As in we tell people 'If you're over a certain age, and you do these things, you're going to do a drive [test] for us.' Because I've had one person—one—who had to take a drive test [who] was under the age of 50. And that was only because he didn't think he needed to read the book and so he failed his [written] test three times."

3-Tier Manager I #3

"I can't help but think that this program—and I know that's not the intention here—but I couldn't help but think that it was almost like we were setting up our seniors to fail. Like we were putting an obstacle in front of them in the hope that would fail, by doing that fog chart."

Hearing Officer #2

"The other thing is that with the physical and mental health of the customer, especially when a customer comes up to me shaking, I'll ask them about it. 'Have you had a stroke in the last 5 or 10 years? Have you had a heart attack in the last 5 or 10 years? Have you had any brain tumors in the past 5 or 10 years?' And I know that's not right, and maybe a little discriminatory, but I'd rather they tell

> me now then leave here, drive their car off our lot and have a stroke in the middle of the street."
>
> MVFR #6

This notion—which was rare—indicates that at least in the view of some staff, specific elements of the 3-Tier process were in fact age-biased. That while functional limitations in visual, cognitive, or even physical health may be correlated with age, this was self-evidently reason to be skeptical of any test that might "catch" senior citizens more often than younger DMV customers. And, therefore, that any kind of testing that appeared to be correlated with age—that is, in a word, non-universal, "targeted" or "discriminatory"—presented a dilemma that somehow needed to be addressed.

Staff managed this dilemma in two ways: by appealing to potential improvements to traffic safety (and especially to the possibility of improving a driver's own individual skill/competency), and by pointing out that the process was, in fact universal. So, when faced with complaints by customers about perceived discrimination, a number of respondents found it useful to explain 3-Tier in terms of preventing future crashes:

> "Once I would talk to them, and even when I didn't have to give them the educational materials, with some of them I did so so that they could really understand why they were being tested in this matter. And once they understood it, they were okay with it. Out of all these months, I probably had one or two that were really against it."
>
> LRE #3
>
> "Well, at first they were a little shocked. We had a lot of customers think that we were just picking on them because of their age or whatever. But then I [would explain] that it wasn't just their age, that it had to do with their vision, or maybe or some other problem that was causing them to be a little slower to react... They would get a little nervous sometimes but then they we would calm them down and let them know what we were doing, that we were just identifying those problems so that they would be aware of the problem. That we want to make sure that they're safe out there and that they're going to be the best driver they can be. And that calmed them down, when they knew that we weren't out to get them."
>
> Upper Management/Headquarters Coordinating Personnel #14

> "One we informed them as to what the goal was, and how they fit into that goal—that we were in fact trying to aid them, to help them—then they were more receptive to taking the time needed to complete everything."
>
> LRE #2

Others pointed out that everyone, not just seniors, had to participate in 3-Tier:

> "I had a lot of my elderly customers say 'Well, they're just doing this to us' because they didn't see that we're actually testing everybody that has to come in for a written renewal test."
>
> Administrative/Back-up Manager #5
>
> "For every person that I did the fog chart with, or sent out on a Drive test, if they were over the age of 75 it was 'This is because I'm old, isn't it?' And my answer to them was always and consistently 'No. If I had a 26 year-old who couldn't read that fog chart, we'd be sitting here doing the same thing.'"
>
> 3-Tier Manager I #3
>
> "Some were pretty upset…and felt right away that this was an age issue: 'You're picking on people our age' and I said 'No, no. That's far from the truth,' I told them. And I would explain to them '*Anybody that comes in here doing a renewal is going to take these test*.'"
>
> 3-Tier Manager I #5

There were thus two methods by which staff managed customer complaints—and perhaps their own concerns—regarding the perceived non-universal, or discriminatory, nature of 3-Tier. The first involved explaining how 3-Tier was tied to traffic safety, and specifically how the program was designed to help drivers improve their own driving skills. The key here appears to have been that 3-Tier both identified potential driving-relevant functional limitations and incorporated education and referral components designed to teach drivers how to compensate for those limitations. It was this second step (education and referral) that made the first step (the identification of potential limitations) acceptable. The other method for answering the charge of discrimination was, simply, to point out that the charge was untrue: in other words, to say that it *was* universal, that *all* renewing customers—not just senior citizens—were required to participate in 3-Tier. To the degree that this answer could work, it needed of course *to be true*; this may be in part the source of the concerns raised regarding the restriction on language. It is perhaps worth noting here that while staff from all levels raised the issue

of discrimination, in the interviews it appears that only representatives from a limited range of job categories who articulated answers to this issue. Specifically, it was those staff who directly assessed customer driving skills—LREs—or who engaged in at-length discussions with customers regarding their driving habits—3-Tier Managers and Administrative/Back-up Managers—who proffered methods to counter charges of discrimination. This likely is a result of on-the-job experience in having to answer such questions regularly.

GOAL TENSION: TESTING AS HURDLE VS. TESTING FOR COMPETENCY

A third source of tension identified in the interviews involved how staff regarded the purpose of the assessment testing. In brief, there appear to be two somewhat different ways of thinking about the purpose of the assessments DMV requires for licensing: firstly, as a series of bars or hurdles each customer must overcome before achieving (or renewing) the privilege to drive ("testing as hurdle"). This view was perhaps best expressed in comments such as the following:

> "You're kind of sitting there and you're hoping everyone gets everything straight the first time through, you know? You just hope it goes smoothly."
>
> MVFR #8
>
> "Ultimately the goal is to get people to drive. To get them to drive safely, of course, but the reality is to get them to pass the test."
>
> Upper Management/Headquarters Coordinating Personnel #7

Secondly, assessment testing was seen as a set of tools to identify potential limitations in competency and skill ("testing for competency"). These views often overlapped, and in the interviews many respondents talked in terms of both, switching between the two views depending on the assessment tool they were discussing at the moment. As a general matter, respondents were most likely to discuss 3-Tier *as a whole* using language that evoked testing as a method of assessing driver competency; however, when discussing the memory recall test many (though by no means all) spoke of this process element as a kind of hurdle that everyone should have been able to pass. In addition, in practice it appears that some staff administered the PRT as if it were a hurdle, rather than as assessment of potential limitations in cognitive health. On the other hand, when discussing the visual acuity or contrast sensitivity charts, most respondents spoke of

these elements in terms that signaled their use in flagging potential visual limitations—i.e., testing for competency.[28]

When asked about 3-Tier *as a whole*, informants noted the importance to the project of evaluating the efficacy of new driver competency tools:

> "[3-Tier was] to determine if the extra testing will reduce crashes. So, if those applicants that were in 3-Tier were more likely to crash or be in violation of the law."
>
> MVFR #4
>
> "Basically, gather information as far as how to…see if there's anything going on that might make them incapable of driving."
>
> MVFR #7
>
> "It's a complete re-evaluation of how we handle the driver license process. Making sure that we fulfill our goal of keeping safe drivers on the streets and also removing those that are potentially unsafe. It's a closer look at their driving skills, with a little bit more detail than what we do no. Instead of just having them take the written test, we take a closer look at they are physically, how they are mentally. Do they still have the memory retention to be able to drive safely? Do they still have the physical skills to drive a motor vehicle?"
>
> Administrative/Back-up Manager #5

In some ways, the language used here repeated itself almost by rote across respondents. This likely echoed the materials presented during training. What was really striking in these comments was the emphasis on using new assessments to test driver competency—to use tools to "see if there's anything going on that might make them incapable of driving." On what might, for lack of a better phrase, be called a *theoretical* level, the majority of respondents saw the value of 3-Tier precisely in terms of the

[28] These two ways of discussing the nature of assessment testing appear to some degree to be correlated with job category: MVFRs were not as likely to use testing-for-competency language as were Managers and LREs. However, this may be a product of the interview protocol more than anything else. As noted in the main body of the test, testing-for-competency views came up when discussing overall project goals (a question asked of all respondents); on the other hand, testing-as-hurdle tended to come up when discussing the memory recall task (a question asked only of front-line MVFRs, SMVTs, and those LREs who had done counter-transactions in addition to drives). In other words, upper management weren't asked directly about the memory recall test or the PRT, though some chose to discuss these assessments in the context of their answers to other questions.

project's goal of refining testing tools for flagging potentially limited drivers. In other words, when framed in general terms, staff spoke of testing as a means for assessing competency.

This was also true for specific project components, particularly the contrast sensitivity chart. With this element, staff saw obvious linkages between the assessment and driver competency, in part because the linkage was regularly confirmed through both the referral process and by customer feedback:

> "I love the fog chart. I absolutely love it. …I've heard nothing but positive about it. Those people that did not pass it—I mean out and out did not pass it?—they all had vision diseases that needed to be treated."
>
> 3-Tier Manager I #3
>
> "The fog chart, I loved it! I think that that worked [well], because like I said, we noticed that the people that weren't passing on the fog chart, that were having trouble, when they came back with their DL62, they had vision problems. Their doctor had noted, if not one, then several things on there."
>
> MVFR #1
>
> "Once they knew if was to help them to see better with fog/smog, and that it was really for their safety and everybody else's, that's when I would get your more mature clientele stating 'You know, I don't like driving in the fog anymore. Thanks for checking that, 'cause I'm color blind too.'"
>
> MVFR #5

That contrast sensitivity was "obviously" linked to driving ability was seen as a given, and therefore the test were seen as useful in its ability to determine potential limitations to competency. In other words, to the degree that the contrast sensitivity chart "discriminated" between different types of drivers, it did so in ways that staff and managers saw as clearly and directly tied to driver competency and skill.

By comparison, fewer respondents discussed the memory recall test in terms of its ability to flag potential limitations:

> "The people who couldn't remember it [their SSN] definitely had something wrong with their memory. Or they just didn't seem like they were able to

comprehend much. So it was good in that it pulled them apart from the rest of the group."

MVFR #8

"Even the memory test for if they could remember their [SSN]. That's good, because it's going to give us a tell-tale sign of any disability they have on the memory loss."

Upper Management/Headquarters Coordinating Personnel #3

Though these two individuals saw the memory recall test as a valuable assessment tool for gauging potential cognitive limitations, they were in the minority. Most staff heavily critiqued this test. But they did so in evocative terms; some held that the problem with the memory recall exercise was that some people couldn't pass it:

"To me, it was a little bit inaccurate because some people were like 'Uhhh, I don't know. I never memorized it.' Even though we'd let them know at the Start Here window 'You're going to be asked' they were still like 'Oh, but I've never memorized it.'"

MVFR #1

Note that the implication of this view is that if an applicant "never memorized" their social security number, it didn't matter if this was a potential flag for cognitive limitations—what mattered is that the test was, in the view of those administering it, discriminatory. The flip side of this view was expressed by others, who held that the test "worked" precisely because everybody passed:

"In my experience, it worked pretty well. I only had maybe three that couldn't recall their [SSN] from memory. All the rest were able to do it without a problem."

MVFR #7

"I thought it worked. Actually, of the applicants I had, I remember only having one that had to look at their documents to remember their [SSN]. Maybe you could use the Driver License number [instead], but most people wouldn't know it [laughter]."

MVFR #4

To the degree that the nearly all of their customers passed, these MVFRs viewed the memory recall test as "working" because it was self-evidently universal; if no-one failed, no-one was discriminated against. On the other hand, if someone did fail this test, it was for the wrong reasons: the test "discriminated" by flagging customers for reasons that were, in the view of most of those interviewed, unrelated to driver competency.

In a sense, then, at least some of the tension between testing-as-hurdle and testing-as-competency is really a version of the tension between universality and discrimination in a secondary form. This tension had two-dimensions: firstly, there was a widespread skepticism regarding the validity of the two cognitive assessments, the PRT and the memory recall exercise. Secondly, and likely as a result of this underlying skepticism, the "solution" to this problem lay in either trying to get all customers to pass a test, or (which amounted to the same thing), dismissing the results of failure as a problem with the test, not with the customer. These "solutions" appeared among both front-line staff and Managers:

> "With the MVFRs, I could see that they kept wanting them to pass it [the PRT] because they have it in their minds that 'That's just a confusing machine.' But it wasn't. And because I've been an examiner for so long, and I've done every drive within the department, I could see that some people were not passing because they had [cognitive] issues."
>
> LRE #3
>
> "I told her [the Office Manager] about the failure rate on the PRT. I said 'The failure rate is not acceptable.' And that I didn't really think the customers understood it."
>
> 3-Tier Manager I #5

Those who administered the PRT certainly found it confusing: as noted above (pp 68-73), there was widespread misunderstanding regarding what it actually tested. Staff also had to manage customer questions regarding both the PRT and the memory recall test; one method of managing these questions appears to have been to administer the assessments as many times as it took for a customer to pass—thus turning a flag for potential limitations in competency into a relatively simple hurdle that anyone could (or should) be able to pass, and eliminating the need for higher-level assessment. It

bears emphasizing that while more highly-trained staff—such as LREs and 3-Tier Managers Is—appear to have been somewhat more likely to discuss the PRT in terms of testing-for-competency, a substantial portion of the staff in these categories also thought that the PRT measured reaction time. Thus, although it may have been valuable as a testing tool, what it tested was misunderstood even by those who found it valuable. Moreover, very few respondents—even among the more highly trained job categories—discussed the memory recall test as a tool for assessing competency.

Changes to Intra-Office and Inter-Division Cooperation and Coordination

Towards the end of each interview, staff and managers were asked a series of questions regarding how 3-Tier did (or did not) change how DMV conducts its work generally. This inquiry arose from the fact that the pilot involved significant changes to the Driver License renewal system, which constitute a significant proportion of the work conducted by the DMV. Depending on job category, respondents were asked (a) what kinds of differences 3-Tier produced to workflows *within* a given office (both between job categories of equal rank as well as between job categories of differing rank), and then (b) what kinds of differences 3-Tier produced to coordination and cooperation *between* offices. This latter question was then broken down to coordination between the field offices and R&D, between Driver Safety and R&D, and between the field offices and Driver Safety.

INTRA-OFFICE CHANGES TO COOPERATION AND COORDINATION

Within the field offices, 3-Tier appears to have encouraged somewhat more communication between front-line employees in the field offices (MVFRs and LREs) and between front-line employees and their direct supervisors (Manager Is). There appears to have been little or no effect on communication between front-line employees and Office Managers, or between Manager Is and Office Managers. Similarly, there does not appear to have been any alteration to the pattern of intra-office communication within the Driver Safety branch. Any changes to intra-office coordination and supervision occurred primarily as a result of the need to manage workload and productivity demands, both because the process was new (and therefore somewhat confusing, especially at the beginning) but more especially as 3-Tier transactions were seen as taking longer than "regular" driver license renewals.

On the one hand, MVFRs often found it necessary to ask for assistance from each other simply to get the new procedures done correctly at the beginning of the project:

> "Initially I felt that some us, well, we were a little bit confused. So we all kind of had to ask each other, you know? 'Is this the way it goes?' And we kind of depended on one another. And I think that helped because it gave a stronger sense of communication between us. To go over what we had learned and what we needed to review in order to make sure that we were all doing it right."
>
> MVFR #1

The increase in coordination continued even after the beginning stages of the project. As noted by a number of respondents, communication between MVFRs occurred in cases where they expected a transaction to take more than an ordinary amount of time to process:

> "By saying 'Hey, I've got a 3-Tier, it's going to take me a minute' to the people that you're sitting next to, they know it's going to take you a little bit more time. So say if they had a customer come back to them, but in the meantime they had called a 3-Tier customer? You would take their other customer because they might be awhile."
>
> MVFR #3
>
> "We would use the phrase 'Hey, I have a 3-Tier.' And what that meant was…to help your fellow team-player or co-worker out, to let the next customer that's sent to the counter know that it'll be a minute."
>
> MVFR #5

Not only did front-line staff find themselves communicating more often with those who sat next to them, but (at least in large offices), they found more reason to get to know people that sat further away:

> "You end up walking to other parts of the office more frequently, so yes, we were able to interact with our co-workers more frequently."
>
> MVFR #10

As part of the physical observation protocol, MVFRs had to walk with customers to the window where their pictures were captured for their new licenses. Particularly in large offices, this could mean walking a significant distance, and therefore communicating with fellow staff whom one might otherwise see only rarely.

This same dynamic occurred for communication between MVFRs and employees in other job categories. For instance, the LREs who would be using the information collected by the MVFRs in their evaluation of a driver's skill:

> "Actually, I never really had too much interaction with any of the LREs before 3-Tier. It was good because you have to know more about what they do, and they were kind enough to share what they knew—which was more than what we did."
>
> MVFR #8

MVFRs also found it necessary to communicate regularly with their immediate supervisors, who held responsibility for quality control of the paperwork produced as part of the pilot:

> "I always ask. I'm the type of person that I'm not afraid to ask for help. If there's something that I don't understand, or that I need to have explained to me, of if they can come and show me and just be there for support? That's a great thing. There was a lot of support from all of our managers on the 3-Tier."
>
> MVFR #3
>
> "It changed in the sense that we had to, sometimes, in our Wednesday morning meetings reiterate things that they were forgetting. So that was more educational."
>
> Administrative/Back-up Manager #1

The increase in coordination within the field offices was generally seen as a good thing by those interviewed, though the reasons for this—an initially confusing change to procedures, the threat of increased processing times—were regarded as avoidable and regrettable.

This increase in within-office communication appears to be confined to the field offices. None of the informants from Driver Safety reported significant changes to reporting or

coordination either between Hearing Officers or between Hearing Officers and the Driver Safety manager:

> "Here we're basically on our own standards and our own set time. And we're in the [cubicles] or we're in the Hearing Rooms. So if it's not break, or lunch, or you're at the fax or the phone, you really don't have a lot of conversing time. We just have our 30-minute lunches where we try to talk about happier times [laughter]…"
>
> Hearing Officer #4

Hearing Officers did report increased pressure on production efficiency because of the lengthened process, but this did not have the same effect as it did in the field. As Hearing Officers work largely independently, there is normally little interaction or communication between them; this did not change as a result of 3-Tier.

INTER-DIVISION CHANGES TO COMMUNICATION AND COORDINATION

In terms of inter-division coordination, there were a number of outcomes that respondents highlighted during the course of the interviews. At the highest level, informants noted that they saw cross-divisional communication as a positive good on its own terms:

> "I went to the initial meetings for 3-Tier, when [R&D] made presentations to all of the people that were going to be involved with it. And that was a real positive meeting; I thought it was really helpful. Because sometimes as a participant there isn't an initial meeting that brings all the players together. So that was very positive…that everyone involved had one focus. Which was to make sure that this pilot was successful."
>
> Upper Management/Headquarters Coordinating Personnel #16

Thus, the pilot brought into contact disparate parts of a large organization that otherwise might not regularly engage in communication:

> "I've been in Field Office for [a number of] years, and I had never, ever before seen anyone from Research and Development. I didn't even know that you guys

> existed, or even really what you did…I didn't know prior to [the project] what all was spent on preparing for a pilot, so that was a good thing. Because most field office staff don't even know who R&D are. But now they do, and that's a good thing. They know another part of DMV."
>
> Upper Management/Headquarters Coordinating Personnel #9

In an agency as large as CA DMV (currently the department employs approximately 9,000 full-time staff), it is entirely possible for members of one division to have only the vaguest of ideas what members of another division do. Because 3-Tier involved multiple divisions and branches within DMV—Field Office, Driver Safety, and Research and Development, among others—the coordination of the project required multiple meetings and, ultimately, the designation of specific liaison staff.

The liaison staff were, according to those interviewed, typically identified as R&D staff with responsibilities for communication with the field offices. Especially at the beginning of the project, they were responsible for overseeing quality control of the output from the pilot's new procedures. Staff within the field offices (and even the FOD Region III administration) specifically singled out these liaison staff for commendation:

> "I want to tell you, I think we lucked into [the liaison staff]. I don't know whose decision it was—I didn't have input into that. But once it was decided to put them on this pilot as the coordinating point between FOD staff, the field offices, and R&D, quite a difference was made there."
>
> Upper Management/Headquarters Coordinating Personnel #1
>
> "I thought it was wonderful that the people that were in charge of 3-Tier would come down periodically from Sacramento and check on how things were developing here. To find out if we had any questions or concerns regarding the program. I felt like they were keeping in touch and so understood that we were working hard to try to develop the program and follow it by the guidelines."
>
> LRE #2
>
> "I think it was really great when the R&D people could be here and fill us in on 'This is another way you can approach this. Maybe you could incorporate this?' They gave us other ideas within the pilot to make it flow smoothly. And it was also helpful to be able to call people that had the expertise on the training…That interaction was really helpful."
>
> Administrative/Back-up Manager #1

Across the board—but especially among Office Managers—it was deemed extraordinarily helpful to have specific individuals who could help solve problems as they were identified. Thus, the liaisons served not only as channels of communication; they also (perhaps more importantly) acted as designated "ombudspersons" who could answer questions, make suggestions, and solve operational problems as they occurred.

The importance of these liaisons to the project stemmed from two qualities: first, when asked what in particular made these liaison staff so valuable, most respondents mentioned (among other things) their many years' experience with DMV field office procedures:

> "It really has to be somebody that has been out in Field; [who] understands how Field works. And therefore, they're not going to make assumptions... But also, they know what we're doing, without us having to explain every step. Yeah, that is absolutely important. And we know [laughter]. We know the minute they walk in the door."
>
> LRE #2
>
> "Their titles didn't matter. They knew their topic—that's the important part. They were available and they were easy to talk to because they knew our lingo. They understood what we were saying when we used it, and what we meant when we said 'You may want to do it this way, not that way.' They had that Field background. It was a Field disposition."
>
> Upper Management/Headquarters Coordinating Personnel #12
>
> "Field people want to talk to Field people. Okay? Break it down for me. I don't need the project draft from two years ago on how R&D got their money…What I need to know is how this is going to affect my office, what my people need to know."
>
> Upper Management/Headquarters Coordinating Personnel #13

The three liaison staff employed during the pilot were only temporarily assigned to R&D: two of them worked (at the time) for Field Office Division (FOD) Staff Services, while the third was a retired annuitant whose most recent DMV post had been within the Training Branch but who also had more than a decade of experience as a field office employee. Hence, while these individuals were identified as employees of Research and Development, it was their experience as former FOD staff that made their services

especially valuable. This meant familiarity with specific procedures—and so very little time had to be spent explaining how things worked "in the real world." But it also meant an ability to find workable solutions to everyday problems.

The second most-commonly identified characteristic of the liaison staff was this "problem-solving" orientation:

> "Sure their personality was an important aspect, but after that initial first impression? It was their ability to follow through, find the answers, and get back to me. Because a lot of problems are about the follow-through. And they did that very well… that was probably the most important [thing]."
>
> Administrative/Back-up Manager #3
>
> "They were very helpful. Every time I called them, they were there. And the best thing was, if they didn't know the answer to the question I was asking them? They were back with the answer within 10 to 20 minutes. They would find the right answer to help me. I was so glad to have those people working with us."
>
> 3-Tier Manager I #4

In addition to knowing how field offices worked, the liaison staff showed the willingness and the ability to solve specific organizational dilemmas in a timely and efficient manner. This likely stems from, and is closely tied to, their many years' experience as FOD staff. Not only were the liaisons able to "break it down" in field office terms, but they knew where and how to find the resources necessary to solve specific operational problems—which could be as simple as restocking the specialized paperwork required for the pilot.

However, among the Driver Safety informants, very little mention was made of these liaison staff. This largely resulted from the fact that during the actual pilot period, communication between R&D and Driver Safety occurred through the Sacramento Branch manager of Driver Safety, rather than through a liaison:[29]

[29] Prior to the pilot, the designated liaison between R&D and Driver Safety was physically located in the Driver Safety Training and Procedures Unit at DMV Headquarters, rather than in either R&D or in the Sacramento Driver Safety branch office.

> "I didn't have any direct contact with anybody in R&D. I think [our manager] handled that mostly."
>
> Hearing Officer #3

While the liaisons did work with the Driver Safety manager to resolve operational problems as they occurred, this did not typically involve direct communication with Hearing Officers. Thus, when problems arose, the solutions to those problems were seen as coming from within Driver Safety, rather than through communication and coordination between Driver Safety and R&D:

> "I don't really have anything like <u>against</u> R&D. Obviously they can't check out all the bugs until you start something. The only thing that I didn't fully grasp is after we'd taken a lot of the bugs out I still had one [question], but I had a little bit of a hard time getting an answer. So I just had to make up my own determination."
>
> Hearing Officer #4

To the extent that there were still questions about new procedures mid-way through the project, these were solved either through communication between Hearing Officers and their own managers, or by individual discretion. On a secondary level, however, a few of the Driver Safety informants noted (typically indirectly) that none of the designated liaison staff had direct experience working in Driver Safety:

> "Like with [our manager], she has a lot of Driver Safety experience. So when there's something that has to be done, she's really good at breaking it down and saying 'Okay, do this and this. And this is why you need to do this and it's like Boom!' You just get it. She's just really good. So when we have questions, she can explain what to do and why, and that helps you not just to understand it, but to remember it."
>
> Hearing Officer #2

The experience and expertise of the Driver Safety manager for Driver Safety procedures operated in a similar manner to the field expertise of the liaisons *in the field offices*. However, the field expertise of the liaisons was, in some sense, foreign to the concerns and problems faced by employees of Driver Safety.

Thirdly, the liaison staff had no responsibility for brokering communication between Driver Safety and the field offices. In respect to the latter, it appears that if and when problems arose, these were resolved through one-on-one communication between individual Hearing Officers and individual field office staff (3-Tier Manager Is, LREs). This led to a level of coordination that had not previously existed:

> "We had to communicate with Driver Safety a lot more. And we had to work in conjunction with them because (a) they have time limits on their hearings where they have to get everything done on a timely basis, so (b) if we can't squeeze them in at the time, then they get pushed back and their caseloads get worked up. So we really had to learn how to work with them a lot more. Not just on Drive Tests but with the whole process. Because they would calls us and ask us questions like 'Well, do you have this information?' Because sometimes they didn't get the whole packet, you know?"
>
> Administrative/Back-up Manager #5
>
> "Well, [3-Tier] pretty much forced [Field] to link up with Driver Safety a little more than normally happens. So it shined a light on areas that maybe needed more work, or needed to be better connected on certain issues…[the staff] got to know each other more on a personal level, you know?
>
> Upper Management/Headquarters Coordinating Personnel #14
>
> "I think [3-Tier] was helpful in that it increased cooperation between Field Office and Driver Safety. We seemed to both be in the same situation in terms of dealing with 3-Tier. So it really enhanced the relationships between the managers in the field offices, the LREs, [and the Hearing Officers]."
>
> Upper Management/Headquarters Coordinating Personnel #2

This one-on-one communication was not seen as efficient, however. As noted by respondents within both Driver Safety and the field offices, it often required multiple phone calls, time that (in the view of most respondents) could have been spent more fruitfully on other tasks. Thus, while the end-product (increased communication) was seen as a good thing, the method of that communication (telephone calls as opposed to faxes or e-mail) was regarded as inefficient. Moreover, because the pilot was temporary, the increased communication did not necessarily lead to changes in standard operating procedures, and was not really expected to outlast the project.

Conclusions

Implications for Analysis of Outcome Data

What implications do these findings have for understanding the data collected during the 3-Tier Pilot? At a minimum, the analysis presented here can provide some confidence bounds regarding the usefulness of different portions of the data. But the insights provided by project participants also point the way toward some new hypotheses regarding previously-unanticipated outcomes from the project.

As reported by those interviewed, there appears to have been substantial variation in the implementation of the memory recall exercise. This variation occurred across at least two dimensions. First there was the substitution of a customer's zip code for their SSN, a substitution that was allowed under specific circumstances, according to pilot protocols. According to the interview data, this substitution probably occurred quite often—though how often cannot be estimated. Secondly, there appears to exist some degree of skepticism on the part of many respondents regarding the face validity of the memory recall exercise, which typically manifested itself by staff treating this process element as a hurdle rather than as an assessment test. In combination, this variation suggests that there are likely a number of type 2 (false negative) errors incorporated into these data. Whether or not this has implications for the predictive value of the data is beyond the scope of this paper, though it suggests that the outcome analysis will have to include sensitivity tests as well as comparisons between the outcomes on this test and outcomes on other cognitive assessments (the PRT especially, but also the cognitive elements of the SDPE).

Respondents also reported some degree of variation in the implementation of the contrast sensitivity test, driven by ambient light levels, glare, and shadows. It is beyond the scope of this module to establish definitively whether or not glare or shadows affected the outcomes on this test. However, there exists some literature (Zhang et al., 1989) that suggests that Pelli-Robson charts are, if anything, *less* sensitive than visual acuity tests to differences in ambient lighting. The author, in Module 4 of this appendix, undertook a formal quantitative analysis of this question.

Those who administered the educational intervention reported having some customers watch these videos—especially the one for contrast sensitivity—even if they were meant to be control cases according to the experimental protocol developed for this element of the pilot. This was driven in part by the quality of the video in question; respondents praised its clarity, brevity, and simplicity. Perhaps more importantly it appears that the video was useful as a customer service tool, for helping (as noted above, p. 75) to explain to customers why they were required to undergo additional assessment. Unfortunately, the very fact that the videos were distributed in a non-experimental fashion compromises the ability to assess comprehensively the causal impact of educational intervention on driver behavior and traffic safety. Even assuming that staff accurately recorded every instance of video distribution[30], the basis upon which they made the decision to give the educational intervention—the reasons why DMV staff felt that customers "didn't understand why they had to take a road test" and so needed to see the video—were not recorded. This introduces large and unknown sources of bias regarding who received (and who did not) educational intervention, which in turn may have had substantial effects on licensure (e.g., number of drive tests taken, number and type of restrictions imposed). This may also introduce substantial (but largely immeasurable) bias into any effects of the educational intervention on crash risk.

The utility of the educational intervention as a short-term customer service tool suggests a second, unforeseen hypothesis. A few respondents anecdotally reported that customers who had seen the educational intervention videos were more likely to pass their on-road driving test, as compared to customers who had not seen the videos. Specifically, informants perceived that customers who received the educational intervention were more likely to study for their test and so less likely to commit errors. Unfortunately, given the aforementioned data-quality problems, it is not possible to test this hypothesis with data collected during the pilot and baseline periods of the pilot.

[30] This appears to be a risky assumption. A number of staff reported over-distributing the videos, or in other words, giving them to customers with even-numbered driver licenses, etc. However, according to other data—see the main body of the Process Analysis—in at least one case a staff member who reported over-distributing the videos in their interview appears, according to the paperwork accumulated from their office, to have actually under-distributed the videos. In other words, the specific piece of paper recording the showing of the DVD was not filed for several dozen customers who ought to have undergone the educational intervention. This contradiction may stem from the respondent's memory, from their understanding of the randomizing experimental protocol, or from the completeness of the submitted paperwork. These three scenarios—which have very different implications for data interpretation—cannot really be untangled.

Finally, a number of respondents reported that the adoption of a structured protocol for the observation of potential driving-relevant physical limitations resulted in an "improvement" in the quality of those observations (see p. 56). If true, one potential measure of this would be in the number of referrals to Driver Safety that originate from a field office. It was not possible to test this hypothesis with data collected during the pilot, largely due to the presence of too many confounding effects associated with other differences between the baseline and pilot cohorts.

Implications for Potential Implementation

What implications do these findings have for potential statewide implementation of the 3-Tier Assessment System? Although specific proposals are beyond the scope of this analysis, the data presented here can provide some guidance for future planning for potential implementation here in California or elsewhere. These implications can be organized into four areas: production efficiency, customer service *qua* personalized attention, universal treatment/discrimination, and the cognitive health assessments. Although inter-related, each area has different implications for potential adoption of the 3-Tier process.

CA DMV places production efficiency among its top priorities—especially as measured by low customer wait times in the field offices. This organizational priority has a substantial impact on field-office level decision-making regarding any new procedures. As a result, the adoption of any or all of the 3-Tier process will likely occur in the context of streamlining and making the process as time-efficient as possible. At the very least, if process elements cannot be substantially shortened, it is likely that new procedures will be shaped to fit as closely as possible with existing processes so as to minimize organizational disruption. So, for instance, the Pelli-Robson contrast sensitivity chart, should it prove useful as a tool for regulating driver competency and improving traffic safety, is quite similar to the existing Snellen visual acuity chart. It takes very little extra time to administer, and is easy to incorporate into existing CA DMV driver license renewal procedures. Other elements of the 3-Tier process—for instance the protocol for observing potential physical limitations in the lower body—are more likely to undergo some kind of revision with an eye towards enhancing production efficiency.

CA DMV also places excellent customer service among its top organizational priorities. As noted in this paper, good customer service admits of substantially different meanings depending on the person interviewed. To the degree that customer service is equated with low customer wait times (and thus production efficiency), the implications laid out in the previous paragraph will likely apply. On the other hand, to the degree that customer service is equated with personalized attention for individual customers, it appears—at least on the basis of these interviews—that staff saw 3-Tier as presenting substantial opportunities for improving DMV's standards in this area, though staff also noted that increased personal attention presents a potential threat to production efficiency. On the basis of preliminary data from a survey of 3-Tier customers (see Module #3), it appears that personalized attention constitutes a substantially positive aspect of the average customer's experience of the pilot program. Hence, leveraging this aspect of 3-Tier may facilitate implementation of the program among both staff and customers.

Staff expressed a serious commitment to the principle of providing universal service—or in other words, with taking seriously any potential for unjustified discrimination. As revealed in the interviews, this concern arose more for some elements than for others. Evidently staff find the physical observation protocol and the two vision tests "obviously" related to traffic safety outcomes and so plainly non-discriminatory (or at least discriminatory in ways that are justifiable vis-à-vis traffic safety). Staff were more concerned with the potential for age-based discrimination when it came to the cognitive assessment tests; this was revealed in a common and substantial degree of skepticism regarding the validity of the memory recall test and the PRT. This degree of skepticism is of significant concern, as the cognitive tests constitute those elements of the 3-Tier process with the highest potential to reduce the kinds of crashes reported in the literature as stemming from declines in executive function (Freund, Colgrove, Petrakos, & McLeod, 2008) and limitations to perception (Koustanaï, Boloix, Van Elslande, & Bastien, 2007).

These concerns may be partly addressed through enhanced training for all positions that might generate, or use, the information gathered from these assessments. This includes not just the basic driver license renewal curriculum given to entry-level MVFRs, but also the training provided to LREs, Hearing Officers, and those positions that would be responsible for administering the PRT. Given the widespread degree of confusion regarding what the PRT actually tests, this must certainly be taken into

account should the PRT be adopted for wider implementation. These tests could also be re-tooled to enhance their face-validity; this is easily done for the PRT, for instance, which could incorporate an introductory segment explaining the nature and purpose of the test.

Increasing the face validity of the cognitive tests introduces a potential secondary problem—from the customer side of the transaction. Several informants noted that (some) customers showed a great deal of anxiety at taking a test "on a computer." These reports are second-hand—no customers were interviewed for this portion of the project—and so cannot be independently confirmed. However, testing in general (regardless of format) typically involves some level of anxiety for the test-taker. This may be particularly true of any test specifically for cognitive health (as opposed to, for instance, knowledge of the rules of the road). In other words, making the cognitive tests more "obvious" may reduce staff skepticism of their utility for assessing driver competency, but at the same time doing so may increase customer anxiety regarding taking such tests. Because the data analyzed here speak to this only indirectly, these must remain speculative comments. However, staff also reported improvement in customer performance on the PRT after repeated attempts—a pattern than may reflect a number of factors, one of which may well be an increase in individual perceptual speed from practice and training (Ball, et al. 2002). Should the 3-Tier Pilot confirm the PRT's predictive validity for assessing driver competency (Edwards, et al. 2002), it may be advisable to (a) determine the additive value of educational materials focused on perceptual speed, such as were used in the pilot (but only rarely, see pp. 73-74), and (b) potentially incorporate a method whereby customers can train and practice their skills in this particular cognitive area. It may also be advisable to assess in a more comprehensive manner potential customer reaction to computer-based cognitive assessment tools—particularly given that CA DMV may soon adopt an automated/computerized system of administering the written law test.

Implications for Future Research and Development Branch Projects

The 3-Tier Pilot involved a substantial number of (temporary) changes to the driver license renewal procedures used in the six pilot offices, and to a lesser extent changes in the procedures used by the Sacramento Driver Safety Branch. This necessitated long-term coordination between R&D and other branches within CA DMV: six months of

actual pilot testing, which came at the end of more than a year of planning. R&D only rarely takes on projects involving this level of intervention into current departmental practice (however, for past studies of the on-road drive test see Shumaker 1994). As noted in this paper, part of what made the pilot a success—especially in the field offices—was the existence of designated liaison staff. These liaisons served not just as channels of communication for the answering of questions, they also oversaw quality control during the implementation of the actual pilot. In this capacity they worked to solve various kinds of day-to-day problems, even (in one case) to the extent of stepping in to act as the 3-Tier Manager I of one of the pilot offices during a regular staff-member's absence. Especially as regards their ability to answer questions and resolve operational problems, the career experience of these staff—in this case, as former FOD employees with actual field office experience—was vital, at least in the eyes of the staff with whom they worked to make the pilot a success. For any future R&D project that involves the same (or similar) degree of intervention into day-to-day office operations, it would seem advisable to designate a specific liaison officer. The particular characteristics of this liaison position will, of course, depend upon the nature of the project at hand.

MODULE #3: THE CUSTOMER SURVEY

THE CUSTOMER SURVEY

Introduction

What did California Department of Motor Vehicles (CA DMV) customers think of the 3-Tier Assessment System (3TAS)? At the conclusions of the field office portion of the pilot, DMV Research and Development (R&D) administered a simple six-question survey to a stratified random sample of customers who had participated in the 3-Tier Pilot. This survey assessed customer attitudes regarding their experiences with these new assessment tests, their confidence in the utility of the program for improving traffic safety, their views on its potential impact on wait times in the offices as well as the quality of customer service, and, finally, their opinions regarding the fairness of 3TAS.

Overall, customers had positive attitudes toward the pilot; they had especially positive remarks regarding the customer service they received from CA DMV field office staff. Customers who experienced enhanced levels of assessment (the 2nd or 3rd tiers of the process) generally regarded the process more negatively, as did customers with two or more crashes in the preceding three years. Other variables, including a respondent's gender or age, appear to bear little in the way of a consistent relationship with attitudes towards the program.

There do exist significant and substantial differences between survey respondents and non-respondents: thus, those who answered the survey were more likely to be older and female, to have fewer crashes, convictions, or negligent-operator points on their record, and to have experienced more extensive assessment in the 3-Tier system. The findings discussed here are thus weighted to account for non-response bias. Due to time constraints, a small portion of customers could not be included in the original sample. These customers were somewhat different from those sampled; the implications for the survey findings are discussed in the analysis section.

Ultimately, our findings suggest that there exists a relatively wide level of acceptance of some sort of safety-focused driver competency assessment system, such as 3TAS. Conclusions include suggestions for (1) a potential follow-up survey with non-3-Tier customers, to further isolate any potential differences between those who regard a new system of driver competency assessment as acceptable versus those who do not, and (2)

a more qualitative assessment of customer attitudes towards 3-Tier—such as might be accomplished with focus-groups—to more precisely specify the means by which customer confidence in 3TAS may be improved.

Method

Background

Very little research has been done on the specific question of public attitudes regarding driver competency assessment systems. Mail surveys are widely used in traffic safety research, of course, to investigate topics such as the relationship between various demographic characteristics (e.g., age, gender, health status) and driving habits (Siren & Hakamies-Blomqvist, 2006) or driving cessation (Hakamies-Blomqvist & Wahlstrom, 1998; Neal, Baggett, Sullivan, & Mahan, 2008). There is some literature in other disciplines on attitudes towards government agencies (Goodsell, 1985; Katz, Gutek, Kahn, & Barton, 1975; National Commission on the Public Service, 1989). In general, research suggests that while there is a widespread presumption of a deep and abiding hostility on the part of the public toward government institutions (Osborne & Gaebler, 1992), surveys of the public regarding specific services tend to be rather positive (Poister & Henry, 1994; Miller & Miller, 1991). That said, there is little published academic research on attitudes towards departments of motor vehicles, and none (that this author is aware of) on CA DMV specifically.

This survey was developed primarily to gauge the potential impact of the 3-Tier assessment system on the perceived quality of service provided to DMV customers. Given the extreme concern within the department regarding field office production efficiency, the survey also included a question regarding customer wait times. As regards 3TAS specifically, there were questions regarding the program's "user-friendliness" (did the instructions make sense, was it easy to follow), its fairness, and the respondent's perceptions regarding the program's potential contribution to improving traffic safety. Taken together, the questions provide some data regarding how acceptable—across a number of dimensions—3TAS will likely be to the public. See Sub-Appendix D for exact question wording.

Variables

Because of the newness of 3TAS, the author did not engage in *a priori* theorizing and hypothesis-generation about what we could expect in terms of customer attitudes. Hence, the analysis presented here is largely of a descriptive nature. That said, we were interested in any potential variation in customer evaluations of 3TAS. Those evaluations could vary, first, by location. Hence, the analysis includes the office at which a customer's application was processed (or the office at which they took their drive test, if they were a Driver Safety case). Also included was the level of assessment experienced by a customer. This was operationalized according to four categories:

(a) Tier 1: These customers were given a simple short-term memory recall test, observed for any potential physical limitations, and took two vision tests (visual acuity and contrast sensitivity).
(b) Tier 2 (Perceptual Response Test): These customers experienced all of the same tests as Tier 1 customers, but were flagged as having a potential visual, cognitive, or physical limitation. These customers were given one additional assessment test for potential cognitive and visual limitations (see Hennessy, 1995 for details on the PRT).
(c) Tier 2 (Educational Intervention): These customers were, on the basis of a randomizing experimental protocol, given the opportunity to view a short (4 minute) educational video on the potential visual or cognitive limitation for which they had been flagged in Tier 1 of the process.
(d) Tier 3: These customers included all those who had been identified as having one or more potentially serious driving-relevant limitations; as a result, they were required to take an on-road drive test.

These categories also served as the primary basis for the sampling stratification (see below).

The other sources of potential variation in customer evaluation derive from the factors which identified a customer as "3-Tier eligible." For the pilot project, 3-Tier eligible customers included all those who were required to visit an office for their driver license renewal (as opposed to renewing their license by mail or online), who sought a class-C (non-commercial) license without special endorsements (e.g., for a motorcycle), and who were required to take the 18-question written law test (but chose to do so in

English, as opposed to some other language). There were essentially two factors which define this class of customers: those with negligent operator points on their record (Gebers & Roberts, 2004), and customers aged 70 or older; both of these groups are required by law to take the written law test at a DMV field office when renewing their license. Because these two factors influenced who was included in 3-Tier, the analysis takes account of both customer age and the number of negligent operator points assigned to the driver's record. However, instead of measuring the number of neg-op points directly, the analysis includes the number of traffic convictions—which may or may not result in points. This taps in a more general way the respondent's involvement with the legal system governing California's roads.[31] Also included is the number of crashes for which a customer was found to be at-fault, and which were reported to the DMV. Convictions and crashes are both measured for the three-year period prior to the date on which the customer began their license renewal process as part of the 3-Tier program.[32] They were both operationalized as dummy variables: 2-or-more crashes in the prior 3 years, and 1-or-more convictions in the prior 3 years. Age was measured as a dummy variable, with seniors (those aged 65+) coded as 1.[33] Finally, the analysis includes a dummy variable for any customer who was enrolled in 3-Tier as a result of a Driver Safety referral.

[31] This substitution was chosen on both theoretical and methodological grounds. Theoretically, the number of convictions represents a fuller range of law-breaking activity, while neg-op points generally includes only more serious convictions. Methodologically, these two variables are highly correlated—and thus represent a threat of multi-colinearity if both were included in the regression analyses. The results presented in the body of the paper do not differ substantially if the variable measuring neg-op points is included instead of the variable measuring number of convictions (equations not shown, available upon request from the author).

[32] This date may or may not coincide with the date on which the customer completed the process. Customers who failed one or more written tests, or who had to take an on-road driving test, may have been issued their license days or even weeks after their initial application.

[33] Various operationalizations of age were tested, including a continuous/interval level (age in years), a series of 5-year incremental dummy codes, and a tripartite ordinal variable differentiating non-seniors, "early seniors" (65-74), and "extra seniors" (75+). Where age was significant, it appears that the cut-point for significance is the traditional age of retirement (65), and not the age at which CA DMV begins requiring in-office renewals for all customers (70) (equations not shown, available upon request from the author).

Sampling Procedure

At the conclusion of the enrollment of new customers in the field office portion of the 3-Tier Pilot (10/31/07), the author extracted from the project database a list of all customers with completed files. This list—of 10,699 customers—was then stratified according to one of four possible categories. These categories were derived from the nested hierarchy of assessment tools administered as part of the 3-Tier system:

(i) Tier 1 customers (n = 8,468)
(ii) Tier 2 customers who had taken the PRT (n = 1,159)
(iii) Tier 2 customers who viewed the experimental educational videos (n = 513), and
(iv) Tier 3 customers (n = 559).

Using SPSS (ver. 14), each stratum was then randomly sampled, with over-sampling for all groups except Tier 1 customers. The final sample consisted of: 4,167 Tier 1 customers (49% of this stratum); 770 PRT customers (66% of this stratum); 453 customers receiving educational intervention (88% of this stratum), and 481 Tier 3 customers who had to take an on-the-road drive test (86% of this stratum).[34]

On the basis of evidence presented in Module #2 ("The Staff Interviews"), it was determined that there exists a substantial amount of error in the designation of customers receiving educational intervention. In particular, some number of customers who were not supposed to see the educational videos appear, in fact, to have seen them. More worrisome is the fact that there exist also a substantial number of respondents who, according to the paper files, should have seen the videos but did not. Due to the nature of the forms used to construct the original database, there is no way to estimate

[34] During data-cleaning, it was determined that a small number of customers (98) had various sorts of uncorrectable errors in the method by which they were processed in the 3-Tier pilot. This included, for instance, customers that should have taken a drive test but did not, or who should have taken a PRT but did not. Even if they are coded according to how they were actually handled (as opposed to how they were supposed to have been handled), their attitudes towards the program are based upon something other than the 3-Tier process as they should have experienced it. Thus, they are excluded from all subsequent analyses discussed in this paper.

this second source of error.[35] This problem was discovered after the sample was constructed and the surveys mailed. This seriously impairs the degree to which one can reliably distinguish between customers who took the PRT and those receiving educational intervention. As a result, in the results presented below, the sampling design weights (expansion and relative weights; Lee & Forthofer, 2006) reflect the original sampling procedure—since this determined a customer's probability of inclusion in the sample. However, the non-response weights and final logistic regression modeling were modified to consist of 3 categories: Tier 1, Tier 2 (either PRT or educational intervention) and Tier 3 customers. The Tier 2 categories could be collapsed, and reliably distinguished from Tiers 1 and 3, because these customers (a) took the PRT, which automatically generated an additional piece of paperwork (with the customer's score) for inclusion in each customer's file, but (b) did not take an on-road drive test (which would have generated a drive test score sheet).

Surveys were distributed in two waves: the first was mailed on 12/07/07, and the second wave was mailed on 1/16/07. Depending on when a given customer began and ended their application, this meant that there was a maximum of six months between visiting their local CA DMV field office and receiving the survey. Each survey was labeled with a randomly-generated unique identifier; this allowed for later matching of individual responses to demographic variables drawn from CA DMV's driver record database.

Limitations of These Data

There are a number of caveats to the analysis which follows. First, some 3-Tier customers were not included in the sampling frame, and so were not surveyed. Although enrollment of new customers through regular field office renewal procedures ended on 10/31/07, not all customer files were complete on that date: this included, for instance, drivers who enrolled in the program towards the end of October and were scheduled for drive tests in November or December. As the Driver Safety portion of the program continued until the end of the 2007 calendar year, this "unsampled" group

[35] Technically, the problem lies in how to interpret non-response. For customers who did not see the video (but who ought to have), the paper files either do not have that specific response box checked, or they are missing the specific form which included that response box. This may indicate either (a) that the customer did not watch the video, or (b) that they did watch the video, but that the attending staff person failed to fill out the required paperwork. These two scenarios—which have very different implications for data interpretation—are unfortunately indistinguishable in this situation.

also includes some Driver Safety referrals to the field office. Finally, some customers started the 3-Tier process but then, for a variety of reasons, did not finish their application by 10/31/07. Together, this unsampled group includes 1,134 customers. Because the pilot period extended over a number of months, the author decided to sacrifice the potential data represented by these customers (9.6% of the pilot population) to avoid the threat to question relevance (and thus validity of the findings) represented by waiting longer to distribute the surveys. The analysis of the survey results includes some basic demographic comparisons between sampled and unsampled 3-Tier customers (see Table M3.1). This can provide some estimation of potential sources of bias. However, for technical reasons (i.e., because their probability of being sampled was zero), the findings presented here cannot be generalized to these customers.

Secondly, while the surveys were distributed as soon as possible after the completion of the project, some time elapsed between a customer's visit to the field office and their answering of the study questions. Unfortunately, due to resource constraints, it was not possible to conduct a survey at the time of each customer's office visit; this might have produced quite different results from those reported here.

Finally, in order to encourage a high rate of survey completion and return, the number of questions was kept to a minimum (6), and phrased so as to maximize brevity and simplicity. This necessarily compromised the degree to which the analysis could include topics in-depth, as might happen with face-to-face interviews or even focus groups.

Analysis Techniques Used

For both descriptive and inferential analyses, the author used SPSS/PASW (ver. 14). In assembling the descriptive statistics (see Table M3.1), the author determined through the exploratory use of chi-square and ANOVA procedures that survey respondents differed in statistically significant ways from non-respondents (and both, in turn, from unsampled customers) on a number of dimensions. These included age, gender, having crashes or violations on the driver record, and assessment level within the 3-Tier Pilot. These differences indicated the necessity for constructing sampling weights to correct for non-response bias, in addition to the inclusion of design weights to correct for over-sampling of Tier 2 and Tier 3 customers.

The sampling design weights were relatively straightforward to construct (Lee & Forthofer, 2006). All customers (whether or not they replied to the survey) were weighted by the inverse of the probability of inclusion in the sample (the expansion weight).[36] This "reduced" the statistical impact of answers by customers who were oversampled (such as Tier 2 and Tier 3 customers). These weights were then divided by the mean of the expansion weights, to construct the relative weight—which scaled down the weighted survey sample size to the actual *N* of the sample.

Creating the non-response weight necessitated the construction of a logistic regression predicting the probability of not replying to the survey. The following variables were used as predictors in this regression: gender, having 2 or more at-fault crashes on one's driver record for the previous 3 years, having 1 or more convictions in the previous 3 years, the office in which the customer was processed, whether or not a customer had contact with the Driver Safety branch, assessment level within the pilot (combining both PRT and educational intervention customers into a single Tier 2 category), and age. In this regression, age was operationalized as a tripartite ordinal variable (< 40, 41-65, and 65+) to maximize variance explained, as reflected in the size of the Nagelkerke pseudo-R-squared statistic.[37] The regression output (predicted probability of responding) was then added to the relative weight to produce the adjusted relative weight according to the following formula: Adjusted Weight = Relative Weight/(1-Predicted Probability).[38] The predictive analysis thus used data weighted by the adjusted relative weight, to account for non-response bias.

[36] Since customers whose files were later determined to have uncorrectable errors in processing—and so who were excluded from the descriptive and inferential analyses—*had* been included at the time of sampling, the *N*s used to construct the stratum design weights include these "uncorrectable" customers.

[37] As with the regression equations used for the predictive analysis, various operationalizations of age were used. A dummy variable (set at age 65) produced a slightly lower Nagelkerke pseudo-R^2. Setting the dummy at age 70, however, produced unacceptable goodness-of-fit Hosmer-Lemeshow statistics. Equations not shown, available upon request from the author.

[38] This was then double-checked to ensure that the sum of the adjusted weights (5955) was approximately equal to the sum of the relative weights (5869), as well as the actual survey sample *N* (5871). The maximum weights assigned were 61 and 51. These were for a pair of respondents in their early 20s who each had eight convictions and a crash on their records. Incidentally, they both answered "agree" or "strongly agree" on all questions.

In the binary logistic regression models used to predict customer attitudes, the responses "disagree strongly" and "disagree" were combined, while the responses "agree" and "agree strongly" were combined. All data are presented as the likelihood of a customer responding "disagree" or "disagree strongly" *as opposed to* "agree," or "agree strongly" In other words, the analysis is structured so as to highlight the factors which predict *negative* attitudes towards 3-Tier.

Stepwise elimination procedures were used to trim all models. As a result, two variables were discarded from the analysis. This included the office at which customers were processed and whether or not they were a Driver Safety referral. The elimination of these variables has some substantive theoretical implications: customer attitudes towards 3-Tier do not appear to relate to where they were processed in the field, nor even to which division (FOD or LOD/Driver Safety) was responsible for processing their transaction. At the most basic level this may indicate that implementation of the 3-Tier process was relatively uniform, or at least that any deviations from training protocols were unrelated to the reported opinions (and thus, potentially, the experiences) of customers. Secondly, it does not appear that Driver Safety customers—who are more likely than regular renewal customers to have driving-relevant health issues, or to have had prior involvement with the law—had substantially different attitudes from other customers. That said, there were very few Driver Safety cases; they represent less than 2% of the 3-Tier customer population; hence the lack of significance for this variable may simply reflect a lack of statistical power.

Results

Descriptive Statistics

Table M3.1 presents descriptive statistics for the sampled and (for comparison) unsampled customers participating in the 3-Tier Pilot. As a general matter, the overall response rate was good: 49%. The response rate rises for customers who experienced more advanced levels of assessment within the 3-Tier system. Approximately two-

Table M3.1: Descriptive Statistics (Unweighted) on 3-Tier Customer Survey Respondents, Survey Non-Respondents, and Unsampled Customers

		Survey respondents	Survey non-respondents	Unsampled customers
Office	Carmichael	895 (31.5%)	749 (25.6%)	282 (24.6%)
	Fairfield	218 (7.7%)	312 (10.6%)	99 (8.6%)
	Folsom	564 (19.9%)	471 (16.1%)	161 (14.0%)
	Sac – Broadway	463 (16.3%)	671 (22.9%)	215 (18.8%)
	Sac – South	479 (16.8%)	502 (17.1%)	312 (27.2%)
	Vacaville	226 (7.9%)	227 (7.7%)	77 (6.7%)
Assessment level	Tier 1	1,688 (59.3%)	2,397 (81.8%)	731 (63.8%)
	Tier 2 - PRT	516 (18.1%)	244 (8.3%)	215 (18.8%)
	Tier 2 - educational intervention	335 (11.8%)	122 (4.2%)	38 (3.3%)
	Tier 3 - drive test	306 (10.8%)	169 (5.8%)	162 (14.1%)
Gender	Male	1,358 (47.7%)	1,687 (57.5%)	533 (46.5%)
Prior 3 years':				
Convictions	Mean (*SD*)	0.19 (0.59)	0.92 (1.42)	0.62 (1.20)
Crashes	Mean (*SD*)	0.16 (0.46)	0.23 (0.56)	0.23 (.56)
Neg-op points	Mean (*SD*)	0.16 (0.61)	0.82 (1.45)	0.53 (1.13)
Driver safety referral	Referred to or from Driver Safety Branch	45 (1.6%)	44 (1.5%)	39 (3.4%)
Age	Mean (*SD*)	68.97 (13.76)	45.43 (19.11)	57.26 (20.79)
Total *N*		2,845	2,932	1,146

Table M3.2: Descriptive Statistics (Unweighted) on Customer Attitudes Towards the 3-Tier Process[a,b]

Question #	Blank/ no answer	Disagree strongly	Disagree	Agree	Agree strongly
1: Time reasonable?	28 (1.0%)	68 (2.4%)	133 (4.7%)	**1774 (62.4%)**	842 (29.6%)
2: System easy?	56 (2.0%)	64 (2.2%)	185 (6.5%)	**1725 (60.6%)**	815 (28.6%)
3: Instructions easy?	56 (2.0%)	58 (2.0%)	217 (7.6%)	**1672 (58.8%)**	842 (29.6%)
4: Staff was courteous?	40 (1.4%)	42 (1.5%)	80 (2.8%)	1192 (41.9%)	**1491 (52.4%)**
5: System fair?	99 (3.5%)	113 (4.0%)	253 (8.9%)	**1506 (52.9%)**	874 (30.7%)
6: Will it improve traffic safety?	184 (6.5%)	101 (3.6%)	334 (11.7%)	**1440 (50.6%)**	786 (27.6%)

[a] Modal Answers in **bold**.

[b] Unweighted total *N*: 2845

thirds of those customers who took the PRT (68%), or were given an on-road drive test (64%) participated in this survey.[39]

Predictive Analysis

"THE TIME I SPENT DURING MY OFFICE VISIT WAS REASONABLE"

The first question on the survey was intended as a partial measure of customer reaction to 3-Tier's potential impact on time spent at a DMV office. In the main body of the process analysis, the author examined the objective impact of 3-Tier on production efficiency and wait times in the field offices. However, the question included in the survey was intended to tap the *subjective* impact of the new process on customer perceptions of processing and wait times. Although there is no way to compare these survey results to non 3-Tier customers (or to customers who visited non-pilot field offices), the question does ask customers to implicitly link their experience of the program with any potential concerns in this area. An overwhelming majority of customers—over 90%—agreed or strongly agreed with the statement "the time I spent in the office was reasonable" (see Table M3.2). Clearly, whatever the objective impact of 3-Tier on production efficiency, customers were largely satisfied in this respect with their experience of the new process.

As noted in Table M3.3, those customers who experienced higher levels of assessment (Tiers 2 or 3) were substantially more likely to disagree with this statement. In other words, having to undergo additional screening is correlated with a negative appraisal of 3-Tier's impact on wait-times in the field offices. This is perhaps not surprising, as higher levels of assessment did, in fact, require that a customer spend more time in the office. This was especially true for those customers who had to make one or more return visits to, for instance, take an on-road drive test. Those with crashes and convictions on their records were also more likely to report negative assessments of 3-Tier's impact on wait times. This is somewhat less straightforward to explain, as these customers did not necessarily experience higher levels of assessment, and hence did not necessarily spend more time in the field office as compared to other customers, or as compared to what they would have experienced under a "normal," non 3-Tier, renewal process. Senior

[39] Nearly three-quarters (73%) of those marked as receiving educational intervention returned their surveys.

Table M3.3: Binary Logistic Regression Results; Predicted Odds Ratios of Answering "Disagree" or "Disagree Strongly" for Six Questions on Attitudes Towards 3-Tier[5]

	Model #1: time[1]		Model #2: ease		Model #3: instructions		Model #4: courtesy		Model #5: fairness		Model #6: safety[2]	
	β (*SE*)	Expβ 95% *CI*	β (*SE*)	Expβ 95% *CI*	β (*SE*)	Expβ 95% *CI*	β (*SE*)	Expβ 95% *CI*	β (*SE*)	Expβ 95% *CI*	β (*SE*)	Expβ 95% *CI*
Assessed at tier 2[3]	0.35 0.15*	1.41 1.06-1.89	0.88 0.14**	2.41 1.82-3.18	1.13 0.13**	3.08 2.38-3.99	0.25 0.19	1.29 0.89-1.86	0.74 0.11**	2.09 1.67-2.61	0.67 0.10**	1.95 1.59-2.38
Assessed at tier 3[3]	1.47 0.16**	4.33 3.15-5.97	2.22 0.15**	9.25 6.88-12.42	2.22 0.15**	9.22 6.89-12.32	1.51 0.19**	4.52 3.12-6.55	1.91 0.13**	6.77 5.21-8081	1.32 0.13**	3.74 2.88-4.86
Crashes (2+ in 3 yrs.)	1.03 0.18**	2.80 1.97-3.98	-0.01 0.32	0.99 0.53-1.86	-0.16 0.31	0.86 0.47-1.57	1.47 0.20**	4.35 2.96-6.40	0.96 0.20**	5.60 1.77-3.81	0.81 0.18**	2.24 1.59-3.16
Convictions (1+ in 3 yrs.)	0.40 0.11**	1.49 1.20-1.85	0.06 0.15	1.06 0.80-1.41	0.22 0.13	1.24 0.96-1.61	-0.02 0.14	0.98 0.75-1.29	0.19 0.11[4]	1.21 0.98-1.50	0.16 0.09[4]	1.17 0.98-1.40
Gender (male = 1)	0.04 0.10	1.04 0.86-1.27	-0.05 0.18	0.95 0.76-1.19	-.22 0.11*	0.80 0.65-0.99	0.11 0.75	1.11 0.87-1.42	0.16 0.09[4]	1.18 0.99-1.40	---	---
Senior (65+ = 1)	-0.75 0.12**	0.47 0.37-0.60	0.26[4] 1.41	1.30 0.98-1.71	0.18 0.13	1.20 0.92-1.55	-0.72 0.15**	0.49 0.36-0.65	-0.02 0.11	0.98 0.79-1.21	-0.18 0.09[4]	0.84 0.70-1.00
Constant	-2.53 0.17**	0.08	-3.24 0.21**	0.04	-2.87 0.19**	0.06	-3.04 0.21**	0.5	-2.74 0.16**	0.07	-1.95 0.08**	0.14
-2 Log likelihood	3166.14		2512.44		2781.60		2229.20		3717.11		4557.12	
Nagelkerke R^2	0.069		0.104		0.106		0.065		0.08		0.04	

* = Significant at the 0.05 level

** = Significant at the 0.01 level

[1] This model has a poor goodness-of-fit, as indicated by its Hosmer-Lemeshow statistic ($p < 0.05$). See text for explanation

[2] Gender was omitted from this model in order to ensure acceptable goodness-of-fit, as indicated by the Hosmer-Lemeshow statistic

[3] Tier 1 is the omitted reference category

[4] These predicted odds ratios approach significance (<.10). See text for discussion

[5] All models incorporate sampling design and non-response weights

citizens were somewhat more likely to report a positive evaluation of 3-Tier's impact on field office wait times. Those aged 65 and older have a lower odds of disagreeing with this first question than those under the age of 65. Gender is both statistically non-significant and substantively close to zero.

3-TIER PROCESS: EASY TO FOLLOW, EASY TO UNDERSTAND?

The next two questions on the survey asked a respondent how easy the 3-Tier system was, both as an overall process (Question #2), and in terms of the instructions provided for each sub-component (Question #3). For both questions, those respondents who experienced higher levels of assessment were more likely to report negative attitudes. However, having crashes on one's record had no statistical relationship to reported attitudes. Gender and having convictions were significant only for question #3; female respondents, and those with convictions, were both more likely to disagree with this statement. Age, however, was not statistically significant for either question—though it approaches significance for the statement regarding whether or not the 3-Tier system was "easy to follow" (p value = 0.067).

"THE DMV OFFICE STAFF TREATED ME WITH COURTESY AND RESPECT"

Very few respondents disagreed with the statement "The DMV office staff treated me with courtesy and respect." Indeed, the modal answer on this question was "strongly agree." The results shown in Model #4 are largely consistent with those found for other questions: those who experience higher levels of assessment were more likely to report negative evaluations of 3-Tier's impact in this area. Also, those with crashes on their records were more likely to disagree with the notion that they were treated with courtesy by DMV staff. On the other hand, senior citizens were substantially *more likely to agree* with this statement. Gender, and a record of traffic convictions, had no statistically significant relationship to respondent attitudes regarding the quality of customer service.

"IN MY OPINION, THIS NEW ASSESSMENT SYSTEM IS FAIR TO ALL CUSTOMERS"

In regards to the statement "In my opinion, this new assessment system is fair to all customers," there were two significant predictors: level of assessment, and number of crashes. Customers who were assessed at Tiers 2 and 3 of the process were more likely

to disagree with this statement as compared to other customers. In addition, customers with more at-fault crashes on their record were less likely to regard the 3-Tier system as fair. Other variables—including the number of convictions, gender, and age—all had odds ratios close to 1 (indicating no direction of effect) and were statistically not significant.

"I AM CONFIDENT THAT THIS NEW ASSESSMENT SYSTEM WILL IMPROVE DRIVER SAFETY"

The final statement asked respondents how confident they were that 3-Tier would improve traffic safety. Customers were, overall, somewhat more likely to disagree with this statement as compared to other statements on the survey; "only" 78% (as opposed to 83-94%) marked "agree" or "strongly agree." That said, the pattern of predicted responses is substantially the same for this question: those who experienced higher levels of assessment were more likely to disagree, as were customers with two or more recent at-fault crashes on their records. Conviction record had no statistically significant effect. Age comes quite close to the standard statistical cutoff ($p = 0.057$); however, the direction of effect is somewhat surprising: senior citizens were somewhat *more likely to agree* with the statement that they have confidence that 3-Tier will improve traffic safety. Gender was dropped from the analysis to improve the goodness-of-fit; this had only a marginal effect on both the size and the statistical significance of the beta coefficients for the other variables. When included, gender is significant: men are less likely to disagree with statement #6 as compared to women (odds ratio = 0.807).

Potential Sources of Bias to the Predictive Models

All regression models included sampling design and non-response weights. These correct, to some extent, for potential bias to the findings. However, a small proportion of the 3-Tier customer base (approximately 9% of all customers) was not included in the sampling frame. As described in Table 1, these customers were somewhat different from those that were sampled. This "unsampled" group was slightly more likely to have been female, to have been a Driver Safety Referral, to have had a recent at-fault crash, and to have been assessed at Tier 3 of the process. In respect to the latter two variables in particular, the exclusion of these customers may understate to a small degree the proportion of customers who evaluate 3-Tier negatively: customers with at-fault crashes, as well as those required to take a drive test, were consistently most likely to disagree with positively-phrased statements about the pilot.

Conclusions

The findings presented here indicate a potentially broad level of acceptance on the part of DMV customers for the 3-Tier Assessment System. In excess of three-quarters of those surveyed agreed with six positively-phrased statements regarding the process. These statements included: the time customers spent in the field offices while participating in the pilot, the functional ease of the 3-Tier process as a whole as well as the instructions for each assessment, the program's impact on the quality of customer service and, perhaps most crucially, the fairness of 3-Tier and its potential for improving traffic safety. Within these overall trends of broad public support are two items worth highlighting; the first regards who is most likely to view 3-Tier negatively. The second trend involves a non-finding; which is to say, there exists consistent evidence that senior citizens are not more likely to view 3-Tier negatively.

Intriguingly, those who experienced the second and third assessment tiers were substantially and significantly more likely than Tier 1 customers to disagree with positively-phrased statements about the system. This was true for all six statements. Although the size of the odds ratios (i.e., the likelihood of holding negative views about 3-Tier) varied from question to question, the direction of effect is always the same. Theoretically, this may indicate one of two things, which have very different implications. First, it may indicate that increased exposure to the 3-Tier system produces negative sentiments, and thus reduces the likelihood of acceptance on the part of individual customers. On the other hand, this correlation may simply be identifying those who regard driver competency assessment *in general* in a negative light—and so view 3-Tier with skepticism not because of anything about the program itself, but simply because it (for example) represents a risk of non-renewal of their driver license and thus the potential loss of their driving privilege. The first possibility suggests that the 3-Tier program may be altered or revised to enhance public acceptance; the second possibility suggests that some portion of the motoring public will always regard driver competency assessment with skepticism.

This second possibility—that negative sentiment towards 3-Tier stems not from nature of the program itself, but rather from the fact that it involves driver competency assessment—is supported by a second piece of evidence. Those with recent at-fault crashes and/or convictions on their records were substantially and significantly more likely to disagree with positively-phrased statements about 3-Tier. This population

consists precisely of those drivers with demonstrated risk for unsafe driving; yet these drivers were also more likely to disagree with the statement that 3-Tier was fair, or that it would improve traffic safety. That individuals with poor driving records regard assessment of driver competency with skepticism may or may not be alterable—this point is most properly the subject of further investigation. Regardless, those who regard 3-Tier with skepticism are somewhat rare: the vast majority of those surveyed agreed with positive statements about the pilot.

On the other hand, a specific population of concern to this study—senior citizens—appears no more likely than other customers to have concerns about the fairness or traffic-safety improvement potential of the 3-Tier program. If anything, those aged 65 or older are more likely to view 3-Tier *positively*—other effects held constant—especially when it comes to the pilot's potential for improving the quality of customer service at CA DMV. There is even some potential evidence that seniors were somewhat more likely to regard 3-Tier as having the potential to improve traffic safety. In a perhaps related point, these customers were also least likely to regard any associated increase in wait times as a negative outcome of 3-Tier.[40] When taken in conjunction with the results presented in Modules #1 and #2 it appears that 3-Tier's potential for exciting (or exacerbating) concerns about age-related discrimination are, at least according to the evidence discussed here, somewhat overblown. Other factors held constant, those senior citizens surveyed for this paper were no more likely than respondents of other ages to regard 3-Tier as unfair. Far more crucial than age was whether a respondent was assessed at Tiers 2 or 3, as well as the extent to which they had a record of at-fault crashes or convictions. Thus, although there is a relationship in these data between age and level of assessment—one third of seniors in the survey, versus 8% of non-seniors, experienced Tiers 2 or 3 of the process—once controlled for, age has no effect. Said another way, non-seniors who took the PRT or an on-road drive test were no less (nor more) likely than senior citizens to regard 3-Tier as unfair.

These results indicate the need for additional research in at least two directions. In the first instance, a survey that included a broader range of independent variables—as might occur with a survey of the general population—might more precisely identify those populations particularly concerned about the institution of this new system of driver competency assessment. Although the survey discussed here suggests that

[40] This may indicate a number of things, not the least of which is that persons over the age of 65 are more likely to be retired—and thus somewhat more likely to have flexible schedules with which to fit in a visit to a DMV field office.

relatively few customers (≈10-20%) viewed 3-Tier in a negative light, it was difficult to say with any precision who these customers were—other than that they tended to have poor driving records, and were those most likely to face higher levels of competency assessment under the new pilot procedures. Secondly, the collection of qualitative evidence—such as through focus groups—would perhaps more precisely identify what kinds of information and education about 3-Tier would reduce concerns and increase acceptance. It appears from these data, for instance, that there exists some room to provide clarification regarding 3-Tier's potential to improve traffic safety. However it is not clear what form the communication of this information ought to take: distribution of pamphlets and videos, not to mention increasing the face-validity of the component assessment tests are all possible methods for conveying 3-Tier's potential contribution to the safety of California's drivers.

In terms of implications for potential statewide implementation, it appears from these data that there exists substantial room for public education regarding 3-Tier's potential contribution to public safety. Statement #6—"I am confident that this new assessment system will improve driver safety"—had both the highest rate of question-specific non-response (6%) and the highest proportion of respondents who disagreed or disagreed strongly (15%). This suggests that the face-validity of the 3-Tier process—the connection between the assessments and a given customer's perception that these assessments relate specifically and obviously to driver competency—may be improved at the margins. Given the data presented here it is not clear which specific element(s) of the 3-Tier process could be altered in this way; this may be the subject of further research. In a related point, it would appear that the results of the outcome analysis—where CA DMV will measure the actual traffic safety benefits of the 3-Tier system—will be an important part of any public education campaign. To the extent that 3-Tier provides a reduction in deaths on the road, and/or a reduction in traffic violations, these benefits to California's motorists should be made widely available to the public.

MODULE #4: ROBUSTNESS OF THE PELLI-ROBSON CONTRAST SENSITIVITY CHART BY LOCATION AND TECHNICIAN

ROBUSTNESS OF THE PELLI-ROBSON CONTRAST SENSITIVITY CHART BY LOCATION AND TECHNICIAN

Introduction

How robust under different light conditions were the Pelli-Robson contrast sensitivity charts used by the California Department of Motor Vehicles (CA DMV) during the 3-Tier Assessment System Pilot? As noted in Modules #1 and #2 of this Appendix, a number of staff who participated in the 3-Tier Pilot raised concerns over the robustness of these charts under varying light conditions. In keeping with best practice standards for evaluation studies (Rossi, Lipsey, & Freeman, 2004; Patton 2008), the author undertook a formal analysis of customer outcomes on the Pelli-Robson chart, using hierarchical logistic regression modeling. This analysis revealed that outcomes on this assessment were primarily a product of customer age, having been previously diagnosed with a serious vision condition, and which technician administered the assessment. All other effects held constant, a small number of charts were found to be associated with statistically different probabilities of passage/failure on this assessment, compared with other charts within the same office. Where chart location was statistically significant, it was usually the case that only one or two technicians processed the majority of customers at that chart—meaning that any independent effect of chart location could not be statistically disentangled from the effect associated with which technician administered the assessment. Thus, while chart location within a field office may bear some relationship to customer outcomes, it is difficult to determine to what extent this variation is a product of variable light conditions *per se*—as opposed to some other possible explanation, such as variation across technicians, or unmeasured bias in the distribution of different types of customers across technicians and charts. The implications for potential implementation of this vision screening tool in the event of statewide adoption of the 3-Tier Assessment System are discussed in the conclusions.

Background

In both the staff surveys (Module #1) and interviews (Module #2), a number of 3-Tier Pilot staff expressed concern about the degree to which customer outcomes on the Pelli-

Robson chart were a product of variable light conditions. These reservations were expressed in a number of ways. In a few offices (especially South Sacramento and Vacaville), staff reported "glare" from sunlight streaming through large banks of windows onto the charts—especially at particular times of day (e.g., late afternoon). In other offices (Carmichael and Broadway), staff reported the presence of "shadows" cast by nearby visual acuity charts. In the view of those who raised the issue, the letters on some charts—those on which glare or shadows were perceived to fall—were more difficult to read than the letters on other charts, thus affecting test outcomes. In all cases staff were concerned about the universality (or fairness) of the test—the degree to which the test was applied equivalently to all customers. In brief, some staff questioned the degree to which outcomes on the contrast sensitivity assessment were a product of customer vision health, as opposed to the location at which they took the test (i.e., a factor ostensibly unrelated to safe driving).

During the pilot period, these concerns were addressed in an ad-hoc manner in three ways. Current CA DMV field office procedures allow a customer to take the Snellen visual acuity test from any location in an office—i.e., not just on the chart associated with the terminal at which they happen to be conducting their transaction. This was adapted for the use of the Pelli-Robson chart as well, and customers were allowed/encouraged to take their time, look at the chart from a different angle, or switch stations altogether and test on a different chart. As an extension of this, individual staff reported "angling" (i.e., bending) certain charts to make them more readable for customers. In at least one case (Carmichael), a single chart was physically moved mid-project—this was done to resolve complaints that the shadows falling on it (from nearby Snellen visual acuity charts) were obscuring the letters and making it more difficult to read, relative to the other Pelli-Robson charts.

Within the clinical and academic literature, there is little evidence to suggest that contrast sensitivity, particularly as tested on the Pelli-Robson chart, varies substantially by lighting conditions—assuming normal visual health and a lighting range typically experienced in an indoor office environment. Zhang, Pelli, and Robson (1989), for instance, find that individual contrast sensitivity varies almost not at all across a wide range of vision conditions, from quite dim (7 cd/m^2, approximately twilight) to well-lit (500 cd/m^2, at the upper end of most indoor office lighting). Even brighter conditions (900 cd/m^2, approximately equivalent to outdoor light on an overcast day) may have

some effect, though the effect appears to *increase* contrast sensitivity, rather than decrease it (Cox, Norman, & Norman, 1999). In other words, very bright light may make it somewhat *easier* for a person to pass a given chart, assuming normal visual health. Even in the presence of spatial noise (e.g., the blurriness of an image, especially of edges), there appears to be little in the way of an independent effect of differing luminance levels on contrast sensitivity (Rovamo, Kukkonen, Thppana, & Näsänen, 1993).

To the degree that individuals experience variation in their contrast sensitivity under different light conditions, this may be a product not of the light *per se* but rather of some underlying vision health condition. Brown and Garner (1983) found that patients with senile macular degeneration may have trouble adapting to changes in luminance, specifically in terms of contrast detection, and so complain of vision problems at both high and low luminance levels. Other authors (Owsley, Sekular, & Siemsen, 1983; Richards, 1977) indicate that sensitivity to changes in luminance—as measured by contrast sensitivity as well as by visual acuity—are to some degree age-dependent. Furthermore, these changes are potentially a product of age-related changes to the neural mechanisms associated with vision (Sloane, Owsley, & Jackson, 1988).

The relationship between age and declines in contrast sensitivity shows up especially in studies of sensitivity to glare. Several studies have shown that older drivers (variously defined) report more discomfort in the presence of glare on a given target (Sheedy, Smith, & Hayes, 2005), that glare impacts contrast sensitivity in a manner that interacts with age and underlying vision conditions (Anderson & Holliday, 1995), and that this has direct implications for individuals' comfort (and safety) driving at night (Puell, Palomo, Sánchez-Ramos, & Villena, 2004; Rassow, 1999; Sturgis & Osgood, 1982). Intriguingly, declines in contrast sensitivity may be attributed to glare (rather than contrast), though this may still result in self-restricting driving—so that while individuals may not precisely understand the nature of their own vision problems, they may still take steps to limit the conditions under which they drive (West, et al., 2003).

According to the academic literature, then, it would appear unlikely that variable lighting conditions (whether consisting of shadows or glare) affected outcomes on the Pelli-Robson contrast sensitivity assessment *independent of customers' vision health*. In other words, prior studies suggest that even in those instances where customers noted

the presence of glare or shadows, if they were flagged for further assessment they likely possessed some kind of vision condition for which they may have needed examination and treatment by a health professional. They may also have benefitted from education as to the relevance of contrast sensitivity for driving. That said, it is clear from the interviews and the surveys that some 3-Tier staff had reservations regarding how to administer this assessment. Furthermore, it appears that among staff (and perhaps among customers, although there is only indirect evidence for this, based on comments from staff) there was substantial variation in understanding what failure on the contrast sensitivity assessment implied for either customer health or driving safety. Given current best practices standards within the field of program evaluation (Rossi, Lipsey, & Freeman, 2004; Patton, 2008), it was therefore determined to conduct a formal analysis of the distribution of outcomes on the Pelli-Robson chart by location and other potential confounding factors. This analysis is intended to achieve two goals for effective program evaluation. First, it incorporates stakeholder comments into the evaluation of the 3-Tier Pilot. Secondly, this analysis may reveal potential actions necessary to prepare for successful statewide implementation.

Method

Matching and Merging of Datasets, Data Cleaning

In order to estimate the potential impact of chart location on customer outcome, it was first necessary to identify individual customers according to the Pelli-Robson chart on which they were assessed for potential limitations to their contrast sensitivity. This required the matching of information from multiple data sources: the raw paperwork used to process customers during the pilot, a database collected as part of CA DMV's internal auditing for potential fraud, and individuals' permanent driver records.

The raw data files from the pilot—which consisted of electronically scanned pdf documents of the physical paperwork used during customer processing—were first raked for the following variables:

- The customer's driver license (DL) number,
- The field office in which the customer's application for license renewal was initiated,

- The date on which a customer's Tier 1 Score Sheet was completed (this form contained their contrast sensitivity and visual acuity assessment outcomes, as tested in the field office),
- Whether or not the customer had been enrolled into the 3-Tier Pilot as a result of a Driver Safety referral,
- The employee identification number of the person who filled out the 3-Tier Score Sheet,
- The customer's outcome on the Snellen visual acuity test,
- Whether the customer had a "long-standing" vision condition (and so was not tested on the Snellen visual acuity chart),
- The customer's outcome on the Pelli-Robson contrast sensitivity test, and
- Whether or not the customer was referred to a vision specialist for professional evaluation (i.e., given a CA DMV DL62 form).

As a general matter, 3-Tier Pilot procedures called for customers to be tested first on the Snellen chart and secondly—assuming a satisfactory result regarding visual acuity—on the Pelli-Robson chart. If a customer had difficulty passing the department's visual acuity standards[41], they were then usually tested on an Optec 1000 Vision Tester (a small, desk-mounted binocular device). If they continued to fail the visual acuity standard on the Optec 1000, they were then referred to a vision specialist (optometrist, ophthalmologist, etc.) for a comprehensive professional examination. Ultimately, a customer could either pass or fail the visual acuity standard; in the case of failure the customer would typically be required to take an on-road driving test to demonstrate their ability to drive safely despite having some kind of vision condition. In some instances, however, a customer possessed a "long-standing vision condition" for which DMV already had a record; this meant that the customer had failed the visual acuity standard at some point in the past, but had then demonstrated their ability to drive safely despite their vision condition. If this was the case, a notation was made to this effect on the 3-Tier paperwork, and the customer was neither referred to their vision specialist, nor were they required to take another drive test. In other instances, a

[41] Current CA DMV vision standards require a minimum corrected acuity of 20/40 in both eyes, and no worse than 20/70 in one eye alone; if a customer cannot achieve 20/200 vision in at least one eye with correction, they cannot be licensed. If a customer requires correction to achieve the minimum standard (i.e., if they must wear glasses or contact lenses), their license is amended to include a vision restriction. This states that the customer must use their corrective lenses while operating the vehicle class for which they are licensed.

customer may have been suffering from a progressive or unstable vision condition that had been diagnosed by a vision specialist. If this were the case, the length of the renewal period was shortened to some period recommended by their eye doctor—e.g., 2 years as opposed to the standard 5 years—and their record changed to note that they were on a "limited term" license. These customers are also required to take an on-road test to demonstrate their ability to drive safely despite their condition. However, unlike "long-standing condition" customers, "limited-term" drivers are generally required to pass a road test every time they renew (or until such time as their eye doctor diagnoses a stabilization of their condition, and recommends a return to a standard 5-year renewal term).

If a customer passed the department's Snellen standard (or if they had a "long-standing condition"), they were then assessed for potential contrast sensitivity limitations using the Pelli-Robson chart. Here, a customer could achieve one of three outcomes: passing (reading all the letters on lines 1, 4, and 5 of the chart), a "somewhat fail" (missing at least one letter on line 5 of the chart), or a frank/extreme fail (missing at least one letter on lines 1 or 4). If a customer had a frank/extreme fail, they were referred to a vision specialist for a professional examination. If they "somewhat failed," they were assessed at Tier 2 of the 3-Tier process; this involved the receipt of educational materials regarding the implications of a limitation to contrast sensitivity for driving, as well as additional assessment of potential limitations to other driving-relevant skills.

This two-stage process of vision assessment had important implications for the current analysis. Those customers who failed the visual acuity standard—either in a DMV field office at their initial assessment, after evaluation by a vision specialist, or both—typically took the contrast sensitivity assessment in a DMV field office at a different time, at a different location, or as administered by a different field office employee. Customers who were thus referred to a vision specialist for a visual acuity failure were therefore excluded from further analysis for two reasons: first to eliminate a potentially confounding explanation for their outcome on the contrast sensitivity assessment, and secondly because of the difficulty in matching of these customers to the specific Pelli-Robson chart on which they were tested.[42] Also excluded were customers enrolled in the 3-Tier Pilot as a result of a referral from the Driver Safety branch; not all of these

[42] 191 customer-cases were excluded from the analysis for this reason.

customers were assessed for vision limitations, and even if they were so assessed the chart on which they were tested could not be determined with confidence.[43]

Using these data as a starting point, the DMV Audits database was then raked for the following variables:

- Customer DL number,
- The field office number denoting the location in which the customer's application was initiated,
- The employee identification number of the person who cashiered the transaction,
- The unique code for the computer terminal at which the renewal fees were collected,
- The date on which the customer's application for license renewal was initiated (the application date), and
- The customer's birth date.

Individual customer data were then matched on customer DL number, the office at which their renewal application was initiated, and the ID number for the DMV employee who cashiered their transaction (Audits) and filled out the 3-Tier Score Sheet (3-Tier database). Customers who could not be matched on all three of these variables were excluded from the analysis.[44] If customers appeared more than once in the Audits database (this might happen if they made multiple visits to an office before completing their license renewal), the records were matched on the basis of the "application date" (Audits database) and the line/date stamp on their Tier 1 Score Sheet.

Once these matches were made, each customer was assigned a data point denoting the specific Pelli-Robson chart on which they were presumed to have been tested, based upon the terminal at which their renewal fees were collected/cashiered. The match between terminals and charts was established by the author during the final month of the pilot project (October, 2007) by mapping each field office's arrangement of terminals and charts, and through interviews with DMV field office staff regarding their usual use

[43] 116 customer-cases were excluded from the analysis for this reason.

[44] Approximately 10% of customers were thus excluded. There does not appear to be any common variables to the customer-cases so excluded—except perhaps that the line date stamp on their Tier 1 Score Sheet (which contained the employee ID number of the staff person who processed the transaction) was often blurry or otherwise illegible.

of charts at specific computer terminals. Specifically, each staff member working at a given terminal was asked "when seeing a 3-Tier customer, which Pelli-Robson chart do you usually use when sitting at this desk? Do you ever use a chart different from that one?" This mapping allowed for the qualitative determination of three additional key pieces of information. In some cases, DMV employees noted that, when processing customers at a particular terminal, they sometimes used one Pelli-Robson chart and sometimes used another. This may have been due to considerations of glare and shadows, or done for customer convenience—but in any case this meant that a customer who renewed their license at such an "indeterminate" terminal could not, with confidence, be linked to one specific chart. Three terminals (one in South Sacramento, and two in Vacaville) were so identified. All customers seen at these terminals were excluded from the analysis.[45]

Secondly, in some offices specific terminals are marked as the "J ticket" line—i.e., those terminals which have a lower counter, or an associated seating space, for ease of use by customers using assistive mobility devices such as wheelchairs. Customers with various kinds of disabilities are often routed to these terminals by the automatic office ticketing queuing system—though it should be noted that DMV staff working these terminals may see a wide range of customers, depending on the mix of customers being processed that day as well as the individual employee's skill set (license renewal, vehicle registration, commercial transactions, etc.). Thus, although these terminals (and their associated charts) may have been somewhat more likely to be associated with customers possessing disabilities (including visual limitations), the customers seen at these charts were <u>not</u> necessarily excluded from the analysis, though they were assigned a dummy code to test for potential effects. In the final analysis, this variable was dropped due to a lack of statistical significance.

[45] At the South Sacramento office, this meant the exclusion of 102 customer-cases. At Vacaville, this meant the exclusion of 165 customer-cases. The matching of customers to terminals also revealed that some customers were seen at terminals that had not been mapped during the October 2007 series of field office visits. Further investigation determined that these terminals were at the Corrections stations at the Sacramento-South office (37 customers) and at the Sacramento-Broadway office (102 customers). It is likely that these were customers who had made appointments online. While it is unlikely that these customers were substantially different from others, they could not be definitely linked to a specific Pelli-Robson chart. Hence they were excluded from the analysis.

Finally, one additional variable was drawn from the driver record database: whether or not they possessed a "limited-term" license. The latter, as described earlier, is normally issued to customers with progressive/unstable vision conditions; some of these conditions might impact contrast sensitivity.

Customers were excluded from the analysis for a variety of reasons, all related in one way or another to missing data on key variables. A grand total of 1,173 customers were excluded, which constitutes 10.6% of the original pilot cohort sample of 10,999 (which, in turn, does not include "non-correctable" customers or customers who later changed license classes).

Procedure for Quantifying Potential Variation Association with Technician

The staff interviews suggested the existence of some variation in the implementation of the contrast sensitivity assessment across individual technicians. It was therefore necessary to develop a method to quantify this potential variation. However, it was also deemed advisable to develop a method that avoided the potential identification of specific individual staff. To accomplish both of these goals, technicians were first ranked according to the proportion of customers they observed with any limitation at Tier 1 (physical function, visual acuity, contrast sensitivity, or memory). Because customers failed the physical observation and memory recall assessments more rarely than the contrast sensitivity assessment, approximately 80% of the variation in customers' Tier 1 scores derived from their outcome on the contrast sensitivity test. Cross-tabulations were then developed which compared staff *within offices*, and each technician was assigned an adjusted standardized residual.[46] Any technician two or more standard deviations away from the mean fail rate for their office was flagged as an outlier. For lack of a better phrase, technicians that were *above* the mean for their office were labeled "over-orthodox"—meaning they were significantly more likely than their

[46] An adjusted standardized residual is a measure of the observed value (in this case the actual proportion of a given technician's customers observed with limitations at Tier 1) minus the expected value (the mean proportion of customers observed with Tier 1 limitations at that technician's office) divided by an estimate of its standard error. The resulting statistic is expressed in standard deviation units above or below the mean proportion of customers assessed at Tier 2 and 3 for that office.

colleagues to identify customers as possessing potential limitations in physical function, vision, or memory. Those who were *below* the mean were labeled "under-orthodox."[47]

The cross-tabulation was calculated at the office level for the simple reason that pilot staff had expressed concern about variation within their offices, as opposed to across offices. This had the secondary effect of controlling to some extent for variation across offices in the mix of customers. Staff also varied a great deal in the number of customers they processed over the course of the pilot. In calculating the proportion of customers each technician observed with Tier 1 limitations, any employee who processed fewer than 10 customers over the course of the pilot was excluded from the calculation. This had the effect of assuming that staff who saw only a few 3-Tier customers did so in a manner that approximated the contrast sensitivity customer fail rate of the "average" employee. Approximately a third (64/200) of the staff who processed 3-Tier customers were thus excluded from the calculation of potential outliers.[48] Descriptive statistics on the proportion of customers observed with Tier 1 limitations, with the number of outlier technicians (in either direction), are given for each office in Table M4.1.

Customers that were seen at an "indeterminate" terminal—meaning that there was some doubt about which chart they were tested on—were excluded from the analysis, and thus were not part of the calculation of outlier technicians. This had some effect on the results presented in Table M4.1. In most cases a given chart was used by multiple technicians seated at multiple terminals; or in other words, one chart typically served anywhere from two to four terminals. Similarly, in most cases, technicians appear to

[47] This is a somewhat clunky phrase, as it conflates "conforming to established doctrine" (Merriam-Webster) and "adhering to what is commonly accepted, customary, or traditional" (American Heritage). In this particular instance, it conflates (a) the "established doctrine" of 3-Tier procedures, as delivered in training modules to all six pilot offices by Departmental Training Branch and overseen by R&D, and (b) the implementation of those standards at the pilot office level by Field Office management, in the context of "customary and traditional" (not to mention complex and quite demanding) field office procedures for processing different kinds of CA DMV customers. It was not possible to incorporate this distinction into the formal equation modeling, however.

[48] These staff saw a combined total of 218 customers, or approximately 2% of the customers for whom chart locations could be fixed.

Table M4.1: Failure/Passage Rates of DMV Staff Assessing Vision Using Pelli-Robson Contrast Sensitivity Charts, by Office and Employee, with Number of Outlier Employees per Office

Pilot office (*N* of staff)	Fail rate for staff seeing ≥ 10 customers (somewhat + extreme fails combined)	*N* of customers per staff	*N* of under/over/average orthodox staff
Sac.- Broadway (29 staff)	Mean (*SD*): 16.82% (0.37) Minimum: 2.44% Maximum: 41.67%	Mean (*SD*): 65.38 (55.21) Minimum: 10 Maximum: 224	4 / 3 / 22
Carmichael (30 staff)	Mean (*SD*): 12.93% (0.34) Minimum: 0% Maximum: 61.26%	Mean (*SD*): 97.17 (88.33) Minimum: 10 Maximum: 418	6 / 7/ 17
Fairfield (17 staff)	Mean (*SD*): 11.00% (0.31) Minimum: 0% Maximum:27.62%	Mean (*SD*): 57.24 (39.50) Minimum: 10 Maximum: 126	1 / 1 / 15
Folsom (23 staff)	Mean (*SD*): 18.01% (0.38) Minimum: 3.88% Maximum: 34.02%	Mean (*SD*): 75.78 (81.32) Minimum: 11 Maximum: 284	4 / 3 / 16
Sac.- South (27 staff)	Mean (*SD*): 9.47% (0.29) Minimum: 0% Maximum: 28.13%	Mean (*SD*): 58.30 (38.59) Minimum: 10 Maximum: 145	1 / 2 / 24
Vacaville (11 staff)	Mean (*SD*): 14.26% (0.35) Minimum: 2.47% Maximum: 36.00%	Mean (*SD*): 56.09 (33.96) Minimum: 10 Maximum: 124	2 / 2 / 7
Total: 137 staff	Mean (*SD*): 13.93% (0.35)	Mean (*SD*): 70.93 (64.41)	18 / 18 / 101
Total including those with < 10 customers (200 staff)	Mean (*SD*): 13.93% (0.35)	Mean (*SD*): 49.67 (61.86)	

Note: See text for description of procedures for calculating "under-" vs. "over-orthodox" staff. Customers seen at "indeterminate" charts or for whom terminal locations could not be fixed were excluded from these calculations (*n* = 1,173).

have moved around, processing customers at different terminals.[49] However, in a handful of cases employees saw their customers largely at a single, "indeterminate,"

[49] On average, staff saw approximately a quarter of their customers at one terminal. This varied quite a bit from office to office and technician to technician. The degree to which technicians move around the office from desk to desk depended largely on office-specific managerial decisions.

terminal.[50] If these customers were included in the analysis, the coding of outliers would change for three technicians, all in the Vacaville office. The results presented in the rest of the paper were calculated using the identification of outlier staff as presented in Table M4.1 (excluding customers seen at indeterminate terminals); results using calculations based upon alternative methodologies are available upon request from the author.

Variables and Analysis Techniques Used

Once technicians were identified as outliers, individual customers were coded on a series of three dummy variables denoting the "orthodoxy" of the attending staff person: under-orthodox/low fail rate, over-orthodox/high fail rate, and all else coded as "average" and serving as the referent category.

Chart location was entered as a series of dummy variables denoting specific charts on which a given customer was imputed to have been tested. The referent category for each office was that chart whose fail rate was closest to the mean for that office. Descriptive statistics on the fail rate by chart (and by office) are contained in Table M4.2.

The customer's age was coded first as a continuous interval variable in whole years, calculated from the difference between their birth date (as listed in their permanent driver record) and the day on which they enrolled in the 3-Tier Pilot (as determined by day on which they initiated their license renewal application). This was then re-coded as a series of five dummy variables: ≤30 years old, 31-45, 46-60, 61-75, and 76 or older.[51]

A dummy variable noting whether or not a customer was seen at a terminal which typically serviced "J-tickets" (customers potentially in need of extra assistance due to a

[50] It is worth emphasizing that we can draw this distinction for the simple reason that we know the contrast sensitivity assessment outcome (pass/fail) for all customers, regardless of which chart they were seen on. Thus, assuming we can identify which technician processed their application (which is true for approx. 90% of the customers in the pilot), we can then aggregate these results to calculate the pass/fail rate each technician—even if we aren't 100% sure which chart a given customer was tested on.

[51] Various transformations of this variable were tried; this scheme allowed for model convergence while also showing the non-linear effect of age upon likelihood of failing the contrast sensitivity assessment. Other schema (including coding age linearly or ordinally) showed results broadly similar to what are presented here; models not shown, but available upon request from the author.

visible disability) was included in early versions of the analysis. As this had no significant effect, the variable was dropped for parsimony in the final model (results not shown, available upon request from the author).

Customers with "long-standing" vision conditions were combined with those possessing "limited-term" licenses into a dummy variable, with all other customers coded as zero.

Table M4.2: Failure/Passage Rates on the Pelli-Robson Contrast Sensitivity Assessment, by Office and Chart

Pilot office (*N* of charts)	Fail rate (somewhat + extreme fails combined)	Number of customers per chart
Sac.- Broadway (13 charts)	Mean (*SD*): 16.67% (0.37) Minimum: 7.41% Maximum: 22.96%	Mean (*SD*): 149.62 (69.48) Minimum: 38 Maximum: 269
Carmichael (14 charts)	Mean (*SD*): 12.94% (0.34) Minimum: 5.11% Maximum: 42.19%	Mean (*SD*): 210.71 (208.69) Minimum: 45 Maximum: 704
Fairfield (5 charts)	Mean (*SD*): 11.07% (0.31) Minimum: 6.76% Maximum:14.91%	Mean (*SD*): 195.20 (92.69) Minimum: 117 Maximum: 322
Folsom (6 charts)	Mean (*SD*): 18.18% (0.39) Minimum: 12.36% Maximum: 27.46%	Mean (*SD*): 298.00 (238.19) Minimum: 52 Maximum: 737
Sac.- South (6 charts)	Mean (*SD*): 9.26% (0.29) Minimum: 6.55% Maximum: 14.93%	Mean (*SD*): 268.33 (142.82) Minimum: 67 Maximum: 443
Vacaville (5 charts)	Mean (*SD*): 14.44% (0.35) Minimum: 3.70% Maximum: 17.92%	Mean (*SD*): 133.00 (44.28) Minimum: 81 Maximum: 181
Total: 49 charts	Mean (*SD*): 13.93% (0.35)	Mean (*SD*): 202.73 (157.74)

Note: Customers seen at "indeterminate" charts (see text), or for whom terminal locations could not be fixed were excluded from these calculations (n = 1,173).

The dependent variable was coded nominally, with persons who "somewhat failed" or "extreme/frank failed" coded as "1" and those who passed the assessment coded as "0.". There were two few cases of "extreme/frank fails" (n=63) to model the dependent variable ordinally across all six offices (at three of the offices, fewer than three persons extreme failed the contrast sensitivity test).

Because office location was associated statistically with the dependent variable (χ^2 = 77.404, 2-tailed significance at the <.001 level), it was determined that a two-level nested model was required. Offices served as the grouping variable, with all other variables

(age, known indicator of a long-standing or provision vision condition, technician orthodoxy, and chart location) set at level 1 and modeled as fixed effects. Office location was entered as a level-2 random effect. The regression was modeled using the NLMIXED procedure of SAS (version 13), the Quasi-Newton algorithm, and 20 quadrature points.

Limitations of These Data

There exist a number of limitations to these data. First, there is simply no method available to retroactively estimate the primary variable of concern identified by staff—namely, the amount of light associated with a chart. No photometer measurements were taken during the course of the pilot, and there is thus no objective measurement of chart luminance. Also, the pilot took place over the course of five months (June through October). To the degree that staff were often concerned about the amount of *natural* light falling on charts—in the form of sunlight streaming through windows—this changed over the course of any given day as well as over the course of the pilot period, which included both the summer solstice and the vernal equinox. At the most basic level, this means that any effects of chart location on test results are potentially *underestimated*, as the data here do not distinguish between bright and less bright times of day (afternoon versus morning), sunny versus cloudy days, or summer versus autumn. That said, to the degree that *natural* light in particular was a source of concern for staff, these effects should be most likely to appear in those offices that contain substantial banks of windows: South Sacramento and Vacaville in particular. The other offices (Sacramento-Broadway, Fairfield, Folsom, and Carmichael) had, at the time of the pilot, fewer and narrower windows and so, presumably, fewer sources of natural, variable, sunlight.

Secondly, the method by which customers were linked to specific Pelli-Robson charts was based on the assumption that customers could be linked to specific, known charts on the basis of reports from staff regarding their usual practice when processing 3-Tier customers. However, this is an assumption, and hard data do not exist that identify the specific chart on which each individual customer was tested. Nor do data exist regarding which customers complained of glare or shadows, much less whether a given customer was seen on some chart *other* than the one associated with the terminal at which they initiated their renewal transaction, whether because of their concerns about

glare and shadows or some (unknown) other reason. Instead we have survey-based data on the procedures used by staff regarding which chart they normally used when sitting at a specific terminal. This introduces some doubt—specifically an unknown quantity of random error—regarding the data linking individual customers and specific Pelli-Robson charts. More problematically, it is possible that customers with age-based limitations in contrast sensitivity, or other vision conditions affecting their contrast sensitivity, were more likely to complain of glare or shadows. They may consequently have been more likely to be tested on a chart *other* than the one associated with the terminal at which they had initiated their license renewal. This would constitute a form of systematic error in the data, and so potentially bias the estimation of the effects of chart-based variation in customer outcomes on the Pelli-Robson contrast sensitivity test.

Finally, as with many varieties of hierarchical regression estimation (Tabachnick & Fidell, 2007) it was difficult to properly specify a model that would converge according to generally accepted statistical standards. In particular it was necessary to exclude *all* data from the Vacaville office in the final model to ensure model convergence. A review of the data suggested that the problem may derive from the relatively greater degree of overlap in that office between chart location and technician, leading to a kind of colinearity that prevented disentangling these effects from each other. In layperson's terms, staff were often moved from desk to desk at 5 of the 6 pilot offices, which had the statistical benefit of ensuring that (a) most technicians saw customers at multiple charts, and (b) most charts were used by multiple staff. This was less true for Vacaville, where a substantial number of staff saw all or most of their customers at one chart, and some of the charts were used almost entirely by single technicians. This has the unfortunate side-effect of removing from the analysis one of the two offices (South Sacramento being the other) at which complaints of "glare" were most serious, and tied specifically to the existence of a large bank of west-facing windows.

Thus, the results presented here cannot entirely confirm (or disconfirm) the robustness of the Pelli-Robson chart by location—which is to say under variable light conditions or with other (undetermined) factors affecting customer outcomes. However, the results can provide some indication of the degree to which the use of these charts in an agency setting is subject to certain limitations potentially associated with inherently variable light conditions, but perhaps more importantly with (as we shall see) variation

associated with individual staff. The implications of these limitations are discussed in the conclusions.

Results

Descriptive Statistics

As can be seen in Tables M4.1 and M4.2, there exists substantial variation in failure rates on the contrast sensitivity assessment both by technician and by chart. Across offices, chart fail rates varied from a low of just under 4% to a high of just over 40% while technician fail rates ranged anywhere from 0 to 60%. To some degree the variation *across offices* is to be expected; CA DMV field offices draw from substantially different customer populations with different constellations of variables that might affect the overall pass/fail rate on the contrast sensitivity assessment. This is particularly the case for age; as can be seen in Table M4.3, the mean customer age varied by more than a decade, from a low of 49.3 years (Sacramento-Broadway) to a high of 59.8 years (Carmichael).

It is less clear why this variation in fail rates should occur at the chart or technician level, however. Each pilot office has an automatic queuing system (the DMV Customer Service Queuing Management System, or DMVQ) that routes customers to technicians and terminals based on algorithms set by the Office Manager at the start of each workday. These algorithms include factors such as the overall (office) wait time, the speed/productivity of a given technician, a given technician's training in processing different types of transactions, and the number of customers in line awaiting different transaction types. Transaction types include (among others): "basic" license and identification cards, individual vehicle registration, and commercial/dealer vehicle registration.[52] To the degree that these algorithms might affect the "mix" of customers that a staff person processes, this effect would largely derive from whether or not a staff member was trained to process transactions *other* than basic ID card/driver license original and renewal transactions (which, in the DMVQ system, are designated under

[52] Most commercial/dealer transactions now occur at Consolidated Registration and Industry Service Centers. At the time of the 3-Tier Pilot, however, some of these transactions still occurred at the field office sites used by the general public.

Table 4.3: Descriptive Statistics

Pilot Office (N)	Mean age, in years (*SD*)	N of customers						
		Failed on fog chart (%)	Seen by under-orthodox staff (%)	Seen by over-orthodox staff (%)	Seen by average staff (%)	Seen at "J ticket" terminal[a] (%)	Seen at moved chart[s] (%)	Limited term license or "long-standing vision condition"[c]
Sac-B'way (1945)	49.27 (20.26)	324 (16.66%)	288 (14.81%)	168 (8.64%)	1489 (75.56%)	142 (7.3%)	N.A.	28
Carmichael (2951)	59.81 (19.88)	382 (12.94%)	1340 (45.41%	552 (18.71%)	1059 (35.89%)	N.A.	192 (6.5%)	30
Fairfield (976)	52.39 (20.77)	108 (11.07%)	122 (12.50%)	105 (10.76%)	749 (76.74%	86 (8.8%)	N.A.	3
Folsom (1788)	57.44 (20.15)	325 (18.18%)	446 (24.94%)	470 (26.29%)	872 (48.77%)	138 (7.7%)	N.A.	30
Sac-South (1609)	54.94 (20.34)	149 (9.26%)	80 (4.97%)	72 (4.47%)	1457 (90.55%)	N.A.	N.A.	4
Vacaville (665)	54.62 (21.03)	96 (14.44%)	174 (26.17%)	103 (15.49%)	388 (58.35%)	200 (30.1%)	N.A.	7
Total (9934)	55.46 (20.60)	1384 (13.93%)	2450 (24.66%)	1470 (14.80%)	6015 (60.54%)	566 (5.7%)	192 (1.9%)	102

Note: See text for description of the calculation of "over-" versus "under-" versus "other" staff. "Other" staff includes technicians who processed fewer than 10 customers. N.A. indicates that there were no customers that fit into this cell.

[a] "J ticket" terminals are those specially designated for processing customers in need of physical assistance. Not all customers seen at such terminals require such assistance. Customers with "J tickets" are seen at any terminal in the Carmichael and South Sacramento offices.

[b] One chart was moved at the Carmichael office, at the request of staff concerned about light conditions. This chart was moved on approximately August 1st, 2007.

[c] This category includes both (i) customers identified on the 3-Tier Score Sheet as possessing a "long-standing" vision condition and (ii) customers who, according to the CA DMV driver record database, held a limited-term license. Many customers with either of these characteristics were excluded during data-cleaning because there was some doubt as to which chart they were tested on (e.g., if they had multiple visits to the office due to a DL62 referral to a professional vision specialist).

one category). As 3-Tier customers constituted a subset of basic license renewal transactions, any staff member trained in basic driver license and ID card transactions (and identified as such in the DMVQ system) theoretically may have seen 3-Tier customers. Certainly the *number* of 3-Tier transactions any individual technician saw may have varied—as a product of what kinds of customers happened to be in the office on a given day, the DMVQ algorithm set by the manager, but especially in regards to how many technicians shared the work of processing basic license/ID card original and renewal transactions. However, it should not have been possible within the limits of the automatic queuing system to route different *types* of 3-Tier customers (e.g., customers of different ages, or customers with long-standing vision conditions) to different technicians, for the simple reason that all 3-Tier customers were, by definition, basic license renewal transactions.[53] In other words, assuming that a technician was designated by the DMVQ system to receive basic DL renewal and origination transactions, they may have seen 3-Tier customers. But *which* 3-Tier customers they saw—and thus, whether those customers were more or less likely to have a vision condition as flagged by the Pelli-Robson contrast sensitivity chart—should have been essentially randomized *at the office level*. Thus, at least within a given office, the variation in technician fail rates should not have been a product of variation in customers seen by that technician. The same logic holds for charts; to the degree that a given chart typically drew customers from multiple terminals, and to the degree that most staff moved around the office (processing customers at different terminals and charts), the variation in chart fail rate should not have been a product *either* of the customers seen at that chart *or* of the mix of technicians who used the chart.

Predictive Analysis

Table M4.4 presents the results of the hierarchical logistic regression equation predicting individual customer likelihood of failing the contrast sensitivity assessment. A number of statistical relationships stand out.

[53] The one exception consisted of customers enrolled in 3-Tier as a result of a referral to the Driver Safety Branch. However, when these customers visited field offices they did so as a result of an appointment made directly with an LRE, and so thus were not part of the queuing system.

Table M4.4: Hierarchical Logistic Regression, Predicting Odds Ratios (Expβ) of 3-Tier Customer Failure (Somewhat or Extreme Fail) on the Pelli-Robson Contrast Sensitivity Chart

		β (s.e.)	Expβ (95% c.i.)
Age, in years (ref. cat. is ≤30 years old)		(.000)	
	31-45 years old	0.216 (.282)	1.241 (0.567 – 2.716)
	46-60 y.o.	1.030 (.273)*	2.800 (1.314 – 5.969)
	61-75 y.o.	3.259 (.219)**	26.021 (14.167 – 47.794)
	76+ y.o.	3.916 (.223)**	50.219 (27.029 – 93.317)
Confirmed vision disorder	On limited term license or has "long-standing vision condition"	2.499 (.287)**	12.168 (5.489 – 26.969)
Staff orthodoxy (ref. cat. is average)		(.000)	
	Customer tested by "under-orthodox" staff	-0.730 (.137)**	0.482 (0.330 – 0.705)
	Customer tested by "over-orthodox" staff	1.415 (.104)**	4.114 (3.085 – 5.488)
Broadway charts	11 of 13 charts not significant		
	Chart L	0.761 (0.346)†	2.141 (0.820 – 5.588)
Carmichael charts	12 of 14 charts not significant		
	Chart F	1.122 (0.327)*	3.072 (1.240 – 7.610)
	Chart H	0.907 (0.368)†	2.477 (0.891 – 6.885)
Fairfield charts	5 of 5 charts not significant		
Folsom charts	2 of six charts not significant		
	Chart A	-0.667 (0.266)†	0.513 (0.245 – 1.074)
	Chart E	-0.585 (0.268)†	0.557 (0.265 – 1.172)
	Chart F	-0.604 (0.227)†	0.546 (0.291 – 1.026)
South-Sacramento charts	6 of 6 charts not significant		
Office	(entered as random effect)	0.173 (0.152)	1.188 (0.780 – 1.1811)

Note: Effect sizes and odds ratios for non-significant charts are not reported in this table, but were included in the equation. *N* = 9268 (nested within 5 subjects, Vacaville office excluded, see text for explanation). -2 Log Likelihood = 5294.3. β of intercept = -18.595** (expβ = <0.001)

† Significant at the .10 level

* Significant at the .05 level

** Significant at the .01 level

Age clearly has the largest relationship to the likelihood of failing the contrast sensitivity assessment. When treated as a ratio variable (regression not shown, available upon request from the author) the effect seems small, with a β value of 0.09. However, the effect of age on likelihood of failing the contrast sensitivity assessment is in fact somewhat non-linear. To demonstrate this more clearly, age was coded as an ordinal variable in 15-year increments starting at age 30. As shown in Table M4.4, the odds ratio

of failing become significantly higher around age 45, and increase substantially thereafter.[54] In fact, for the oldest cohort (those 76 and older), the odds of failing this assessment were about 50 times greater than the odds for someone 30 or younger.

The next largest effect, after age, consists of whether or not a customer had been previously identified as having either a "long-standing" vision condition, or was on a limited-term license for a progressive vision condition. The odds of failing the contrast sensitivity assessment for this group are about 12 times greater compared to individuals on full-term licenses and with no long-standing vision condition on record. Given previously-published findings in the ophthalmological literature, these findings regarding the effects of age and medical conditions (as reflected by DMV-recorded stable or progressive vision conditions) are entirely expected.

Office was entered as a level-2 random effect, and the estimate of the effect size is not significant. This suggests that there does not exist further unexplained variance at the office level.

Once these effects have been included, any pure effects of chart location largely (though not entirely) disappear. In two offices (Fairfield and South Sacramento) no charts are significantly different from the office mean. In one office (Broadway), a single chart approaches, but does not reach, conventional standards of statistical significance (α <.10). At another office (Folsom) half the charts (3 out of 6) approach significance (α <.10). At a third office (Carmichael), one chart approaches significance (α <.10) while a second exceeds significance at the α <.05 level. For those charts where α <.10, the confidence intervals of the β values include 0 (this appears in the exponentiated β values column as a confidence interval range that includes the value of 1). This indicates the possibility that there is, in fact, zero effect on customer outcomes associated with the location of these charts.

There is one chart (at the Carmichael office) were α <.05. This chart was moved approximately halfway through the pilot due to concerns about shadows from nearby Snellen charts that (by report) made the letters on the Pelli-Robson chart relatively more difficult to read. Controlling for other variables, we find that there appears to be an

[54] Coding age as a ratio or ordinal variable does not materially change the β coefficients of other variables in the model.

effect associated with this chart, and in the expected direction—customers tested on that chart appear to have been substantially and significantly more likely to fail the contrast sensitivity assessment. It is not clear, however, whether we can be precisely sure that the statistical effect we find here is a pure product of luminance. In the first instance, the fail rate did not statistically change after it was moved; it remained at approximately 41%. A χ^2 calculated for the fail rate on this chart before and after August 6th (the approximate date of the move) yields no significant differences.[55] Perhaps more to the point, of the 196 customers seen at this chart, 62% (122) were processed by two technicians identified as "over-orthodox" (having a high fail rate on all Tier 1 assessments). One of these two technicians saw nearly all of their customers at this chart, and so the effects of technician orthodoxy and chart location cannot be disentangled in this case. However, the other technician saw only one-third of their customers at the chart that was moved. This technician saw the majority of their customers at a different chart, one that had a fail rate statistically indistinguishable from the overall mean for the Carmichael office. There, this second technician failed an even greater proportion of their customers than they did at the chart that was moved.

Of greater statistical impact than which chart one was tested on was the effect of which technician processed the transaction. Seeing an "under-orthodox" (low fail rate) technician was associated with 50% decrease in the odds of failing the contrast sensitivity assessment; by contrast, seeing an "over-orthodox" (high fail rate) technician was associated with a four-fold increase in the odds of failing, as compared to being assessed by the referent category of technician. Some technicians were substantially more likely to record failures—in excess of 1/3 of their customers, in some cases. On the other hand, some technicians never failed anyone at all (including, in one case, a technician who processed almost 80 applications). Unfortunately, it was not statistically possible to model interaction terms (the models failed to converge); it is therefore impossible to say whether technician orthodoxy varied by chart location, or customer age.

Conclusions

How robust are the Pelli-Robson contrast sensitivity charts used as part of the 3-Tier Assessment System Pilot? To the degree that customer outcomes on the chart can be

[55] The fail rate on this chart prior to August 6th was 40.6%. After the move, it actually *increased* to 43.1%.

linked to pre-existing vision conditions (stable or progressive) or to age, it appears that the charts operate substantially as predicted by the clinical and research literature. These two variables exert the largest effect sizes on individual likelihood of failing the contrast sensitivity screen. However, even controlling for these variables, we cannot eliminate the potential for chart-level variation. Moreover, it appears that there exists some substantial variation associated with which technician administered the assessment.

It cannot be precisely determined from these data what produces the variation by chart location. This may be a matter of variable light conditions, as suggested by staff. However, the empirical outcomes demonstrated in these data do not match up in any predictable way with staff concerns. First, customer outcomes are less reliably a product of which chart they were tested on—at least as indicated by these data—as compared to which technician administered the assessment. This suggests that *if* variation in light levels affected customer outcomes, this variation occurred not so much between charts (which staff noted and were concerned about) as between staff.

Given that no variables were included in this analysis which would have captured the time of day, the proximity to a window (or the direction faced by a proximate window), or even the season of the year, it is still possible that luminance differences across charts are not captured properly in the final model. That said, the most brightly-lighted office (Sacramento-South), with the largest number of windows and consequently the most complaints of glare, had the lowest overall fail rate. This suggests that whatever else complaints of "glare" may signify, they do not correlate in any direct way with a higher rate of failing the contrast sensitivity assessment screen (which was the brunt of staff concern). Other offices, with fewer windows and thus lower levels of natural light, had either a higher fail rate (as at Sacramento-Broadway), or a substantially greater spread of outcomes by chart and by technician (as at Carmichael). The outcomes at these offices—especially at Carmichael—may be due to the presence of shadows created by nearby visual acuity charts. It appears, for instance, that the one chart that was moved at the Carmichael office in fact had an abnormally high fail rate compared to other charts in that office. However, this abnormally high fail rate may plausibly be attributed to the over-orthodoxy of the technicians who processed customers at the terminals associated with this particular chart. The contribution of variation in luminance to chart- (or office) level variation in customer likelihood of failing the contrast sensitivity

screen must therefore remain speculative. No formal measurement of light levels was ever taken of any chart at any office, and thus the effects of variability in natural sunlight or indoor artificial light cannot be directly or precisely quantified.

The absence of any substantial variation on customer outcomes, as predicted by chart location may be a product of either random or systematic error in the linking of individual customers with specific charts. The data supporting these links derive from indirect evidence, namely person-to-person surveys conducted by the author with individual staff regarding their usual procedures in testing 3-Tier customers on the Pelli-Robson chart. Thus, there is no direct evidence in the case of any individual customer as to which chart they were tested on. Instead there is hard evidence regarding the terminal at which they initiated their transaction, and on the basis of this an imputed link to the chart normally used by the technician operating that terminal. There is no evidence at all, unfortunately, regarding whether or not a given customer complained of glare or shadows on the chart on which they were first tested, or whether or not the customer was moved to a second chart as a result of their concerns. As a subsidiary point, there is no evidence regarding how many customers complained of glare/shadows, nor of how many customers switched charts. To the degree that substantial numbers of customers switched charts, and especially if particular kinds of customers (i.e., those with conditions affecting their contrast sensitivity) did so, the estimates presented here of the effect size, significance level, and direction of effect of chart location on customer outcomes may be biased or otherwise inaccurately specified.

As it happens, the likelihood of failing the contrast sensitivity screen was at least as much a product of which technician processed a customer transaction as it was a product of which office that customer visited; furthermore, technician variation had a stronger relationship to customer outcomes than which chart the customer was tested on. It is difficult, for at least two reasons, to attribute variation by technician to differences in the type of customer processed by a given staff member. First, age alone, along with age-related vision conditions (e.g., macular degeneration, cataracts, and diabetic retinopathy) are the primary sources of contrast sensitivity declines in individuals, according to both the results presented here and the clinical and research literature (see Elliot, 1998; Hawkins, Szlyk, Ardickas, Alexander, & Wilensky, 2003; Stavrou & Wood, 2003). As these two variables are included in the final model, any technician variation produced by differences in the mix of customers seen (i.e., in their

average age or their likelihood of possessing contrast-related vision disorders) is statistically controlled for. Thus the effects of technician variation found in the final model are *net* of such differences (if they exist). Secondly, the DMVQ automated ticketing system employed by California DMV field offices during the 3-Tier Pilot routed customers to staff terminals on the basis of *transaction type* rather than *customer type*, and did not differentiate among 3-Tier customers (all of whom were, by definition, conducting basic license renewal transactions). Even if some customers had the kind of assistance needs that resulted in the issuing of a "J-ticket," this had no statistically significant effect on contrast sensitivity outcomes (equations not shown, available upon request from the author). For these reasons, it is unlikely that staff were seeing different kinds of 3-Tier customers (at least within a given office).

It is also difficult to attribute technician variation to differences in light levels in their work areas (i.e., on the charts that they used to process customers). Among staff that processed at least 10 customers, technicians used an average of 4.5 charts over the course of the pilot project and saw approximately 22% of their customers at any one chart. While the number of charts used varied from office to office, in at least four offices no technician used fewer than two different charts.[56] Thus, because of DMV field office procedures entirely unrelated to 3-Tier, staff members moved from terminal to terminal, processing customers under (potentially) quite different lighting conditions. This has the statistical effect of removing some (though no all) of the potential overlap between chart location and technician, and supports the argument that in most cases whatever is being measured here as "technician orthodoxy" was independent of the location at which a given staff person conducted their work.

Finally, the link between individual customers and specific technicians depends upon much more reliable data than that linking customers to charts. When conducting a 3-Tier transaction, each technician would individually stamp the paperwork recording that customer's outcome on the contrast sensitivity screen. In addition, every time a customer tendered their renewal fees, this produced a transaction record which included the identification number for the technician cashiering the fees. According to

[56] The mean number of charts used by technicians varied from a high of 6.1 (Sacramento-Broadway) to a low of 3.5 (Folsom). All technicians at Sacramento-Broadway, Carmichael, Fairfield, and Sacramento-South used at least two charts.

training procedures for the pilot, these two activities—conducting the contrast sensitivity assessment and cashiering the renewal fees—occurred within a very short span of time, and by the same person. While a customer may have switched charts, only rarely would they have switched technicians.[57] For these reasons, the reservations regarding data interpretation (mentioned above) regarding chart location do not apply with the same force regarding the estimated effects of what is here termed "technician orthodoxy."

What are the implications of these findings? At the very least they suggest the likely necessity of further research. Given that the 3-Tier Pilot occurred in a naturalistic, quasi-experimental, agency environment, it would have been difficult to eliminate (let alone measure and control for) all confounding sources of influence over an individual's likelihood of passing the contrast sensitivity screening test. It should certainly be possible to conduct research projects in the future that would more rigorously address these questions. In the tightly-controlled lab environments used in previously published studies, it has been shown that luminance has little effect independent of age and vision health conditions (Zhang, Pelli, & Robson, 1989; Cox, Norman, & Norman, 1999; Rovamo, Kukkonen, Thppana, & Näsänen, 1993). Testing in an agency setting—for example, taking photometer readings in real time while customers are taking the assessment screen—would potentially require more sensitivity to the vagaries of quasi-experimental research.[58] However, the potential benefits of capturing real-time variation in seasonal light differences, time-of-day differences, proximity to windows (and direction of said windows), etc., would lend enormous face validity to any findings regarding the utility of this instrument in an inherently variable agency environment. It might also allow for distinguishing between chart-associated and staff-associated differences in outcomes, which in this study are to some extent conflated (as in the main equation) or left unanalyzed due to excess colinearity (as in the data collected from the Vacaville office).

[57] Circumstances where one staff person administered the contrast sensitivity test, but a different staff person cashiered the transaction would include (for instance) driver safety referrals. These were excluded from the analysis at an early stage of data-cleaning. This might also occur in those rare instances where a 3-Tier customer was accidentally "queued" to a technician who was not trained in 3-Tier procedures. This was normally prevented by the DMVQ algorithm set by the Office Manager each day.

[58] It would also require attention to potential Hawthorne effects, especially as regards staff implementation of training protocols.

Secondly, these findings introduce a cautionary note regarding the ease with which clinically-reliable assessments may be transferred for use within an agency setting. As noted elsewhere in this process analysis, the Pelli-Robson chart was one of the most popular elements of the 3-Tier Pilot. In interviews (see Module 2 of this Appendix) staff reported that the chart possessed an intuitive, face-valid connection to safe driving that was easily explained to customers. Regardless of the ease with which staff administered the contrast sensitivity assessment using this tool, and independent of any face (or criterion) validity, the data reported here indicate the existence of substantial variation across staff in the reliability of the administration of this assessment.

To some extent this question of chart robustness must remain predicated on the ultimate traffic safety utility of this screening tool. Whether the chart is robust (or not) only matters if the Pelli-Robson chart is useful for identifying drivers at risk of crashing, for identifying drivers in need of vision-related health referrals, and/or for identifying drivers in need of education to improve their driving skill. Given that prior authors have already found this screening assessment useful for precisely these questions (Hennessy, 1995; Owsley, et al., 1991), it would seem appropriate to suggest at least two possible avenues for controlling any concerns about chart-level or technician-level variation associated with the likelihood of passing this screening assessment.

One possibility lies in the content of training offered to staff in the use of the Pelli-Robson contrast sensitivity chart. For instance, the clinical literature indicates that both sensitivity to glare and the effect of glare on contrast sensitivity are indicative of (a) age-related changes to vision health, and (b) potential vision problems that may impact driving. Thus, when faced with customer concerns about glare on charts, staff may note—assuming they themselves have access to this information—that this is a driving-relevant problem (and therefore not "a problem with the chart") that is best addressed by evaluation from a clinical professional. In other words, the screening nature of this assessment tool may be multi-dimensional: not only are there multiple outcomes possible on the chart (pass, somewhat fail, and frank/extreme fail), but discomfort with taking the test may itself be an indicator of potential vision difficulties.

A second possibility lies in the use of a more controlled environment for testing purposes. According to current CA DMV field office procedures, if a customer fails the Snellen visual acuity standard they are retested on an Optec 1000 Vision Tester. This is a

small binocular device, mounted on an office counter, which allows the customer to eliminate distracting visual cues while at the same time allowing the technician to control more precisely factors such as which eye the customer is using, and the distance at which the customer takes the test. Similar devices exist for the testing of contrast sensitivity, and may be adapted for CA DMV field office use. The primary utility of such a device would likely lay less in the creation of a controlled environment for testing purposes, and more in the creation of an option for staff to use when faced with customer complaints (or their own skepticism). While it is unlikely that variation in light conditions directly influence outcomes on a contrast sensitivity screening assessment, having a controlled environment may provide a necessary alternative for customers wanting a second opportunity to pass this assessment screen.

REFERENCES

Anderson, S. J., & Holliday, I. E. (1995). Night driving: Effects of glare from vehicle headlights on motion perception. *Ophthalmic and Physiological Optics, 15*, 545-551.

Ball, K. K., Beard, B.L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and Visual Search: Expanding the useful field of view. *Journal of the Optical Society of America A, 5*, 2210-2219.

Ball, K. K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., Morris, J. N., Rebok, G. W., Smith, D. M., Tennstedt, S. L., Unverzagt, F. W., & Willis, S. L. (2002). Effects of Cognitive Training Interventions With Older Adults: A randomized controlled trial. *Journal of the American Medical Association, 288*, 2271-2281.

Brown, B., & Garner, L.F. (1983). Effects of luminance on contrast sensitivity in senile macular degeneration. *American Journal of Optometry and Physiological Optics, 60*, 788-793.

California Department of Motor Vehicles. (2009). *Strategic Plan 2009*. Sacramento: Author.

Clay, O. J., Wadley, V. G., Edwards, J. D., Roth, D. L., Roenker, D. L., & Ball, K. K. (2005). Cumulative Meta-Analysis of the Relationship Between Useful Field of View and Driving Performance in Older Adults: Current and Future Implications. *Optometry and Vision Sciences, 82*, 724-731.

Cox, M. J., Norman, J.H., & Norman, P. (1999). The effect of surround luminance on measurements of contrast sensitivity. *Ophthalmic and Physiological Optics, 19,* 401-414.

Edwards, J. D., Wadley, V. G., Myers, R. S., Roenker, D. L., Cissell, G. M., & Ball, K. K. (2002). Transfer of a Speed of Processing Intervention to Near and Far Cognitive Functions. *Gerontology, 48,* 329-340.

Elliott, David B. (1998). Assessment of Patients with Age-Related Cataract. *Ophthalmic and Physiological Optics, 18,* S51-S61.

Freund, B., Colgrove, L., Petrakos, D., & McLeod, R. (2008). In My Car the Brake is On the Right: Pedal Errors Among Older Drivers. *Accident Analysis and Prevention, 40,* 403-409.

Gebers, M. A., & Roberts, R. A. (2004). *Characteristics of Negligent Operators in California*. (Report No. 209). Sacramento: California Department of Motor Vehicles.

Goodsell, C.T. (1985). *The Case for Bureaucracy: A Public Administration Polemic*. Chatham NJ: Chatham House Publishers, Inc.

Hakamies-Blomqvist, L., & Wahlstrom, B. (1998). Why do Older Drivers Give up Driving? *Accident Analysis and Prevention, 30,* 305-312.

Hawkins, A. S., Szlyk, J. P., Ardickas, Z., Alexander, K. R., & Wilensky, J. T. (2003). Comparison of Contrast Sensitivity, Visual Acuity, and Humphrey Visual Field Testing in Patients with Glaucoma. *Journal of Glaucoma, 12*, 134-138.

Hennessy, D. F. (1995). *Vision Testing of Renewal Applicants: Crashes predicted when compensation for impairment is inadequate*. (Report No. 152). Sacramento: California Department of Motor Vehicles.

Hennessy, D. F., & Janke, M. K. (2005). *Clearing a Road to Driving Fitness by Better Assessing Driving Wellness: Development of California's Prospective Three-Tier Driving Centered Assessment System: Summary Report*. (Report No. 215). Sacramento: California Department of Motor Vehicles.

Janke, M.K. (2001). Assessing Older Drivers: Two Studies. *Journal of Safety Research 32*, 43-74.

Katz, D., Gutek, B. A., Kahn, R. L., & Barton, E. (1975). *Bureaucratic Encounters: A Pilot Study in the Evaluation of Government Services*. Ann Arbor, MI: The University of Michigan Institute for Social Research.

Koustanaï, A., Boloix, E., Van Elslande, P., & Bastien, C. (2007). Statistical Analysis of 'Looked-But-Failed-To-See' Accidents: Highlighting the involvement of two distinct mechanisms. *Accident Analysis and Prevention, 40,* 461-469.

Lee, E. S., & Forthofer, R. N. (2006). *Analyzing Complex Survey Data*. Thousand Oaks, CA: Sage Publications.

Miller, T. I., & Miller, M. A. (1991). Standards of Excellent: U.S. Residents' Evaluations of Local Government Services. *Public Administration Review, 51,* 503-514.

National Commission on the Public Service. (1989). *Leadership for America: Rebuilding the Public Service*. Washington DC: The Commission.

Neal, M. B., Baggett, S., Sullivan, K. A., & Mahan, T. (2008). *The Older Driver In Oregon: A Survey of Driving Behavior and Cessation. Final Report*. (Report No. FHWA-OR-RD-08-08). Salem: Oregon Department of Transportation Research Unit.

Osborne, D., & Gaebler, T. (1992). *Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector*. Reading, MA: Addison-Wesley.

Owsley, C., Ball, K. K., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1991). Visual/Cognitive Correlates of Vehicle Accidents in Older Drivers. *Psychology and Aging, 6*, 403-415.

Owsley, C., Sekular, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research, 23,* 689-699.

Patton, M. Q. (2008). *Utilization-Focused Evaluation* (4th ed.). Los Angeles, CA: Sage Publications.

Poister, T. H., & Henry, G. T. (1994). Citizen Ratings of Public and Private Service Quality: A Comparative Perspective. *Public Administration Review, 54*, 155-160.

Puell, M, C., Palomo, C., Sánchez-Ramos, C., & Villena, C. (2004). Mesopic contrast sensitivity in the presence or absence of glare in a large driver population. *Graefe's Archive for Clinical and Experimental Ophthalmology, 242,* 755-761.

Rassow, B. (1999). Effect of luminance on contrast sensitivity and glare in the mesopic range. *Klinische Monatsblätter für Augenheilkunde, 214,* 401-406.

Richards, O. W. (1977). Effects of luminance and contrast on visual acuity. *American Journal of Optometry and Physiological Optics, 54,* 178-184.

Roenker, D. L., Cissell, G. M., Ball, K. K., Wadley, V. G., & Edwards, J. D. (2003). Speed-of-Processing and Driving Simulator Training Result in Improved Driving Performance. *Human Factors, 45*, 218-233.

Rossi, P. H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage Publications.

Rovamo, J., Kukkonen, H., Thppana, K., & Näsänen, R. (1993). Effects of luminance and exposure time on contrast sensitivity in spatial noise. *Vision Research, 33,* 1123-1129.

Sheedy, J. E., Smith, R., Hayes, J. (2005). Visual effects of the luminance surrounding a computer display. *Ergonomics, 48,* 1114-1128.

Shumaker, N. C. (1994). *The California Driver Performance Evaluation Project: An evaluation of the current driver licensing road test*. (Report No. 151). Sacramento: California Department of Motor Vehicles.

Siren, A., & Hakamies-Blomqvist, L. (2006). Does Gendered Driving Created Gendered Mobility? Community-Related Mobility in Finnish Women and Men Aged 65+. *Transportation Research: Part F, 9,* 374-382.

Sloane, M. E., Owsley, C., & Jackson, S.A. (1988). Aging and luminance-adaptation effects on spatial contrast sensitivity. *Journal of the Optical Society of America A: Optics and Image Science, 5,* 2181-2190.

Stavrou, E.P., & Wood, J. (2003). Letter Contrast Sensitivity Changes in Early Diabetic Retinopathy. *Clinical and Experimental Optometry, 86,* 152-156.

Sturgis, S. P., & Osgood, D. J. (1982). Effects of glare and background luminance on visual acuity and contrast sensitivity: Implications for driver night vision testing. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 24,* 347-360.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, MA: Allyn and Bacon.

West, C. G., Gildengorin, G., Haegerstrom-Portnoy, G., Lott, L.A., Schneck, M. E., & Brabyn, J. A. (2003). Vision and driving self-restriction in older adults. *Journal of the American Geriatrics Society, 51,* 1348-1355.

Zhang, L., Pelli, D.G., & Robson, J.G. (1989). The effects of luminance, distance, and defocus on contrast sensitivity as measured by the Pelli-Robson chart. *Investigative Ophthalmology and Vision Science, 30,* 406.

SUB APPENDIXES

SUB APPENDIXES

Sub-Appendix A: Survey Instrument

DMV Staff Survey for 3-Tier Pilot

In keeping with Vehicle Code Section 1659.9, the California DMV is trying out a new assessment system for customers renewing their driver license. Your hard work, cooperation, and input have been vital to the implementation of the "3-Tier" Pilot Program.

As someone who actually implemented the pilot, you have a unique understanding of what worked about the project, as well as what could or should have been different. This survey is intended to gather your input on the pilot.

Please take a few minutes to answer the attached questions. Your response to this survey is anonymous. Only the overall findings of the survey will be used (i.e., your answers will be grouped with those of others).

You will note that the first question asks for your job classification (MVFR, LRE, manager, etc.): this information is useful so that we can better understand how 3-Tier affected different job categories in different ways.

You may, if you so desire, write any additional comments you feel appropriate at the end of the survey.

In addition to this survey, we are also conducting interviews with a select group of staff, to get more "at-length" feedback about the process. **If you would like to participate in the interview process, please write your name and office below** (this sheet will be detached from your survey to assure anonymity). Someone from Research and Development will contact you in the near future regarding scheduling the interview.

Thank you for helping us build a better DMV!

Name: _____ **Office:** _____
(write your name and office here, if you would like to be contacted for an interview with R&D about the 3-Tier Pilot)

DMV Staff Survey for 3-Tier Pilot

1.) What is your job classification? (circle one):

MVFR	LRE	SMVT	Hearing Officer	3-Tier Manager I	Manager (other)

Other job classification (please specify): _____

2.) On average, about how many 3-Tier customers did you see (or process) over the course of a typical day? (circle one)

Very few (1-2) customers per day	3-6 customers per day	7-12 customers per day	13-20 customers per day	More than 20 customers per day

3.) Think for a moment about the various forms and other paperwork that you may have used to collect data on and to process 3-Tier customers (for instance, the Score Sheet or the Tracking Sheet). Is there anything <u>specific</u> that you would suggest for how to improve these forms? Please be as specific as you can in your suggestions.

_____
_____
_____

4.) Aside from the forms and paperwork, if you could pick one part of the **process** that you think should be changed somehow, what would it be, and why?

_____
_____
_____

5.) In your experience, what impact has this pilot had on customer service?

Very Positive	Positive	Neutral	Negative	Very Negative

a. Comments

_____
_____
_____

6.) Did you receive feedback (positive or negative) from customers regarding the 3-Tier Pilot?

a. Yes No
b. Comments

_____
_____
_____

7.) How fair do you think the 3-Tier process was? (circle one)

Very Fair	Fairly Fair	Somewhat Fair	Not Very Fair

a. If you have concerns about the fairness of 3-Tier, is there anything you can think of that would improve the program in this respect?

_____
_____
_____

8.) Did you attend a formal training session for the 3-Tier Pilot?

a. Yes No
b. If the answer is "No," were you given one-on-one training (e.g., by a manager, FOD staff services, or someone from R&D)?
    i. Yes No

9.) How useful did you find the formal training, knowing what you do now about the process? (circle one)

Extremely Useful	Very Useful	Somewhat Useful	Of Limited Usefulness

a. Comments

_____

_____

10.) Once the pilot was implemented, about how long did it take for you to get used to 3-Tier processes and procedures? (circle one)

I was comfortable with 3-Tier procedures right from the start	More than a week but less than a month	1-2 months	More than 2 months	I never really got used to 3-Tier procedures

11.) Think about a "typical" week over the course of the last month or two (so, sometime in September, for example). How often did you find yourself asking someone (a co-worker, a manager, someone from R&D) for advice or help about some aspect of 3-Tier? (circle one)

Several times per day	About once per day	About once per week	Hardly Ever

a. If you answered "several times per day" or "about once per day," what is an example of the kind of question (or questions) you would typically ask? (if you can think of more than one question, please write all of them down)

_____

_____

_____

Finally, if you have any further insights, suggestions, or constructive input, please use the space below to write any comments you think would be helpful to understanding how the 3-Tier Pilot worked.

THANK YOU for all your input on this survey!

Sub-Appendix B: Interview Protocol for the 3-Tier Pilot Process Analysis

Over the past few weeks and months, you have worked in a DMV office where we tested out the 3-Tier Pilot. Your hard work, cooperation and input have been vital to the implementation of this project. For that, we at Research and Development (R&D) would like to thank you for your participation in 3-Tier.

As someone who actually implemented the pilot, you have a unique perspective on what worked about the project, what could/should have been different, and what it will take if 3-Tier is to be implemented statewide. What we'd like to do today is to get your in-depth feedback on the pilot program.

First, let me thank you for agreeing to an interview. Secondly, let me thank you for filling out the survey. Just so you know, the survey went to all 3-Tier staff in the various Field and Driver Safety offices that participated in the pilot. We are doing interviews with staff from all offices, and from all job classifications, to get as full a picture of the implementation process as we can.

This interview should take about a half-hour. Most of the questions are "open-ended"—this means that there is no "set" answer. We hope that you can be as detailed as possible in your responses.

I will tape-record this interview once we are done with this introduction, and everything you say will be confidential. Let me say a few words about what this means. We are interested in your insights into the 3-Tier process *in your capacity as someone who worked on the project*. In other words, we want to know what you, as an MVFR, or an LRE, or a Hearing Officer, or a manager, know about the pilot. For that reason, we do not need to know your name. When we report the results of this research, we may use specific quotes from this interview—however, we will not use your name. Instead, we will most likely identify you by your job classification or position. Secondly, the only people who will listen to the actual interview tape are myself and other members of the Research and Development (R&D) Branch.

We will report our findings from this project at presentations within DMV, at traffic safety and other conferences, in reports and articles. If you would like, we will be happy to provide you with a copy of major publications that use data from this interview.

If you have any questions about 3-Tier or this interview, I'd be happy to do my best to answer them now. I would also be happy find out the answers to any questions that I cannot answer at this time.

Thank you for helping us build a better DMV!

_____ Date: _____

Name of Informant (Please both print and sign your name. Your name will be kept confidential. Signing constitutes consent to use of this interview by DMV R&D for analyzing the process and outcome of the 3-Tier Pilot project).

NOTE 1 to interviewer: these are semi-structured questions. Alter as necessary to elicit full answers, but make sure to cover all subject areas.

NOTE 2: it is not necessary to ask all questions of every respondent. See suggestions below regarding the particular job categories to which you should direct particular questions.

NOTE 3: Some questions include "prompts." Use only when necessary to clarify question meaning and/or to draw forth more substantial answers from a respondent. Do not use to guide respondent to a particular answer.

MVFR/SMVT Questions

Background

1.) If we may, let's start with some basic background. What is your job classification?
    a. How long have you been in this position?
2.) Can you tell me, in your own words, what the goal (or goals) of 3-Tier is?

Workload

3.) About how many 3-Tier customers did you see (or process) over the course of a typical day?
4.) When it comes to 3-Tier, can you give me an example, or a general idea, of what a customer interaction involved for you?

Process and Customer Interaction

5.) Can you give me a sense of your own experience of a 3-Tier transaction? By that I mean, how did a transaction "feel" different—if that was in fact the case—from other kinds of tasks?
    a. For instance, which parts of a transaction or customer interaction, did you find especially difficult or complicated?
        i. What made them complicated?
    b. Which parts of it seemed easy or common-sensical?
        i. What seemed easy about them, compared to other parts?
6.) DMV puts a strong emphasis on customer service. Did 3-Tier affect, in your experience, the service you were able to provide to customers?
    a. Did it affect it in a positive way?
    b. How about in a negative way?
7.) One of the things that 3-Tier is designed to do is provide a structured way for MVFRs to observe each customer regarding their memory, any physical limitations, and their vision [provide copy of score sheet at this point]. Let's take each of those elements in turn:
    a. How well did the memory recall exercise work, in your experience?
        i. Are there other ways you can think of that we might collect this information?
    b. What about the observation of physical limitations—how well did that work, in your experience?
        i. Are there other ways you can think of that we might collect this information?
    c. What about the fog chart? How well did that work, in your experience?

Inter-personal and Inter-division Cooperation

8.) Anytime we change something about one job, we also tend to change the way that job intersects with other people's jobs. Was that true, in your experience, of 3-Tier?
    a. For instance, did 3-Tier lead to a different kind of interaction with your co-workers—the people in the same job classification as you?
    b. What about other people in the office (for instance, your manager or the 3-Tier Manager I?).

Training

9.) The work associated with a project can sometimes be easy, and sometimes hard, depending on the stage of the project.
    a. Can you describe a period when your role in 3-Tier was more difficult than at other times?
    b. Can you describe a period when your work on 3-Tier was easier than at other times?
    c. What changed to make your work easier, or more difficult, in each case?
10.) Let's talk about training for a moment. Obviously this was the first time we've done training for 3-Tier—including the project background, the procedures, and the completion of forms for data collection. We shall almost certainly have to change some things. What suggestions would you have for the future?
    a. Prompts: length of training, specificity/generality, differentiation from "regular" training, others?

Last Thoughts/Suggestions:

11.) Finally, do you have any last suggestions for any changes to or fine-tuning of the 3-Tier process?

That's all the questions I have. Do you have any questions you'd like to ask, or any points you'd like to add to what we discussed?

Finally, let me give you my deepest thanks. This has been a complicated project, and our success depends on feedback from people like yourself. Our analysis and reports won't be out for months, and in some cases years, but we will be sure to contact you when we are ready to share our findings.

LRE Questions

Background

1.) If we may, let's start with some basic background. What is your job classification?
    a. How long have you been in this position?
2.) Can you tell me, in your own words, what the goal (or goals) of 3-Tier is?

Workload

3.) About how many 3-Tier customers did you see over the course of a typical week?
4.) When it comes to the 3-Tier project, can you give me a general idea of what a typical customer interaction involved for you?

Process and Customer Interaction

5.) DMV puts a strong emphasis on customer service. Did 3-Tier affect, in your experience, the service you were able to provide customers?
    a. Did it affect it in a positive way?
    b. How about in a negative way?
6.) One of the things that 3-Tier was intended to do was to set up a standard structure and process for deciding when customers need to be given an on-the-road driving test—especially those with physical or mental limitations who might otherwise be renewed without adequately demonstrating their ability to drive safely. In particular, 3-Tier was meant explicitly to increase the number of Supplementary Driving Performance Evaluations (SDPE) we give.
    a. Did this lead to any sorts of changes in your interaction with customers?
    b. How often did you impose restrictions on a customer's license for safety reasons? [Examples: No freeway driving, no night driving]
    c. Under what kinds of circumstances did you consider assigning a Restriction 50 (a Supervised Instruction Permit, or a Special Restricted License)?
7.) One of the things that 3-Tier was intended to do was to come up with a standard structure for renewing customers to find out about, and to discuss with DMV staff, all their drive test options—including the ADPE.
    d. How often did you recommend, or administer, an ADPE?
    e. At what point would you usually bring up the option of taking an ADPE?
    f. During your observation of a customer during a drive test, what kinds of things would indicate to you that you would recommend an ADPE?
    g. In your experience, how do customers respond to the idea of taking an Area Drive?

Inter-personal and Inter-division Cooperation

8.) Anytime we change something about one job, we also tend to change the way that job intersects with other people's jobs. Was that true, in your experience, of 3-Tier?
    a. For instance, did 3-Tier lead to a different kind of interaction with your co-workers—the people in the same job classification as you?
    b. What about other people in the office, for instance, your managers?
    c. What about people in other branches or divisions with whom you would or do normally communicate? (for instance, communication between you and Hearing Officers regarding referrals)

9.) Implementing the 3-Tier Pilot required the coordinated efforts of at least three different branches and divisions of DMV: Field Office, Driver Safety, and Research and Development. This coordination appears to have gone more smoothly during some parts of the project, and less smoothly at other times.
    a. What was the most helpful or positive aspect of the interactions you were a part of?
    b. Were there any particularly difficult periods, or project components, that come to mind?
    c. Which individuals in other divisions or branches did you communicate with regarding these difficulties?
    d. How did these issues get resolved?
    e. Difficulties in implementing new projects are, at least to some extent, unavoidable. However, there are usually "lessons learned" that can be used to make future projects easier. Are there any specific suggestions or changes you would make regarding future inter-branch cooperation?

Training

10.) The work associated with a project can sometimes be easy, and sometimes hard, depending on the stage of the project.
    a. Can you describe a period when your role in 3-Tier was more difficult than at other times?
    b. Can you describe a period when your work on 3-Tier was easier than at other times?
    c. What changed to make your work easier, or more difficult, in each case?

11.) Let's talk about training for a moment. Obviously this was the first time we've done training for 3-Tier—including the project background, the procedures, and the completion of forms for data collection. We shall almost certainly have to change some things. What suggestions would you have for the future?

a. Prompts: length of training, specificity/generality, differentiation from "regular" training, others?

Last Thoughts/Suggestions:

12.) Finally, do you have any last suggestions for changes to, or fine-tuning of, the 3-Tier process?

That's all the questions I have. Do you have any questions you'd like to ask, or any points you'd like to add to what we discussed?

Finally, let me give you my deepest thanks. This has been a complicated project, and our success depends on feedback from people like yourself. Our analysis and reports won't be out for months, or in some cases years, but we will be sure to contact you when we are ready to share our findings.

Hearing Officer Questions

Background

1.) If we may, let's start with some basic background. What is your job classification?
    a. How long have you been in this position?
2.) Can you tell me, in your own words, what the goal (or goals) of 3-Tier is?

Workload

3.) About how many 3-Tier customers did you see (or process) over the course of a typical week?
4.) When it comes to the 3-Tier project, can you give me a general idea of what a typical customer interaction involved for you?

Process and Customer Interaction

5.) Can you give me a sense of your own experience of the structure of a 3-Tier customer contact? By that I mean, how did a contact "feel" different—if that was in fact the case—from other kinds of tasks?
    a. For instance, which parts of a 3-Tier customer contact did you find especially difficult or complicated?
        i. What made them complicated?
    b. Which parts of it seemed easy?
        i. What seemed easy about them, compared to other parts?
6.) One of the things that 3-Tier is designed to do is to change (somewhat) the nature of customer/staff interaction. This may have affected, positively or negatively, the nature and degree of customer service.
    a. Did that occur, in your view?
    b. If so, can you describe what was different?
7.) Thinking specifically about your 3-Tier referrals, how often did you assign a Special Instruction Permit, or Special Restricted License restriction? Under what kinds of circumstances would you consider doing so?
8.) One of the things that 3-Tier was designed to do was to come up with a standard structure for renewing customers to find out about, and to discuss with DMV staff, all their drive test options—including the ADPE.
    a. How often did you schedule, or administer, an ADPE?
    b. At what point would you usually bring up the option of taking an ADPE?
    c. During the course of a customer contact, what kinds of things would indicate to you that this customer ought to be referred for an ADPE?

d. In your experience, how do customers respond to the idea of taking an Area Drive?

Inter-personal and Inter-division Cooperation

9.) Anytime we change something about one job, we also tend to change the way that job intersects with other people's jobs. Was that true, in your experience, of 3-Tier?
   a. For instance, did 3-Tier lead to a different kind of interaction with your co-workers—the people in the same job classification as you?
   b. What about other people in your office (for instance, your managers).
   c. What about people in other branches or divisions with whom you would or do normally communicate? (for instance, your communication with Field Office LREs regarding referrals)

10.) Implementing the 3-Tier Pilot required the coordinated efforts of at least three different branches and divisions of DMV: Field Office, Driver Safety, and Research and Development. This coordination appears to have gone more smoothly during some parts of the project, and less smoothly at other times.
   a. What was the most helpful or positive aspect of the interactions you were a part of?
   b. Were there any particularly difficult periods, or project components, that come to mind?
   c. Which individuals in other divisions or branches did you communicate with regarding these difficulties?
   d. How did these issues get resolved?
   e. Difficulties in implementing new projects are, at least to some extent, unavoidable. However, there are usually "lessons learned" that can be used to make future projects easier. Are there any specific suggestions or changes you would make regarding future inter-branch cooperation?

Training

11.) The work associated with a project can sometimes be easy, and sometimes hard, depending on the stage of the project.
   a. Can you describe a period when your role in 3-Tier was more difficult than at other times?
   b. Can you describe a period when your work on 3-Tier was easier than at other times?
   c. What changed to make your work easier, or more difficult, in each case?

12.) Let's talk about training for a moment. Obviously this was the first time we've done training for 3-Tier—including the project background, the procedures,

and the completion of forms for data collection. We shall almost certainly have to change some things. What suggestions would you have for the future?

d. Prompts: length of training, specificity/generality, differentiation from "regular" training, others?

Last Thoughts/Suggestions:

13.) Finally, do you have any last suggestions for changes to, or fine-tuning of, the 3-Tier process?

That's all the questions I have. Do you have any questions you'd like to ask, or any points you'd like to add to what we discussed?

Finally, let me give you my deepest thanks. This has been a complicated project, and our success depends on feedback from people like yourself. Our analysis and reports won't be out for months or in some cases years, but we will be sure to contact you when we are ready to share our findings.

3-Tier Manager I and Administrative Manager Questions

Background

1.) If we may, let's start with some basic background. What is your job classification?
    e. How long have you been in this position?
    a. Can you tell me, in your own words, what the goal (or goals) of 3-Tier is?

Workload

2.) About how many 3-Tier customers did you see (or process) over the course of a typical day?
3.) When it comes to the 3-Tier project, can you give me a general idea of the kinds of activities you were involved in?

Process and Customer Interaction

4.) Can you give me a sense of your own experience of a 3-Tier transaction? By that I mean, how did a transaction "feel" different—if that was in fact the case—from other kinds of tasks?
    a. For instance, which parts of a transaction or customer interaction, did you find especially difficult or complicated?
        i. What made them complicated?
    b. Which parts of it seemed easy or common-sensical?
        i. What seemed easy about them, compared to other parts?
5.) One of the things that 3-Tier was intended to do was to change (at least a little bit) the nature of customer/staff interaction during the renewal process. This may have affected, positively or negatively, the nature and degree of customer service.
    a. Did that occur, in your view?
    b. If so, can you describe what was different?
6.) A key piece of the 3-Tier process was the Perceptual Response Test, or PRT.
    a. In your own words, can you describe what the PRT was designed to measure?
    b. How often (say, how many times per day) did you administer a PRT for a customer?
    c. How did customers react to taking the PRT? What sorts of questions did you get from customers about this test?
7.) One of the things that 3-Tier was designed to do was to set up a standard structure and process for deciding when customers need to be given an on-the-road driving test—especially those with mental or physical limitations, who

might otherwise be renewed without adequately demonstrating their ability to drive safely. In particular, 3-Tier was meant explicitly to increase the number of Supplementary Driving Performance Evaluations (SDPEs) we give.

   a. In your experience, how do customers react to having to take a drive test?
   b. In most cases, a customer would take an SDPE for their first drive test. How did you use the Scope of Driving Questions on *subsequent* drives, for making a determination about what kind of test (ADPE vs. SDPE) was most appropriate for a customer?

8.) One of the things that 3-Tier was designed to do was to come up with a standard structure for renewing customers to find out about, and to discuss with DMV staff, all their driving options—including the ADPE.

   a. How often did you schedule, or administer, an ADPE?
   b. In your experience, how do customers respond to the idea of taking an Area Drive?

9.) As part of the 3-Tier Pilot, we had some customers watch educational videos about how to be the best driver they can be.

   a. [prompts: contrast sensitivity video, PRT video]
   b. In your experience, what did customers think of the educational videos?
   c. Are there other ways you can think of that we could provide this service?

Inter-personal and Inter-division Cooperation

10.) Anytime we change something about one job, we also tend to change the way that job intersects with other people's jobs. Was that true, in your experience, of 3-Tier? [make sure to get a sense of all levels of interaction]:

   a. For instance, did 3-Tier lead to a different kind of interaction with the front-line staff in your office—the MVFRs, SMVTs, and LREs?
   b. What about people to whom you report (for instance, the Office Manager, or folks from the regional office)?
   c. What about people in other branches or divisions with whom you would or do normally communicate?
   d. Prompts: Research and Development, FOD Headquarters or Region III, Driver Safety.

11.) Implementing the 3-Tier Pilot required the coordinated efforts of at least three different branches and divisions of DMV: Field Office, Driver Safety, and Research and Development. This coordination went more smoothly during some parts of the project, and less smoothly at other times.

   a. What was the most helpful or positive aspect of the interactions you were a part of?

b. Were there any particularly difficult periods, or project components, that come to mind?
c. Which individuals in other divisions or branches did you communicate with regarding these difficulties?
d. How did these issues get resolved?
e. Difficulties in implementing new projects are, at least to some extent, unavoidable. However, there are usually "lessons learned" that can be used to make future projects easier. Can you share some of these lessons learned?

Training

12.) The work associated with a project can sometimes be easy, and sometimes hard, depending on the stage of the project.
    a. Can you describe a period when your role in 3-Tier was more difficult than at other times?
    b. Can you describe a period when your work on 3-Tier was easier than at other times?
    c. What changed to make your work easier, or more difficult, in each case?
13.) Let's talk about training for a moment. Obviously this was the first time we've done training for 3-Tier—including the project background, the procedures, and the completion of forms for data collection. We shall almost certainly have to change some things. What suggestions would you have for the future?
    a. Prompts: length of training, specificity/generality, differentiation from "regular" training, others?

Last Thoughts/Suggestions:

14.) Finally, do you have any last suggestions for tinkering or fine-tuning the 3-Tier process?

That's all the questions I have. Do you have any questions you'd like to ask, or any points you'd like to add to what we discussed?

Finally, let me give you my deepest thanks. This has been a complicated project, and our success depends on feedback from people like yourself. Our analysis and reports won't be out for months, or in some cases, years, but we will be sure to contact you when we are ready to share our findings.

Upper Management (Including Office Managers) and Headquarters Coordinating Personnel Questions

Background

1.) If we may, let's start with some basic background. What is your job classification and title?

Workload

2.) When it comes to the 3-Tier project, can you give me a general idea of the kinds of activities you were involved in? Perhaps you might start by telling me when you first become involved in the project, and what your responsibilities were?

Inter-personal and Inter-division Cooperation

3.) Implementing the 3-Tier Pilot required the coordinated efforts of at least three different branches and divisions of DMV: Field Office, Driver Safety, and Research and Development. This coordination appeared to go more smoothly during some parts of the project, and less smoothly at other times.
    a. What was the most helpful or positive aspect of the interactions you were a part of?
    b. Were there any particularly difficult periods, or project components, that come to mind?
    c. Which individuals in other divisions or branches did you communicate with regarding these difficulties?
    d. How did these issues get resolved?
    e. Difficulties in implementing new pilots are, at least to some extent, unavoidable. However, there are usually "lessons learned" that can be used to make future projects easier. Can you share any lessons learned from this project?
    f. Are there any specific suggestions or changes you would make regarding future inter-branch cooperation?

Implications and Preparations for Taking Statewide

4.) As you know, if the 3-Tier process is taken statewide, this would not happen before approximately 2012—after all of the research on traffic safety, the implications for the department's budget, and the policy impacts have been considered. <u>If</u> 3-Tier is taken statewide at some future date:
    a. What challenges would we (the DMV as a whole, your office, other branches) need to overcome?

b. What resources—financial, institutional, personnel-wise—would we need to ensure, or facilitate, success in implementation?
c. In your view, are there advantages that could be capitalized upon, or leveraged, for (potentially) taking 3-Tier statewide?

That's all the questions I have. Do you have any questions you'd like to ask, or any points you'd like to add to what we discussed?

Finally, let me give you my deepest thanks. This has been a complicated project, and our success depends on feedback from people like yourself. Our analysis and reports won't be out for some months, but we will be sure to contact you when we are ready to share our findings.

Sub-Appendix C: Scope of Driving Questions Used for Drive-Test Counseling During the 3-Tier Pilot

Pre-Drive Scope of Driving Questions:

Y N (1) Do you ever drive 15 miles or more from your residence?
Y N (2) Do you drive on longer trips, for 45 minutes or more at a time?
Y N (3) Do you ever drive on the FWY or HWYs with speed limits of 55 mph or more?
Y N (4) Do you ever drive during heavy traffic?
Y N (5) Do you ever drive to unfamiliar areas?

- If YES to one or more questions, schedule SDPE.
- If NO to all 5 questions, discuss how the SDPE differs from the ADPE (Use Script).

If customer requests an ADPE, then ask the 2 ADPE Questions.

ADPE Questions:

Y N (1) Do you drive at least once a month?
Y N (2) Do you limit your driving to certain locations, routes and destinations?

- If NO to either of these 2 questions, then schedule SDPE
- If YES to both, ask the 5 Additional Scope of Driving Questions

Additional Scope of Driving Questions:

Y N (1) When you are driving normally, are you able to avoid heavy traffic conditions?
Y N (2) Do you drive only in familiar areas?
Y N (3) Are the speed limits on the roads you use less than 55 mph?
Y N (4) Is most of your driving within 15 miles of your residence?
Y N (5) Do you take shorter trips whenever possible (30 minutes or less one-way)?
If YES to all 5 Additional Questions, then offer to schedule ADPE for First Drive Test

Y N Since you found out that you needed to take a drive test, did you get some behind the wheel training?
Who did you drive with? □Driving Instructor □Occup/Rehab Specialist □Friend or Relative
□Other:_____

Sub-Appendix D: Customer Survey Instrument

Customer Service Survey for New DMV "3-Tier" Pilot Assessment System

In accordance with a new state law (Assembly Bill 2542), and California Vehicle Code Section 1659.9, the California Department of Motor Vehicles (DMV) is currently conducting a pilot study of a new driver license assessment system. Recently, you visited a DMV field office and participated in this pilot study (please see the bottom of this page for a brief description of the new driver license screening tests).

We would like to know what you think about this new assessment system. Your opinion is important to us, and vital to creating a safe and efficient driver license renewal process. Your response to this survey is confidential—please do not write your name or driver license anywhere on this document. Your answers will be grouped with those of others for comparison of, for instance, people that experienced different screening tests.

Please take a few minutes to review the statements below. In each case, please circle the number that best reflects your opinions, then put this form in the enclosed, postage-paid envelope and mail it back to us.

Thank you for helping us build a better DMV!

	Disagree Strongly	Disagree	Agree	Agree Strongly
1) The time I spent during my office visit was reasonable.	1	2	3	4
2) I found the new assessment system easy to follow.	1	2	3	4
3) I found the instructions for each test easy to understand.	1	2	3	4
4) The DMV office staff treated me with courtesy and respect.	1	2	3	4
5) In my opinion, this new assessment system is fair to all customers.	1	2	3	4
6) I am confident that this new assessment system will improve driver safety.	1	2	3	4

As part of this new assessment pilot, you may have been asked to participate in one or more screening tests in addition to the regular visual acuity testing you are probably familiar with. This included reading a vision chart with faded letters (a "fog chart"). Depending on how you did on these exercises, you may also have been asked to: (a) take a perceptual response test—a.k.a. the "PRT"—which involved choosing between images of a truck and a car on a computer screen; (b) view an educational video; and/or (c) take a behind-the-wheel drive test.