OPTIMIZATION OF BITTERNESS IN CHOCOLATE THROUGH ROASTING WITH ANALYSIS OF RELATED CHANGES IN IMPORTANT BITTER COMPOUNDS

A Dissertation

Presented to

The Faculty of the Graduate School

At the University of Missouri-Columbia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

ALAN PATRICK MCCLURE

Dr. Ingolf Grün, Dissertation Supervisor

December, 2020

The undersigned, appointed by the dean of the Graduate School, have examined the

dissertation entitled

OPTIMIZATION OF BITTERNESS IN CHOCOLATE THROUGH ROASTING WITH ANALYSIS OF RELATED CHANGES IN IMPORTANT BITTER COMPOUNDS

Presented by Alan Patrick McClure

A candidate for the degree of Doctor of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

_____

Dr. Ingolf Grün, Food Science

_____

Dr. Misha Kwasniewski, Food Science

_____

Dr. Mark Ellersieck, Food Science

_____

Dr. Erin Schliep, Statistics

ACKNOWLEDGEMENTS

It is certainly not unusual for doctoral students to feel positively influenced by their primary advisor. However, the impact of Dr. Ingolf Grün's advice goes well beyond simple positive influence. He was *always* there to discuss research or other academic difficulties without question. Yet, it is more than simply being supportive that has made Dr. Grün a peerless mentor; he is quite literally one of the finest people I have ever known. Intelligent and wise, kind and ethical, and such a positive force in the department and the classroom, where he is unquestionably the best teacher I have ever had; I would absolutely not have been able to achieve what I have without his constant presence over the past 5 years. I owe him an enormous debt of gratitude that quite simply I will never be able to repay.

I would also like to thank Dr. Helene Hopfer for her generosity, direction, creativity, and a great deal of her valuable time, as we worked through a large and complex consumer sensory trial that was, to the best of my knowledge, the first of its kind in terms of scale and objectives. I am certain that I will never forget my enlightening and very happy time at Penn State, and without Dr. Hopfer, it would not have been possible. Next, I must thank Dr. Misha Kwasniewski for his wisdom and direction, and for making the instrumentation of his impressive lab available to me over the course of my project. Importantly, I would also like to give thanks to a brilliant cadre of statisticians including (alphabetically): Dr. Mark Ellersieck, Dr. Erin Schliep, and Dr. Christine Spinka, each of whom advised me over the course of several years on subjects such as experimental

design, data analysis, data visualization, and much more, to be sure that I would have the best chance of obtaining the quality data necessary to answer the questions that have interested me most, and the tools to do so. I would like to express gratitude to Lakdas Fernando, MS longtime technician (retired) of the Milton Bailey Flavor Chemistry laboratory, both for his friendship, and for his hands-off approach that allowed me to learn so much about HPLC analysis, extractions, and chromatography in general, while at the same time always being available to give direction if I was unable to satisfactorily solve any specific chromatographic problem on my own. As for the expertise of those beyond the university, it would be hard to find anyone who has been quite as instrumental in moving the project forward as Ed Seguine, Professional Manufacturing Confectioners Association (PMCA) member and chocolate industry consultant, who virtually singlehandedly acquired cacao donations for the project (with the exception of a donation from Marañon Chocolate's Dan Pearson, to whom I also owe huge thanks), and who acted as an efficient point person for contact with Guittard Chocolate, and as a deep well of experience when it comes to the practical side of cacao processing, especially as it pertained to the deceptively simple act of winnowing raw cacao, and roasting small amounts of cacao. As for the PMCA, this project was funded almost in its entirety through the generosity of their members, and none of the exciting possibilities, such as performing sensory analysis at the Sensory Evaluation Center at Penn State's storied Food Science department would have been possible without their help. The quality of our data is certain to have increased immeasurably as a result of their generosity, and I would like to especially acknowledge the primary PMCA contacts for

our project, Annelie Verbon and Brandy Kresge, who were both so pleasant and helpful to work with, and absolute positive forces for the project. Finally, thanks are in order for Dr. Gregory Ziegler, who I was honored showed sincere interest in our project, and generously offered advice and feedback, while also inviting me to his table on more than one occasion for amazing food and food-science discussion. Of course, it is impossible in a few short paragraphs to say “thank you” to everyone to whom I owe a debt of gratitude. But for anyone I have left out for the sake of rendering this section manageable in size, whether faculty, or other academic or industry advisor, and of course all friends and family members who have supported me throughout this process, please know that I do remember and value your contributions, and I am constantly in awe of the kindness and generosity that you have all shown me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

OPTIMIZATION OF BITTERNESS IN CHOCOLATE THROUGH ROASTING WITH ANALYSIS OF RELATED CHANGES IN IMPORTANT BITTER COMPOUNDS

Alan McClure

Dr. Ingolf Grün, Dissertation Advisor

ABSTRACT

Chocolate is made from the fermented, dried, and roasted seeds of the *Theobroma cacao* tree, an important agricultural food crop which contains bioactive flavonoid polyphenols with beneficial health effects. Such effects include improvement of antioxidant status, positive impacts on cardiovascular health and endocrine system function, association with cancer prevention, LDL cholesterol reduction, and reduction of obesity and related conditions. However, products which have the highest levels of cacao flavonoids of all eating-chocolate, such as high-cacao-percentage dark chocolate, are known to be quite bitter, a taste modality that is not readily appreciated by humans. Though the complex causes of bitterness in cacao are still not completely understood, it has long been known that the methylxanthines theobromine and caffeine impart bitterness, as do certain flavan-3-ols, sometimes called catechins, which are a class of the aforementioned healthy bioactive polyphenolic flavonoids, also found in tea. Yet, what else is known of bitterness in cacao is sparse and even contradictory. Work on cacao bitterness has described the importance of cyclic dipeptides called 2,5-diketopiperazines (DKPs), while suggesting some form of interaction between theobromine and DKPs as well. Yet these earlier assertions have only been confirmed with mixed results by others, in part due to the incredible complexity of bitterness in

roasted cacao, which has been said to require further sensory evaluation. More recent work on bitterness in cacao suggested for the first time that a DKP called cyclo(Pro-Val) is the most important bitter compound. However, even while seeming to confirm the importance of previously known important bitter compound classes, this research was based upon only a single cacao sample from a single origin of cacao, and with an undefined roasting treatment, even though previous work had noted that differences in DKP formation are dependent upon roast profile. Additionally, sensory work was based in part on recombinants of bitter compounds in aqueous solution, allowing for potentially biased estimation of the contribution of the different compounds to finished chocolate bitterness, since the varying kinetics of dissolution of the diverse bitter compounds from low-moisture, high-fat cacao matrix into saliva were not considered, nor were interactions with aroma compounds present in chocolate. Therefore, much was still to be learned about the variation in bitter-compound composition in cacao and related sensory characteristics, within and between different cacao origins and across different roast profiles. This fact, combined with a growing desire for healthy, functional versions of foods such as chocolate makes research into the impact of cacao roasting on consumer perceptions of bitterness and overall liking in chocolate, and the underlying chemical changes, all the more timely. This research project has resulted in findings covering a significant range of chocolate topics. First of all, a new efficient method for extraction and analysis of important bitter compounds in cacao and chocolate was developed. A custom response-surface methodology (RSM)-based design for the roasting treatments, with emphasis on I-optimality for minimizing prediction variance,

was created. Chemical and sensory analysis of the roasted chocolate treatments were carried out, followed by in-depth data analysis and interpretation in the context of current chocolate science. Specifically, the aqueous 70% N,N-Dimethylformamide solvent system and HPLC method developed for fast and efficient extraction, followed by analysis, of important bitter compounds from three different chemical classes (i.e., methylxanthines, flavan-3-ols, and 2,5-diketopiperazines) simultaneously, functioned successfully, resulting in acceptable standard curves, % RSD values, and % recovery values. As for quantitative chemical findings, our work generally supports previous studies as regards changes in chemical concentrations during roasting. However, even with the large number of roasting treatments (i.e. 24, or 8 for each of 3 origins) across a reasonably large experimental region, we did not confirm the presence of concentrations of cyclo(Pro-Val) similar to that of previous research. As for sensory evaluation findings, we discovered that reduction of bitterness, sourness, and astringency are all correlated with increased liking in our chocolates. We also noted that consumers appear to have a preference for increased cocoa intensity. Roast profiles that minimize and maximize these characteristics respectively can vary by origin, but temperature and time combinations such as 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C were generally effective, whereas, raw and lightly roasted treatments (i.e., 0 minutes at 24°C, 11 minutes at 105°C, or 55 minutes at 64°C) were not, resulting in the lowest liking ratings. As with any complex food system, caveats do exist. Additions of sugar, salt, and other ingredients would likely introduce significant effects relevant to overall sensory characteristics and consumer liking, and intensity of various other aroma

profiles not yet analyzed could do the same (e.g., floral, fruity, nutty). One additional sensory finding is that we can now say that perception of chocolate aroma is likely to play a large role in the perception of taste modalities (i.e., bitterness, sourness, sweetness), and astringency, as well as liking in chocolate. Finally, regarding the relationship of bitter chemical concentrations in the treatments, and consumer bitterness perception thereof, while the analysis is somewhat complicated by the stability of theobromine and caffeine during roasting, we can say that we have little evidence to suggest that theobromine concentration is strongly correlated to bitterness in chocolate. There is far more evidence that caffeine may play a role in the increase of bitterness in cacao, though the magnitude of its importance is not yet known, and to better understand the impact of both theobromine and caffeine, study of many more origins will be required. As for epicatechin and procyanidin B2, as already known, they are quite well correlated, and of all the chemicals we studied, they were, as a pair, the most correlated with changes in bitterness in our data across all treatments, including all three origins. Given that epicatechin has previously been shown to be a more important contributor to bitterness than higher molecular weight procyanidins (e.g., procyanidin B2), the overall importance of epicatechin could be the greatest of all the compounds that we studied. In contrast, catechin and cyclo(Pro-Val), do not appear to be particularly important for changes in bitterness. More specifically, we have found no evidence supporting the claim that the DKP cyclo(Pro-Val) is the most important bitter compound in cacao or chocolate. This does raise additional questions about the

importance of diketopiperazines (DKPs) as a class as they relate to bitterness in chocolate.

CHAPTER 1

1 INTRODUCTION

1.1 BACKGROUND

Chocolate is made from the fermented, dried, and roasted seeds of the *Theobroma cacao* tree (Aprotosoaie, Luca, & Miron, 2016a), which in addition to being an important agricultural food crop (Aprotosoaie et al., 2016a), is known to contain bioactive flavonoid polyphenols with beneficial health effects in humans (Andújar, Recio, Giner, & Ríos, 2012; Aprotosoaie, Miron, Trifan, Luca, & Costache, 2016b). Such effects include improvement of antioxidant status (Cooper, Donovan, Waterhouse, & Williamson, 2008), positive impacts on cardiovascular health (Andújar et al., 2012; Ellam & Williamson, 2013) and endocrine system function (Andújar et al., 2012; Bowser et al., 2017; Strat et al., 2016), and association with cancer prevention (Ishaq & Jafri, 2017), LDL cholesterol reduction (Ellam et al., 2013) and reduction of obesity and related conditions (Matsui et al., 2005; Min et al., 2013; Sun, Tao, Gu, & Lambert, 2015). However, despite strong evidence for the healthful qualities of cacao, products which have the highest levels of cacao flavonoids of all eating-chocolate, such as high-cacao-percentage dark chocolate (Langer, Marshall, Day, & Morgan, 2011; Miller et al., 2009), are known to be quite bitter, a taste modality that is not readily appreciated by many humans (Drewnowski & Gomez-Carneros, 2000; Fischer, Gilad, Man, & Pääbo, 2005). Though the complex causes of bitterness in cacao are still not completely understood, it has long been known that two methylxanthines, theobromine and caffeine, impart

bitterness to cacao and chocolate (Stark, Bareuther, & Hofmann, 2006; Woskresensky, 1842). Additionally, certain flavan-3-ols, sometimes called catechins, which are a class of the aforementioned healthy bioactive polyphenolic flavonoids, also found in tea (Scharbert & Hofmann, 2005), are likewise bitter and present in cacao at levels above the detection threshold (Stark, Bareuther, & Hofmann, 2005a)(e.g., (-)-epicatechin (Stark et al., 2006)). Yet, what else is known of bitterness in cacao is sparse and can seem contradictory. Pickenhagen's (1975) work was the first to describe the importance in cocoa of the cyclic dipeptides called 2,5-diketopiperazines (DKPs), while suggesting some form of interaction between theobromine and DKPs, leading to the "typical bitterness" (Pickenhagen et al., 1975) of cacao. Yet Pickenhagen's assertions were only confirmed with mixed results by Bonvehí and Coll (2000), who ultimately determined that the incredible complexity of bitterness in roasted cacao, and the role of DKPs within it, required further sensory evaluation. The most recent significant work on bitterness in cacao (Stark et al., 2006) suggested for the first time that a DKP called cyclo(L-Pro-L-Val) is the most important bitter compound in cacao. In this work (Stark et al., 2006), the contribution to bitterness in cacao of compounds in the methylxanthine, flavan-3-ol and 2,5-diketopiperazine classes was confirmed; however, the data and results were based upon a single cacao sample from a single origin of cacao—Ghana—and with an undefined roasting treatment, even though previous work had already noted that differences in DKP formation are dependent upon roast profile and probably also origin (Bonvehí et al., 2000; Rizzi, 1989). Stark's work (2006) also relied on sensory analysis of recombinants of bitter compounds found in cacao in aqueous solution (Stark et al.,

2006), allowing for potentially biased estimation of the contribution of the different compounds to finished chocolate bitterness, since the varying kinetics of dissolution of the diverse bitter compounds from low-moisture, high-fat cacao matrix into saliva were not considered.

Therefore, fourteen years after Stark's (2006) work on bitterness in cacao, much is still to be learned about the variation in bitter-compound composition in cacao and related sensory characteristics, within and between different cacao populations/origins and across different roast profiles, even considering that the significance of the three primary classes of bitter compounds—DKPs, flavan-3-ols, and methylxanthines—appears clear. Indeed, the continued dearth of work in the area of cacao bitterness over the last decade, while faced with a growing desire for healthy, functional versions of foods such as chocolate (Mintel, 2016; Pacyniak, 2014; 2016), makes research into the impact of cacao roasting on consumer perceptions of bitterness and overall liking in chocolate, and the related chemical changes, all the more timely.

1.2 OBJECTIVES

Given the goal of increasing the consumption of healthful dark chocolate, this project has aimed to determine the impact of roasting on bitterness, given its status as the most important step in cacao processing (Aprotosoaie et al., 2016a), especially as it is related to the optimization of overall liking, and to consider its relationship to other sensory characteristics such as astringency. Sub-objectives are as follows:

1) Roast three different origins of cacao with multiple treatments (time and temperature combinations) chosen in accordance with Response Surface Methodology (RSM). Produce 100% unsweetened chocolate (chocolate liquor) from each cacao treatment.
2) Perform quantitative analyses of important bitter chemicals in the chocolate liquors made from all cacao treatments.
3) Perform consumer sensory analysis of the same chocolate liquors to determine intensity of liking, bitterness, astringency, sourness, sweetness, cocoa intensity, and other relevant characteristics.
4) Perform mixed-effects modeling and other statistical analyses to determine the relationship of origin and roasting effects (i.e., time and temperature) to the sensory characteristics.
5) Perform mixed effects modeling and other statistical analyses to determine the relationship of specific bitter compound concentrations to consumer bitterness intensity specifically.

CHAPTER 2

2 LITERATURE REVIEW

2.1 Cacao and Chocolate

2.1.1 Cacao

Also known as cocoa, cacao consists of the fermented and dried seeds of the fruit of the tropical *Theobroma cacao* tree in the *Malvaceae* family (Aprotosoaie et al., 2016a). Cacao is a significant food commodity, with annual global consumption reaching approximately 4.6 million metric tons as of 2018 with an increase in demand of 3.9% over 2017 (Barchart, 2019), and plays an important role as the primary ingredient in the manufacture of chocolate (Aprotosoaie et al., 2016a).

2.1.2 Chocolate

Chocolate is a usually sweetened, uniquely flavored, solid paste that melts smoothly at human body temperature due to the presence and unique fatty acid composition of cacao fat, called cocoa butter (Aprotosoaie et al., 2016a). The aforementioned unique chocolate aroma is due mostly to the presence of cacao, required in American-made chocolate by FDA Title 21 (2019). Prior to transformation into chocolate, cacao is roasted to obtain a more complex flavor and character that is generally preferred by consumers over that of raw cacao (Aprotosoaie et al., 2016a).

2.2 Health Effects of Cacao and Chocolate

2.2.1 Cardiovascular-Protective Effects

Numerous *in vivo* human studies suggest that cacao flavonoids have a protective effect on the heart and cardiovascular system through increased vascular function and decreased blood platelet aggregation (Peluso, Palmery, & Serafini, 2015; Vlachojannis, Erne, Zimmermann, & Chrubasik-Hausmann, 2016), as well as decreased platelet reactivity in their capacity as antioxidants (Aprotosoaie et al., 2016b; Keen, Holt, Oteiza, Fraga, & Schmitz, 2005). Observational studies also suggest that chocolate consumption may be beneficial for the heart (Aprotosoaie et al., 2016a; Mostofsky et al., 2017), including a recent study by Harvard's Chan School of Public Health, which found in an epidemiological sample of 55,502 men and women between the ages of 50 and 64 that chocolate consumption is inversely correlated with atrial fibrillation (AR) (Mostofsky et al., 2017), AR being associated with "higher risk of stroke, heart failure, cognitive decline, dementia and mortality" (Mostofsky et al., 2017).

2.2.2 Anti-Obesity Effects

In vitro studies (Gu, Hurst, Stuart, & Lambert, 2011; Stanley et al., 2018), along with *in vivo* mice studies (Dorenkott et al., 2014), have determined that cacao flavonoids, including flavan-3-ol monomers and oligomers, have anti-obesity effects (Dorenkott et al., 2014; Stanley et al., 2018). Flavan-3-ol oligomers are produced during cacao roasting (Stanley et al., 2018), and the mechanisms by which these anti-obesity effects occur include competitive and non-competitive inhibition of pancreatic lipase (Gu et al., 2011).

Additionally, further *in vitro* studies have shown that higher molecular weight fractions of roasted cacao extract inhibit intestinal α-glucosidase, which has been shown to result in a slower rate of carbohydrate digestion, and therefore fewer calories absorbed, which is correlated with weight loss (Bellesia & Tagliazucchi, 2014). The high molecular weight fractions analyzed were found to contain polysaccharides, proteins, phenolic compounds, and melanoidins (Bellesia et al., 2014), with melanoidins themselves being a complex of a variety of smaller compounds such as polyphenolic flavan-3-ols (Morales, Somoza, & Fogliano, 2012).

2.2.3 Anti-Diabetic Effects

In vitro studies have described improvements in insulin function and even insulin mimetic properties of cacao flavonoids (Bowser et al., 2017), while they have also shown that high molecular weight cocoa melanoidins are able to effectively scavenge excessive carbohydrate breakdown products (e.g., α-dicarbonyls) that otherwise contribute to neuropathy, nephropathy, and other chronic disease (Zhang, Zhang, Troise, & Fogliano, 2019). Interestingly, a study of rodent animal models using extracts of both raw and roasted cacao showed that both had anti-obesity and anti-diabetic properties, and functioned effectively at protecting hepatic cells from oxidative stress and modulating insulin signaling (Żyżelewicz et al., 2016). This finding showed, specifically, that roasting did not diminish the beneficial properties of the cacao extracts in these regards (Żyżelewicz et al., 2016).

2.2.4 Miscellaneous Positive Health Effects

In vitro studies have shown that cacao flavonoids function as antioxidants, scavenging free radicals, resulting in overall reduced oxidative stress, inhibiting DNA damage, and reducing oxidation of LDL-cholesterol and lipids (Andújar et al., 2012). Additionally, though certain flavan-3-ol concentrations decline during roasting, the high molecular weight melanoidins formed as Maillard reaction products contain large concentrations of phenolic compounds (Quiroz-Reyes & Fogliano, 2018), and melanoidins have been found to function effectively as antioxidants (Langner & Rzeski, 2014; Morales et al., 2012; Quiroz-Reyes et al., 2018), as well as resulting in synergistic effects with phenolics, allowing cacao to maintain antioxidant activity despite roasting (Quiroz-Reyes et al., 2018). Several *in vitro* studies have also shown other beneficial effects due to melanoidins, including antimicrobial effects for certain pathogenic bacteria, and have suggested that melanoidins are also antihypertensive, prebiotic, and antiallergenic, as well as being antimutagenic, with tumor-growth-inhibiting effects (Langner et al., 2014). Therefore, even as research in the field of bioactive cacao flavonoids and melanoidins continues, a relatively persuasive case has so far been made that cacao and chocolate, whether raw or roasted, can be healthful additions to a balanced diet.

2.3 Bitterness and its Sensation and Perception

2.3.1 Bitterness

Bitterness is one of the five taste modalities (i.e., salty, sweet, sour, bitter, umami) sensed by the tongue (Gaudette & Pickering, 2013; Keast & Breslin, 2003b). It has been noted that toxic compounds are often bitter (Keast et al., 2003b), and the ability of humans to taste bitter substances is likely to have evolved as a form of toxin detection (Keast, Bournazel, & Breslin, 2003a). This may explain why bitterness is generally disliked by humans (Drewnowski et al., 2000; Fischer et al., 2005) and even rejected in most food (Gaudette et al., 2013), despite famous exceptions such as coffee, beer, red wine, and dark chocolate (Gaudette et al., 2013; Keast et al., 2003b; Roy, 1997) that highlight the sometimes complex nature of human food choices (Gaudette et al., 2013).

2.3.2 Bitterness Sensation

Bitter sensation begins as a ligand/taste molecule contacts a bitter receptor cell, and bitter taste transduction results, with nerve impulses being sent to the brain (Drewnowski, 2001; Lawless & Heymann, 2010b). Bitter taste receptor cells are clustered as part of taste buds, located on the papillae across the tongue, with more located on the palate, and in the throat (Drewnowski, 2001; Lawless et al., 2010b). Of the basic tastes (i.e., bitter, sweet, sour, salty, umami), bitter is the most complex (Drewnowski, 2001), and it is now known that in humans there are approximately 25 subtypes of G-protein-coupled receptors called TAS2Rs which are responsible for the transduction of bitter taste from many thousands of compounds (Dagan-Wiener et al.,

2018; Maehashi & Huang, 2009). Each bitter taste sensation is complex and unique, with four properties to define it: quality, intensity, temporal patterns and spatial patterns (Keast et al., 2003b).

2.3.3 Bitterness Perception

Bitterness perception starts with the sensation of bitter compounds, but also includes processing by the brain of incoming signals from other sensory modalities (i.e., other tastes, aromas, and somatosensory, aural, and visual inputs) (Lawless & Heymann, 2010a). In short, bitterness perception is a combination of sensation and central cognitive effects resulting from other concurrently processed sensory information (Keast et al., 2003b). For example, aural stimulation (i.e. music) (Carvalho, Wang, van Ee, Persoone, & Spence, 2017) can even affect bitterness perception.

2.3.4 Variation in Bitterness Sensation/Perception

There is a great deal of variation across human individuals regarding sensation of bitterness (Drewnowski, 2001; Mennella, Pepino, & Reed, 2005; Negri et al., 2012), and the ability to sense certain bitter compounds at all, which can be inherited (e.g., phenylthiocarbamide (PTC) and 6-n-propylthiouracil (PROP) (Drewnowski, 2001)). Genotype therefore leads to a large variance between individuals regarding sensitivity to bitter compounds, where approximately 1/3 of people are unable to sense certain bitter compounds, while others, so-called supertasters, show hypersensitivity to bitter compounds, due in part to larger numbers of taste buds, with all other individuals falling upon a spectrum between these extremes (Lawless et al., 2010b). Additionally, children

avoid bitter foods to a larger extent than do adults (Negri et al., 2012), and sex-based differences have also been described (Bartoshuk, Duffy, & Miller, 1994).

Variation in bitterness perception depends on more than genotype, age, and sex, however. It is an interesting case that perception of overall bitterness intensity of a mixture of bitter compounds at known concentrations is less than that of the sum of intensities of the individual compounds at the same concentrations (Keast et al., 2003b). It is also the case that bitterness may be suppressed by certain compounds with sweet, salty, and umami tastes, and enhanced by those with sour ones (Calviño, García-Medina, Cometto-Muniz, & Rodríguez, 1993; Drewnowski, 2001; Fischer & Noble, 1994). Short-term variance in perception also exists, when exposure to bitter compounds can lead to adaptation (i.e., decreased responsiveness) to bitterness (Lawless et al., 2010b), and bitterness can even be altered by the concentration of calcium ions in an individual's saliva (Neyraud & Dransfield, 2004).

2.4 Known Important Bitter Compounds in Cacao

2.4.1 Methylxanthines

It is well known that the bitter compounds theobromine and caffeine are present in cacao at levels above their sensory detection thresholds (Bonvehí et al., 2000; Stark et al., 2006). However, the ratio of theobromine to caffeine can change substantially depending upon cacao variety (Bonvehí et al., 2000; Trognitz et al., 2013; Ziegleder, 2009) covering a ratio range from approximately 1 to 6 or more (Trognitz et al., 2013). Still, theobromine is present in cacao at a higher concentration than caffeine (Stark et

al., 2006), which one might assume makes it a more important contributor of bitterness in cacao, though this assumption is complicated by the much lower solubility of theobromine than caffeine in near-neutral aqueous solutions (Spiller, 1997; Stark et al., 2006). Furthermore, any potential differences in solubility kinetics between the two compounds are likely to be important for overall bitterness, as chocolate passes through the mouth only briefly, and such a comparison has not been published as far as this author is aware. However, theobromine is known to dissolve quite slowly (Wadsworth, 1921), an observation also noted by this author, suggesting the possibility that a given amount of theobromine in chocolate could result in a lower perceived bitterness than that of an equivalent amount of caffeine. This does not even consider the potential difference in bitterness in equimolar solutions of the two compounds due to possible variation in bitterness receptor binding.

2.4.2 Flavan-3-ols

Flavan-3-ols are the main polyphenolic compounds in cacao, even when origin and varietal are considered (Oracz, Nebesny, & Żyżelewicz, 2015). These healthful cacao flavonoids (Aprotosoaie et al., 2016b), which are known to be bitter (Stark et al., 2006), are present in cocoa at levels above their sensory detection threshold, especially the monomer (-)-epicatechin (Natsume et al., 2000; Stark et al., 2006; Ziegleder, 2009), and its epimer (+)-catechin (Kothe, Zimmermann, & Galensa, 2013), and oligomers such as procyanidin B2, which are also present above their taste threshold concentrations (Stark et al., 2006). However, approximately 35% of the polyphenol content of cacao is made up of the (-)-epicatechin monomer alone (Cooper et al., 2007). Interestingly,

epicatechin appears to decrease during roasting, while catechin and some larger molecular weight polymers of epicatechin increase (Kothe et al., 2013; Stanley et al., 2018), as do melanoidins (Quiroz-Reyes et al., 2018), themselves containing phenolic compounds (Morales et al., 2012). Overall, changes in these compounds during roasting appear to lead to an overall reduction in bitterness and astringency (Aprotosoaie et al., 2016a).

2.4.3 2,5-Diketopiperazines

2,5-Diketopiperazines (DKPs) are cyclic dipeptides that may be found in certain foods and beverages (Borthwick & Da Costa, 2017; Ginz & Engelhardt, 2000), and that in the case of cacao, are created predominantly during roasting, and have a bitter taste (Bonvehí et al., 2000; Stark et al., 2006). The mechanism of formation is heat-induced cyclization of linear alpha-amino-acid oligomer precursors (i.e., peptides) (Rizzi, 1989). It has been claimed that the bitter taste of DKPs is augmented in the presence of theobromine (Borthwick et al., 2017), an interaction (Drewnowski, 2001) resulting in a characteristic cocoa-like bitterness (Pickenhagen et al., 1975). A variety of DKPs have been found in varying amounts in cacao after having undergone various roasting treatments (Bonvehí et al., 2000; Rizzi, 1989; Stark et al., 2006) as well as in a variety of other roasted foods (Bonvehí et al., 2000; Gautschi et al., 1997; Ginz et al., 2000; Roudot-Algaron, Le Bars, Einhorn, Adda, & Gripon, 1993). It has been claimed that the DKP cyclo(L-Pro-L-Val) is the most important bitter compound in roasted cocoa, based upon cacao sourced from Ghana with an unknown roast treatment (Stark et al., 2006),

though this has apparently not yet been verified by further research, nor for a variety of roast profiles or cacao origins/populations beyond the single sample of Ghana.

2.5 Causes of Bitterness Variation in Cacao

Factors that impact bitterness in cacao, and therefore chocolate, are numerous, including varietal, growing conditions (e.g., hours of sunlight and rainfall), ripeness at harvest, and post-harvest processing (Kongor et al., 2016). Geographical location, even within a single country, appears to play a role in overall methylxanthine concentration and theobromine to caffeine ratio, probably due to altitude of the specimens from which the cacao is derived (Carrillo, Londoño-Londoño, & Gil, 2014). However, cacao varietal, ripeness at harvest, and the multi-stage post-harvest processing (i.e., fermentation and drying) have perhaps been most commonly noted as playing important roles (Afoakwa, Paterson, Fowler, & Ryan, 2008; Aprotosoaie et al., 2016a; Beckett, Fowler, & Ziegler, 2017; Kongor et al., 2016).

2.5.1 Varietal

Regarding cacao varietal, while ongoing research has led to a better understanding of the cacao genome and genetic variation in cacao generally (Johnson et al., 2009; Motamayor et al., 2008; Takrama et al., 2014; Zhang et al., 2011; Zhang et al., 2012), the sensory characteristics of the 10 identified genetic clusters of cacao (Motamayor et al., 2008) have yet to be thoroughly characterized, though an ongoing sensory evaluation project at Penn State will likely result in useful data on this topic (Hopfer & Brown, 2018, personal correspondence). Certainly, it is already clear that specific varietals show

variation in bitterness (Kongor et al., 2016), with theobromine concentration (Bonvehí et al., 2000) and ratios of theobromine to caffeine being significantly different across origins/varietals/populations (Timbie, Sechrist, & Keeney, 1978; Trognitz et al., 2013; Zoumas, Kreiser, & Martin, 1980). Additionally, there appear to be differences in the DKP concentrations of the roasted cacao prepared from fermented and dried cacao of different origins (Rizzi, 1989).

2.5.2 Ripeness

Ripeness of the cacao fruit at the time of harvest is known to play an important role in cacao flavor in general (Afoakwa et al., 2008) and in bitter methylxanthine concentration in particular (Bonvehí et al., 2000; Timbie et al., 1978) as caffeine and theobromine both increase substantially between 12 weeks from flowering and fruit harvest at 5-6 months (Senanayake & Wijesekera, 1971). Cacao flavonoid profile, including flavan-3-ols, also depends upon ripeness (Nazaruddin, Seng, Hassan, & Said, 2006; Rusconi & Conti, 2010), as unripe cacao has 29% less epicatechin than ripe cacao (Payne, Hurst, Miller, Rank, & Stuart, 2010).

2.5.3 Post-Harvest Processing

2.5.3.1 Fermentation

While post-harvest processing, which consists of fermentation and drying, is a well-known cause of bitterness reduction in cacao (Afoakwa et al., 2008; Aprotosoaie et al., 2016a; Beckett et al., 2017), it is a complex one with many unknown factors of its own (John et al., 2019). It is known that during cacao fermentation in the presence of oxygen and polyphenol oxidase, flavonoid oxidation and polymerization occur, leading to

browning, and reduction of bitterness and astringency (Ziegleder, 2017). Additionally, fermentation may result in an up to 40% reduction in bitter methylxanthines due to their diffusion through the cacao seed coat and loss in fermentation runoff/waste (Bonvehí et al., 2000), and flavonoids are also likely to be lost in runoff (Cooper et al., 2007). It has also been noted that the ratio of catechin to epicatechin shifts during fermentation, increasing during the process (Porter, Ma, & Chan, 1991), likely due to increases in (-)-catechin as (-)-epicatechin concentration decreases (Cooper et al., 2007; Payne et al., 2010). Overall, there is a >80% decrease in total catechin monomers (i.e., epicatechin and catechin) during fermentation (Payne et al., 2010).

2.5.3.2 Drying

Well-managed drying practices result in further reductions in bitterness via additional flavonoid oxidation (Barišić et al., 2019; Ziegleder, 2009). Interestingly, and perhaps unexpectedly, very low concentrations of certain DKPs have been found in fermented and dried, but *unroasted*, cacao (Rizzi, 1989).

2.5.4 Roasting

Roasting, considered by some to be the most important step in processing cacao (Aprotosoaie et al., 2016a), results in the creation of bitter diketopiperazines (DKPs) from peptides (Rizzi, 1989; Ziegleder, 2017), and darker roasts, particularly at higher temperatures, appear to increase DKP levels the most (Bonvehí et al., 2000), whereas unroasted cocoa contains virtually no DKPs (Bonvehí et al., 2000). However, roasting is required to alter the harsh, unpleasant flavor of raw cacao, and to develop characteristic cocoa aroma (Ziegleder, 2009). Additionally, roasting alters the concentrations of

epicatechin, and its epimers and oligomers (Kothe et al., 2013; Stanley et al., 2018), compounds which are both bitter and astringent (Stark et al., 2006), sometimes in unexpected ways seemingly related to varietal (Kothe et al., 2013). Loss of epicatechin at temperatures over 70°C occurs, and at a roasting temperature of 120°C catechin content has been seen to increase by approximately 650% in fermented cacao (Payne et al., 2010). This is probably due to epimerization of epicatechin to catechin (De Taeye, Kankolongo Cibaka, Jerkovic, & Collin, 2014; Oracz et al., 2015; Payne et al., 2010) and decomposition of procyanidins (Oracz et al., 2015), probably first to epicatechin followed by epimerization (Zhu et al., 2002). For this reason, the epicatechin/catechin ratio helps to understand the previous processing of cacao (Payne et al., 2010). Interestingly, roasting also leads to production of high molecular weight melanoidins that bind to polyphenols such as catechin, epicatechin, and procyanidin B2 (Oracz, Nebesny, & Żyżelewicz, 2019). As for the epimerization of epicatechin during roasting, it appears to be from (-)-epicatechin to (-)-catechin, even though the naturally present form of catechin is (+)-catechin (De Taeye et al., 2014; Hu, Kim, & Baik, 2016; Hurst et al., 2011; Payne et al., 2010). Cooper et al. (2007), was the first to show this (-)-epicatechin to (-)-catechin epimerization due to cacao processing, using a chiral column with UPLC (Cooper et al., 2007), and this was later confirmed by Kothe et al. (2013). Cooper et al. (2007) also showed that there is a strong linear relationship between epicatechin and procyanidins, as well as (+)-catechin, so that they can all be predicted from the epicatechin concentration, while (-)-catechin, however, cannot be predicted in the same fashion (Cooper et al., 2007). Finally, it has been known for decades that a

small loss of methylxanthines from the cotyledons to the cacao shell is expected during roasting, apparently increasing with degree of roast (Wadsworth, 1922).

2.5.5 Alkalization

Alkalization, sometimes called “Dutch Processing” (Miller et al., 2008) or simply “Dutching” (Aprotosoaie et al., 2016a), is the application of basic/alkaline aqueous solution to cacao or its products, either before or after roasting, in order to alter color and other sensory characteristics (Aprotosoaie et al., 2016a; Ziegleder, 2017). For example, reduction of astringency and bitterness occur due to flavonoid oxidation (Li et al., 2012), with bitterness further reduced due to apparent methylxanthine degradation (Li et al., 2012), and Maillard reaction product increase (Li et al., 2012; Ziegleder, 2017) leading to a profile deemed typical for cacao (Afoakwa et al., 2008), and reduction of sour notes also occurs (Afoakwa et al., 2008), perhaps due to significant acetic acid neutralization (Li et al., 2012). However, origin characteristics are generally also altered during this process, depending upon the pH of the basic solution employed (Afoakwa et al., 2008), leading to a cacao product deemed mellow and containing undertones of what are described as notes of alkali (Ziegleder, 2017). Interestingly, unlike in the case of non-alkalized cacao roasting, even larger molecular weight flavonoid polymers decrease with alkalization treatment, leading to overall reduced polyphenol concentrations and reduced Oxygen Radical Absorbance Capacity (ORAC) when compared to non-alkalized roasted cacao products (Miller et al., 2008), with the processing resulting in a loss of 98% of epicatechin and 80% of catechin overall (Payne et al., 2010). Therefore, it may be suspected, given the known beneficial qualities of

cacao flavonoids (Peluso et al., 2015; Vlachojannis et al., 2016), that alkalization could have a detrimental impact on the healthful qualities of cacao in chocolate.

2.5.6 Conching

Conching is a mixing/kneading and aerating process of the chocolate, where it is heated via friction or other means, resulting in significant flavor and textural changes during which there may be some further slight oxidation of cacao flavonoids (Afoakwa et al., 2008). The final result is a more balanced tasting, less intense, and less bitter tasting finished product with a smoother texture and better mouthfeel (Ziegleder, 2017).

2.6 Reduction of Bitterness in Cacao

As noted previously, several of the steps in the chocolate manufacturing process are well known to reduce or minimize bitterness, chief among these in overall impact being post-harvest processing (i.e. fermentation and drying), alkalization, and roasting (Afoakwa et al., 2008; Ziegleder, 2017). However, bitterness minimization through manufacturing processes such as these is not the only possible way to reduce bitterness or perceived bitterness in cacao and chocolate. Bitterness masking (Gaudette et al., 2013) and bitter compound extraction (Ramalakshmi & Raghavan, 1999) are also known, though they each also possess certain acknowledged complications.

2.6.1 Bitterness Masking

Bitterness masking is a technique for reducing the sensation or perception of bitterness that has long been known in the pharmaceutical and food industries (Alalor, 2015;

Gaudette et al., 2013; Keast, 2008; Sun-Waterhouse & Wadhwa, 2013). Perhaps the simplest masking agent is sucrose, which leads to mixture suppression effects when combined with bitter foods and beverages, in addition to lending sweetness (Gaudette et al., 2013). Other common compounds, such as sodium chloride, have also been found to have bitterness-masking effects, as have lesser known compounds such as homoeriodictyol salts (Gaudette et al., 2013; Ley, Blings, Paetz, Krammer, & Bertram, 2006). As a whole, there are a variety of known masking agents that have been shown to reduce the sensation and/or perception of bitterness depending upon the bitter compounds in question, through physicochemical, peripheral physiological, and/or central cognitive effects (Gaudette et al., 2013; Sun-Waterhouse et al., 2013). Certainly, sucrose is already commonly used in chocolate manufacture (Beckett et al., 2017), often in substantial quantities, such as is the case with mass-produced milk chocolates (Beckett et al., 2017). However, masking agents can have disadvantages; for example, it is now known that sucrose consumption, especially in large amounts, can be harmful to human health (Basu, Yoffe, Hills, & Lustig, 2013; Bray & Popkin, 2014; Hu, 2013), and masking agents with hard-to-pronounce names may not match a consumer's ideas of acceptable "clean label" (Zink, 1997) ingredients.

2.6.2 Bitter Compound Extraction

Extraction of specific chemicals from agricultural products, such as the removal of the bitter chemical caffeine from coffee beans, is well-known (Ramalakshmi et al., 1999; Zosel, 1981). In the case of coffee, raw coffee beans may have their caffeine extracted

prior to being roasted using a variety of solvents, such as supercritical CO_2, or by aqueous extraction called the Swiss Water Process, which better maintains flavor compounds and their precursors (Ramalakshmi et al., 1999). Supercritical extraction de-bittering processes have been devised for cacao as well (Margolis, Chiovini, & Pagliaro, 1989). However, application of such methods to cacao may be problematic because post-harvest processing steps (e.g. fermentation and drying) lead to an altered complex flavor profile that is preferable to that of pre-processed raw cacao (Afoakwa et al., 2008), and such positive attributes can be lost during a de-bittering process (Ramalakshmi et al., 1999).

2.7 Quantification of Bitter Compounds in Cacao and Chocolate

2.7.1 From Theobromine to Diketopiperazines in the Pre-HPLC Era

In the mid-19th century, a white crystalline bitter compound was first extracted from the seeds of *Theobroma cacao* by Russian chemist M. Alexander Woskresensky (1842), who thereupon named it theobromine after the plant from which it was derived (Woskresensky, 1842). At this early date, it was already apparent that theobromine is not well soluble in water, alcohol, or ether (Woskresensky, 1842). Some 80 years later, theobromine solubility remained a problem, as British chemist and cacao researcher Raymond V. Wadsworth (1920) wrote, perhaps with some annoyance, that "theobromine is only soluble with difficulty in most liquids." Wadsworth was, however, able to better define the problem; he noted that theobromine would dissolve well in neither cold water, ether, petroleum ether, chloroform, cool ethanol, nor any of a

variety of other organic solvents. Relative success was only achieved when certain solvents were brought to the very point of boiling (e.g., water, ethanol, and trichloroethane)(Wadsworth, 1920). One interesting exception was uncovered in the solvent aniline, which could dissolve more than 10 times as much theobromine as could pure water (i.e., 8 g theobromine/100 mL solution when boiling) (Wadsworth, 1920). In later decades of the 20th century, it would be discovered by other enterprising individuals that N,N-dimethylformamide also performs well, successfully dissolving almost 5 times as much theobromine as pure water (both at 25 °C) (Martin, Paruta, & Adjei, 1981). The difficulty of theobromine dissolution in most solvents would be eventually attributed to strong intermolecular hydrogen bonding compared to more methylated xanthines, such as caffeine, which are more soluble in water (Hockfield et al., 1982; Spiller, 1997). Perhaps it is due to this affinity for itself that theobromine will only dissolve well in cool water when its pH is significantly acidic or basic (Spiller, 1997).

Recent discoveries aside, Wadsworth found that even with the relative issues of theobromine solubility, he was still able to develop a multi-step extraction and titrimetric analytical method for the determination of theobromine in cacao (Wadsworth, 1921; Wadsworth, 1922). Yet his method for extracting and analyzing theobromine (Wadsworth, 1921; Wadsworth, 1922), at one time adopted by the AOAC (Kreiser & Martin Jr, 1978; Timbie et al., 1978), had problems with precision, specificity, and applicability (Kay & Haywood, 1946; Kreiser et al., 1978; Timbie et al., 1978), and methods subsequently developed by others, both titrimetric and spectrophotometric,

were found equally lacking and laborious (Kreiser et al., 1978; Timbie et al., 1978). However, by the late middle of the 20th century, the advent of High Performance Liquid Chromatography (HPLC) promised to make analysis of both caffeine and theobromine in cacao remarkably simple, while at the same time increasing sensitivity and specificity (Timbie et al., 1978). Perhaps it was for this reason that the new analytical tool also led to a flood of articles describing the analysis of other bitter compounds in cacao, such as flavan-3-ols (Kim & Keeney, 1983; 1984) and DKPs (Van der Greef, Tas, Nijssen, Jetten, & Höhn, 1987). Yet, before the HPLC era was truly in full flight, enterprising German chemist Wilhelm Pickenhagen, operating comfortably in the space between the new and old analytical chemistry paradigms, set the tone for the future study of bitterness in cacao (Pickenhagen et al., 1975). Indeed, Pickenhagen's work on bitterness in cacao—of Brazilian origin with unknown roast profile--in addition to confirming the importance of chemicals in the methylxanthine class (i.e., theobromine and caffeine), was the first to suggest that certain bitter cyclic dipeptides called 2,5-diketopiperazines (DKPs), initially described in 1849 (Pickenhagen et al., 1975), were not only present in cacao, but that, together with theobromine, they were at the very heart of the typical and unique cacao bitterness (Pickenhagen et al., 1975). Pickenhagen's analytical method took advantage of aqueous acetone as extraction solvent, various preparatory chromatographic techniques, and finally gas chromatography (GC) paired with mass spectrometry (MS), in the process of extracting and identifying a variety of DKPs (Pickenhagen et al., 1975). Pickenhagen's work also showed that, unlike theobromine, DKPs are soluble in many solvents and their mixtures, including water, acetone, methanol, chloroform, and

interestingly, like theobromine (Martin et al., 1981), also in dimethylformamide (Pickenhagen et al., 1975). Indeed, Pickenhagen's (1975) findings concerning bitterness and DKPs in cacao served as the conceptual foundation for the work of researchers in cacao bitterness for at least the next 30 years (Bonvehí et al., 2000; Stark et al., 2006; Stark & Hofmann, 2005b; Van der Greef et al., 1987). Thus, the stage was set for the aforementioned HPLC instrumentation advances to bear fruit, and quickly at that.

2.7.2 The Modern Era of Cacao Analysis: HPLC and Beyond

2.7.2.1 Methylxanthines

As formerly noted, the apparent first use of HPLC instrumentation to analyze cacao extracts was in the late 1970s, for the purposes of methylxanthine determination (Kreiser et al., 1978; Timbie et al., 1978). The interest in methylxanthines in cacao, and theobromine in particular, since as early as 1842 (Woskresensky) may be due to the fact that theobromine, of all taste-active compounds in cacao, is present at a significantly higher concentration than any other (Stark et al., 2006). Credit must certainly be given to Kreiser and Martin's (1978) work, influenced by D.J. Timbie's doctoral studies under the direction of Philip Keeney (Timbie et al., 1978), which introduced a clear and simple methodology for methylxanthine analysis of cacao and chocolate via HPLC. In fact, the general method, consisting of fat extraction of the sample, boiling water extraction of the resulting residue, filtration and/or centrifugation, and then reversed-phase (RP) HPLC analysis on C18 column, with ultraviolet (UV) detection at approximately 280 nm, has been used almost to this very day, with only minor modifications, by many

researchers interested in methylxanthine analysis in cacao and chocolate products (Bonvehí et al., 2000; De Camargo & Toledo, 1999; Stark et al., 2006; Vries, Johnson, & Heroff, 1981; Zoumas et al., 1980).

2.7.2.2 Flavan-3-ols/ Catechins

Analysis of (-)-epicatechin, a flavan-3-ol, was the next HPLC-based analytical method for a bitter compound in cacao to be developed, also by the Keeney research group at Penn State (Kim et al., 1983; 1984). This method was similar to that of methylxanthines, as it also called for defatting of the sample followed by solvent extraction—80% aqueous acetone--of the residue, along with final analysis by RP-HPLC-UV on C18 column at 280 nm (Kim et al., 1983; 1984). This method did also include intermediate steps of acetone evaporation followed by solid-phase sample cleanup (Kim et al., 1983; 1984), both of which are similar to Pickenhagen's (1975) pre-HPLC era sample cleanup methodology using preparative chromatography. Interestingly, though this method, a precursor to simultaneous analysis of methylxanthines and flavan-3-ols, had clearly visible caffeine peaks in the chromatograms produced by Kim and Keeney, where they noted specifically that there was "excellent" (Kim et al., 1983) separation between caffeine and (-)-epicatechin, methylxanthine determination was not formally a part of the method. Future methods for extraction and analysis of one or more flavan-3-ols in cacao were predominantly modified versions of this Kim and Keeney (1983) method, including those by Villeneuve (1989), Porter (1991), Bonvehí (1997), Adamson (1999), and Natsume (2000). Addition of fractionation was common when a larger number of flavan-

3-ols, including dimers, trimers and other oligomers were of interest (Adamson et al., 1999; Natsume et al., 2000; Porter et al., 1991). Interestingly, like theobromine and DKPs, lower molecular weight flavan-3-ols have also been found to be readily soluble in N,N-dimethylformamide (Del Rio et al., 2010; Turkmen, Sari, & Velioglu, 2006), and the resulting extracts have been noted to exhibit better chromatographic separation than those extracted with equivalent solvents containing acetone, ethanol, or methanol (Turkmen et al., 2006).

2.7.2.3 2,5-Diketopiperazines

Twelve years after Pickenhagen's seminal work (1975) on analysis of DKPs in cacao and their importance as bitter compounds for chocolate, J. Van Der Greef used HPLC paired with MS to identify and quantify ten DKPs in cacao of unknown origin and roast profile. His method defatted ground samples, extracted them with refluxed hot water followed by solid-phase extraction sample cleanup, and finally fractionation with preparatory liquid chromatography, giving a final fraction containing the peptides injected onto a Polygosil 5-60 silica column and analyzed with HPLC-MS (Van der Greef et al., 1987). Thus, Van Der Greef's method (1987) borrows significantly from Pickenhagen (1975), while also showing that pure aqueous extraction of DKPs is effective. Soon after, Prodolliet (1993) confirmed that aqueous extraction of DKPs would work for a variety of foods, including chocolate, with recoveries of 96.1% to 105.0% using RP-HPLC-UV instrumentation set to 214 nm, while including Carrez I and II reagents for sample cleanup of lipids and proteins (Prodolliet et al., 1993). Several years later, J. Serra

Bonvehí (2000), showed success with this method for cacao, further supporting the idea that aqueous extraction of DKPs can be effective, while also confirming that detection in the lower part of the UV range (i.e. 210 nm) works well for DKPs (Bonvehí et al., 2000). In this substantial work, Bonvehí analyzed a variety of cacao samples with different roast profiles, and revealed that indeed higher roast temperatures increase overall DKP concentrations (Bonvehí et al., 2000). Interestingly, methylxanthines and DKPs are both present in Bonvehí's published chromatograms (2000), but the compounds were not extracted and quantified using equivalent methods for each class. Instead, Bonvehí continued to rely upon the older modified Kreiser & Martin (1978) method for methylxanthine analysis. However, as we shall soon see, the age of HPLC-based simultaneous analysis of multiple bitter compound classes had already begun.

2.7.2.4 Simultaneous Analysis of Multiple Bitter Compound Classes

Several aforementioned methods for analysis of methylxanthines, flavan-3-ols, and DKPs in cacao and chocolate have shown that compounds from more than one class could be extracted at the same time, for example with pure water (Bonvehí et al., 2000) or aqueous acetone (Kim et al., 1983), and these compounds also appeared in the same chromatograms, making simultaneous analysis seem quite reasonable. However, it was Hisaya Terada's work (1992) on catechins and methylxanthines in various teas, rather than chocolate, that appears to be the first record of the simultaneous analysis of both of these classes of bitter compounds. Terada used a 40% aqueous ethanol extraction with C18 solid-phase sample cleanup, and carried out analysis on a C18 column with

HPLC-UV set to 207 nm (Terada et al., 1992); the method was generally effective, with recoveries of most compounds falling within the 90%-100% range (Terada et al., 1992). However, recovery of theobromine seemed to pose a slight problem with a recovery range of 50%-100% (Terada et al., 1992), perhaps reflecting the previously discussed difficulties with theobromine solubility. A decade later, a similar method first used for catechin analysis in various foods (Arts & Hollman, 1998) was attempted with chocolate (Tokusoglu & Ünal, 2002) for simultaneous analysis of methylxanthines and catechins, substituting refluxed 90% aqueous methanol at 80°C for Terada's ethanolic extraction, and eschewing the solid-phase cleanup (Tokusoglu et al., 2002). Recoveries of all compounds ranged from about 95% - 100%, including for troublesome theobromine (Tokusoglu et al., 2002). Several years later, Timo Stark, of Thomas Hofmann's research group, used a two-pronged approach (Stark et al., 2006), extracting flavan-3-ols and DKPs together in a modified Kim and Keeney method (1983) with 70% aqueous acetone, while extracting methylxanthines using a modified version of the original Kreiser and Martin boiling-water extraction method (1978), but with the addition of Carrez solutions for sample cleanup in the manner of Prodolliet (1993). Just two years later, Charles H. Risner (2008) exhibited with incredible detail, including validation based upon a NIST reference chocolate standard, that for simultaneous analysis of methylxanthines, (+)-catechin, and (-)-epicatechin, either water, 15% aqueous methanol, or 80% aqueous methanol, all with 1% acetic acid added, performed relatively well as extraction solvents.

Given the successes and near-successes in simultaneous determination of different classes of bitter compounds in cacao with HPLC instrumentation (Bonvehí et al., 2000; Kreiser et al., 1978; Terada et al., 1992; Tokusoglu et al., 2002), and the significant amount that has been learned over the last 170 years about solubility and extraction of flavan-3-ols and DKPs in general (Kim et al., 1983; Pickenhagen et al., 1975), and theobromine in particular (Tarka Jr & Hurst, 1998; Wadsworth, 1920; Woskresensky, 1842), it is reasonable to suspect that a method for simultaneous analysis of the most important bitter compounds in each class (i.e., theobromine, caffeine, (-)-epicatechin, and cyclo(L-Pro-L-Val)) (Stark et al., 2006), should be possible.

CHAPTER 3

3 METHOD FOR THE SIMULTANEOUS DETERMINATION OF IMPORTANT BITTER COMPOUNDS FROM THREE DIFFERENT CHEMICAL CLASSES IN CHOCOLATE

3.1 Introduction

3.1.1 Quantification of Bitter Compounds in Cacao and Chocolate

3.1.1.1 From Theobromine to Diketopiperazines in the Pre-HPLC Era

In the mid-19th century, a white crystalline bitter compound was first extracted from the seeds of *Theobroma cacao* by Russian chemist M. Alexander Woskresensky (1842), who thereupon named it theobromine after the plant from which it was derived (Woskresensky, 1842). At this early date, it was already apparent that theobromine is not well soluble in water, alcohol, or ether (Woskresensky, 1842). Some 80 years later, theobromine solubility remained a problem, as British chemist and cacao researcher Raymond V. Wadsworth (1920) wrote, perhaps with some annoyance, that "theobromine is only soluble with difficulty in most liquids." Wadsworth was, however, able to better define the problem; he noted that theobromine would dissolve well in neither cold water, ether, petroleum ether, chloroform, cool ethanol, nor any of a variety of other organic solvents. Relative success was only achieved when certain solvents were brought to the very point of boiling (e.g., water, ethanol, and trichloroethane)(Wadsworth, 1920). One interesting exception was uncovered in the solvent aniline, which could dissolve more than 10 times as much theobromine as could pure water (i.e., 8 g theobromine/100 mL solution when boiling (Wadsworth, 1920). In

later decades of the 20th century, it would be discovered by other enterprising individuals that N,N-dimethylformamide also performs well, successfully dissolving almost 5 times as much theobromine as pure water (both at 25 °C) (Martin et al., 1981). The difficulty of theobromine dissolution in most solvents would be eventually attributed to strong intermolecular hydrogen bonding compared to more methylated xanthines, such as caffeine, which are more soluble in water (Hockfield et al., 1982; Spiller, 1997). Perhaps it is due to this affinity for itself that theobromine will only dissolve well in cool water when its pH is significantly acidic or basic (Spiller, 1997).

Recent discoveries aside, Wadsworth found that even with the relative issues of theobromine solubility, he was still able to develop a multi-step extraction and titrimetric analytical method for the determination of theobromine in cacao (Wadsworth, 1921; Wadsworth, 1922). Yet his method for extracting and analyzing theobromine (Wadsworth, 1921; Wadsworth, 1922), at one time adopted by the AOAC (Kreiser et al., 1978; Timbie et al., 1978), had problems with precision, specificity, and applicability (Kay et al., 1946; Kreiser et al., 1978; Timbie et al., 1978), and methods subsequently developed by others, both titrimetric and spectrophotometric, were found equally lacking and laborious (Kreiser et al., 1978; Timbie et al., 1978). However, by the late middle of the 20th century, the advent of High Performance Liquid Chromatography (HPLC) promised to make analysis of both caffeine and theobromine in cacao remarkably simple, while at the same time increasing sensitivity and specificity (Timbie

et al., 1978). Perhaps it was for this reason that the new analytical tool also led to a flood of articles describing the analysis of other bitter compounds in cacao, such as flavan-3-ols (Kim et al., 1983; 1984) and DKPs (Van der Greef et al., 1987). Yet, before the HPLC era was truly in full flight, enterprising German chemist Wilhelm Pickenhagen, operating comfortably in the space between the new and old analytical chemistry paradigms, set the tone for the future study of bitterness in cacao (Pickenhagen et al., 1975). Indeed, Pickenhagen's work on bitterness in cacao—of Brazilian origin with unknown roast profile--in addition to confirming the importance of chemicals in the methylxanthine class (i.e., theobromine and caffeine), was the first to suggest that certain bitter cyclic dipeptides called 2,5-diketopiperazines (DKPs), initially described in 1849 (Pickenhagen et al., 1975), were not only present in cacao, but that, together with theobromine, they were at the very heart of the typical and unique cacao bitterness (Pickenhagen et al., 1975). Pickenhagen's analytical method took advantage of aqueous acetone as extraction solvent, various preparatory chromatographic techniques, and finally gas chromatography (GC) paired with mass spectrometry (MS), in the process of extracting and identifying a variety of DKPs (Pickenhagen et al., 1975). Pickenhagen's work also showed that, unlike theobromine, DKPs are soluble in many solvents and their mixtures, including water, acetone, methanol, chloroform, and interestingly, like theobromine (Martin et al., 1981), also in dimethylformamide (Pickenhagen et al., 1975). Indeed, Pickenhagen's (1975) findings concerning bitterness and DKPs in cacao served as the conceptual foundation for the work of researchers in cacao bitterness for at least the next 30 years (Bonvehí et al., 2000; Stark et al., 2006; Stark et al., 2005b;

Van der Greef et al., 1987). Thus, the stage was set for the aforementioned HPLC instrumentation advances to bear fruit, and quickly at that.

3.1.1.2 The Modern Era of Cacao Analysis with HPLC: Compounds in a Single Class

3.1.1.2.1 Methylxanthines

As previously noted, the apparent first use of HPLC instrumentation to analyze cacao extracts was in the late 1970s, for the purposes of methylxanthine determination (Kreiser et al., 1978; Timbie et al., 1978). The interest in methylxanthines in cacao, and theobromine in particular, since as early as 1842 (Woskresensky) may be due to the fact that theobromine, of all taste-active compounds in cacao, is present at a significantly higher concentration than any other (Stark et al., 2006). Credit must certainly be given to Kreiser and Martin's (1978) work, influenced by D.J. Timbie's doctoral studies under the direction of Philip Keeney (Timbie et al., 1978), which introduced a clear and simple methodology for methylxanthine analysis of cacao and chocolate via HPLC. In fact, the general method, consisting of fat extraction of the sample, boiling water extraction of the resulting residue, filtration and/or centrifugation, and then reversed-phase (RP) HPLC analysis on C18 column, with ultraviolet (UV) detection at approximately 280 nm, has been used almost to this very day, with only minor modifications, by many researchers interested in methylxanthine analysis in cacao and chocolate products (Bonvehí et al., 2000; De Camargo et al., 1999; Stark et al., 2006; Vries et al., 1981; Zoumas et al., 1980).

3.1.1.2.2 Flavan-3-ols/ Catechins

Analysis of (-)-epicatechin, a flavan-3-ol, was the next HPLC-based analytical method for a bitter compound in cacao to be developed, also by the Keeney research group at Penn State (Kim et al., 1983; 1984). This method was similar to that of methylxanthines, as it also called for defatting of the sample followed by solvent extraction—80% aqueous acetone--of the residue, along with final analysis by RP-HPLC-UV on C18 column at 280 nm (Kim et al., 1983; 1984). This method did also include intermediate steps of acetone evaporation followed by solid-phase sample cleanup (Kim et al., 1983; 1984), both of which are similar to Pickenhagen's (1975) pre-HPLC era sample cleanup methodology using preparative chromatography. Interestingly, though this method, a precursor to simultaneous analysis of methylxanthines and flavan-3-ols, had clearly visible caffeine peaks in the chromatograms produced by Kim and Keeney, where they noted specifically that there was "excellent" (Kim et al., 1983) separation between caffeine and (-)-epicatechin, methylxanthine determination was not formally a part of the method. Future methods for extraction and analysis of one or more flavan-3-ols in cacao were predominantly modified versions of this Kim and Keeney (1983) method, including those by Villeneuve (1989), Porter (1991), Bonvehí (1997), Adamson (1999), and Natsume (2000). Addition of fractionation was common when a larger number of flavan-3-ols, including dimers, trimers and other oligomers were of interest (Adamson et al., 1999; Natsume et al., 2000; Porter et al., 1991). Interestingly, like theobromine and DKPs, lower molecular weight flavan-3-ols have also been found to be readily soluble in N,N-dimethylformamide (Del Rio et al., 2010; Turkmen et al., 2006), and the resulting

extracts have been noted to exhibit better chromatographic separation than those extracted with equivalent solvents containing acetone, ethanol, or methanol (Turkmen et al., 2006).

3.1.1.2.3 2,5-Diketopiperazines

Twelve years after Pickenhagen's seminal work (1975) on analysis of DKPs in cacao and their importance as bitter compounds for chocolate, J. Van Der Greef used HPLC paired with MS to identify and quantify ten DKPs in cacao of unknown origin and roast profile. His method defatted ground samples, extracted them with refluxed hot water followed by solid-phase extraction sample cleanup, and finally fractionation with preparatory liquid chromatography, giving a final fraction containing the peptides injected onto a Polygosil 5-60 silica column and analyzed with HPLC-MS (Van der Greef et al., 1987). Thus, Van Der Greef's method (1987) borrows significantly from Pickenhagen (1975), while also showing that pure aqueous extraction of DKPs is effective. Soon after, Prodolliet (1993) confirmed that aqueous extraction of DKPs would work for a variety of foods, including chocolate, with recoveries of 96.1% to 105.0% using RP-HPLC-UV instrumentation set to 214 nm, while including Carrez I and II reagents for sample cleanup of lipids and proteins (Prodolliet et al., 1993). Several years later, J. Serra Bonvehí (2000), showed success with this method for cacao, further supporting the idea that aqueous extraction of DKPs can be effective, while also confirming that detection in the lower part of the UV range (i.e. 210 nm) works well for DKPs (Bonvehí et al., 2000). In this substantial work, Bonvehí analyzed a variety of cacao samples with different

roast profiles, and revealed that indeed higher roast temperatures increase overall DKP concentrations (Bonvehí et al., 2000). Interestingly, methylxanthines and DKPs are both present in Bonvehí's published chromatograms (2000), but the compounds were not extracted and quantified using equivalent methods for each class. Instead, Bonvehí continued to rely upon the older modified Kreiser & Martin (1978) method for methylxanthine analysis. However, as we shall soon see, the age of HPLC-based simultaneous analysis of multiple bitter compound classes had already begun.

3.1.1.3 The Modern Era of Cacao Analysis with HPLC: Simultaneous Analysis of Multiple Bitter Compound Classes

Several aforementioned methods for analysis of methylxanthines, flavan-3-ols, and DKPs in cacao and chocolate have shown that compounds from more than one class could be extracted at the same time, for example with pure water (Bonvehí et al., 2000) or aqueous acetone (Kim et al., 1983), and these compounds also appeared in the same chromatograms, making simultaneous analysis seem quite reasonable. However, it was Hisaya Terada's work (1992) on catechins and methylxanthines in various teas, rather than chocolate, that appears to be the first record of the simultaneous analysis of both of these classes of bitter compounds. Terada used a 40% aqueous ethanol extraction with C18 solid-phase sample cleanup, and carried out analysis on a C18 column with HPLC-UV set to 207 nm (Terada et al., 1992); the method was generally effective, with recoveries of most compounds falling within the 90%-100% range (Terada et al., 1992). However, recovery of theobromine seemed to pose a slight problem with a recovery

range of 50%-100% (Terada et al., 1992), perhaps reflecting the previously discussed difficulties with theobromine solubility. A decade later, a similar method first used for catechin analysis in various foods (Arts et al., 1998) was attempted with chocolate (Tokusoglu et al., 2002) for simultaneous analysis of methylxanthines and catechins, substituting refluxed 90% aqueous methanol at 80°C for Terada's ethanolic extraction, and eschewing the solid-phase cleanup (Tokusoglu et al., 2002). Recoveries of all compounds ranged from about 95% - 100%, including for troublesome theobromine (Tokusoglu et al., 2002). Several years later, Timo Stark, of Thomas Hofmann's research group, used a two-pronged approach (Stark et al., 2006), extracting flavan-3-ols and DKPs together in a modified Kim and Keeney method (1983) with 70% aqueous acetone, while extracting methylxanthines using a modified version of the original Kreiser and Martin boiling-water extraction method (1978), but with the addition of Carrez solutions for sample cleanup in the manner of Prodolliet (1993). Just two years later, Charles H. Risner (2008) exhibited with incredible detail, including validation based upon a NIST reference chocolate standard, that for simultaneous analysis of methylxanthines, (+)-catechin, and (-)-epicatechin, either water, 15% aqueous methanol, or 80% aqueous methanol, all with 1% acetic acid added, performed relatively well as extraction solvents.

Given the successes and near-successes in simultaneous determination of different classes of bitter compounds in cacao with HPLC instrumentation (Bonvehí et al., 2000; Kreiser et al., 1978; Terada et al., 1992; Tokusoglu et al., 2002), and the significant

amount that has been learned over the last one hundred years or more about solubility and extraction of flavan-3-ols and DKPs in general (Kim et al., 1983; Pickenhagen et al., 1975), and theobromine in particular (Tarka Jr et al., 1998; Wadsworth, 1920; Woskresensky, 1842), it is reasonable to suspect that a method for simultaneous analysis of the most important bitter compounds in each class (i.e., theobromine, caffeine, (-)-epicatechin, and cyclo(L-Pro-L-Val)) (Stark et al., 2006), should be possible. Specifically, given the discoveries of researchers over the past 40 years, showing that N,N-Dimethylformamide performs well as a solvent for compounds in all three known classes of important bitter compounds in cacao, i.e., diketopiperazines (Pickenhagen et al., 1975), the methylxanthine theobromine (Martin et al., 1981), and lower molecular weight flavan-3-ols such as epicatechin (Del Rio et al., 2010; Turkmen et al., 2006), even showing better chromatographic separation in those cases (Turkmen et al., 2006), our goal was to develop a method that could simultaneously, precisely, and accurately extract and quantify compounds from all three of these classes using aqueous N,N-Dimethylformamide extraction and HPLC analysis.

3.2 Materials & Methods

3.2.1 Materials

High-purity standard compounds were purchased from several manufacturers (Table 3.1).

Table 3. 1: Standard compound information

Compounds	Manufacturer	Lot #	Part #	Purity Factor
Theobromine	Sigma Aldrich	BCBM9560V	T4500	0.987
Caffeine	Sigma Aldrich	MKBW1243V	W222402	1.000
(+)-Catechin*	Sigma Aldrich	WXBC5812V	C1251	0.868
(-)-Epicatechin	Sigma Aldrich	BCBW4134	E1753	0.951
Procyanidin B2	Cayman	0506392-7	0506392-7	1.000
cyclo(Pro-Val)	BAChem	1026889	1026889	0.999

*COA: (+)-Catechin hydrate 99%, adjusted for water weight (12.35%) to 86.8%

The following solvents were obtained from Fisher Scientific: HPLC Water Lot# 189570, HPLC Methanol Lot# 187803, 88% Formic Acid Lot# 180909, 0.5 M NaOH in methanol Ricca Lot# 4802L81, and 0.5 M HCl Ricca Lot# 4803D64. The following solvent was obtained from Millipore Sigma: N,N-Dimethylformamide Lot# SHBL1229.

Mobile phase mixtures used for HPLC analysis consisted of an aqueous phase (Channel A), containing 1140 µL 88% formic acid, and 50 mL methanol, made to volume in a 1 L volumetric flask with HPLC water, and an organic phase (Channel B), containing 1140 µL 88% formic acid, made to volume in a 1 L vol flask with HPLC methanol.

Solvent for standard mixtures (SMS) and sample extraction solvent (ES) consisted of 700 mL N,N-Dimethylformamide mixed with 340 mL HPLC water, which gives an

approximate volume of 1000 mL once the mixed solvent temperature is at equilibrium with room temperature.

3.2.1.1 Additional Materials and Analytical Equipment

HPLC analysis was performed with an Agilent 1200 series HPLC with Chemstation for LC Rev. B.02.01 [244] software, with Kinetex C18 and Zorbax Eclipse XDB-C18 columns, both 4.6 mm x 250 mm, run in sequence, with a photodiode array detector (DAD).

Ghanaian chocolate liquor prepared from roasted cacao was obtained from Patric Chocolate (Columbia, MO, USA).

Cacao liquor was ground with a whirling-blade coffee grinder (Hamilton Beach model #80335R), and extraction carried out using an ultrasonic water bath (Kendal model # HB-S-23DHT), vortexer (Vortex-Genie model # 12-812-v1), motorized pipette filler (RPI Corp model # 248646) paired with borosilicate 25 mL volumetric pipette, and a centrifuge with 6-slot fixed-angle rotor (Eppendorf model #5804, rotor model # F-34-6-38). 50 mL conical centrifuge tubes (Celltreat #229421) were also used. Subsamples of ground chocolate liquor and all standard compounds were massed with an analytical balance (DeltaRange model # AG204).

3.2.2 Methods

3.2.2.1 Standard Curve Preparation

Stock solutions of all compounds were prepared by quantitatively transferring accurately massed quantities of standard compounds into a single 100 mL volumetric

flask and making to volume with SMS (section 3.2.1) as per the amounts listed in Table 3.2. Calculated stock concentrations, adjusted to account for standard purity, are included (Table 3.2).

Table 3. 2: Stock solution information for each compound

Compounds	Expected Concentration (mg/g) in cacao*	Expected Concentration (mg/mL) in 50 mL extraction of 1 g cacao	Target Stock Concentration (mg/mL)	Accurate Mass in 100 mL Vol Flask (mg)	Purity-Adjusted Stock Concentration (mg/mL)
Theobromine	11.4519445	0.22903889	0.3	30.1	0.297
Caffeine	1.01328342	0.020265668	0.1	10.0	0.100
(+)-Catechin	0.68617464	0.013723493	0.1	10.4	0.090
(-)-Epicatechin	2.50000938	0.050000188	0.15	14.9	0.142
Procyanidin B2	1.20505716	0.024101143	0.1	10.0	0.100
cyclo(Pro-Val)	1.7423075	0.03484615	0.1	11.0	0.110

***Based upon (Stark et al., 2006)**

After stock preparation was complete, stock was diluted in SMS to prepare an external standard curve consisting of 5 approximately equidistant dilutions for each standard compound (Table 3.3).

Table 3. 3: Standard curve pure compound concentrations

Compounds	1 (mg/mL)	2 (mg/mL)	3 (mg/mL)	4 (mg/mL)	5 (mg/mL)
Theobromine	0.297	0.249	0.200	0.152	0.104
Caffeine	0.084	0.067	0.051	0.035	0.020
(+)-Catechin	0.061	0.046	0.032	0.018	0.0005
(-)-Epicatechin	0.096	0.072	0.050	0.028	0.001
Procyanidin B2	0.067	0.051	0.035	0.020	0.0010
cyclo(Pro-Val)	0.074	0.056	0.038	0.022	0.0005

Prepared external standard solutions in amber-tinted 2 mL HPLC autosampler vials were analyzed via HPLC. Mobile phase (section 3.2.1), was purged with high-purity helium for

15 minutes. The system was then purged of mobile phase already present in the lines for 5 minutes at 4 mL/min to avoid potential issues with gas bubble formation. Inlet frits in the mobile phase containers were also visually checked at this time to be sure that no air bubbles were trapped behind them. Initial conditions for the HPLC method (i.e., 0.75 mL/min flow rate at 100% Channel A) were then run for at least 20 minutes until backpressure had stabilized. Finally, the HPLC method was run in full, consisting of a gradient (Table 3.4) with 10-minute post-time for re-equilibration of the columns at 100% Channel A conditions.

Table 3. 4: HPLC gradient for standard quantification

Time (min)	% Channel A	% Channel B
0	100	0
25	70	30
50	50	50
55	0	100
Post (10 min)	100	0

A 0.75 mL/min flowrate was maintained throughout, with column temperature held at 30 °C. Injection volume was 10 µL, with post-injection methanol needle wash between injections. All wavelengths between 190 nm and 300 nm were monitored at 1 nm increments. Ultraviolet absorbances of all standards were ultimately used for quantification purposes using either 205 nm or 280 nm wavelengths (Table 3.5). An initial blank was run to assure that separation conditions would be the same for each

standard injection. Prepared standards were run in order from least concentrated solution to most concentrated solution for each compound.

Table 3. 5: Wavelengths used for quantification of standard compounds

Compounds	UV Wavelength (nm)
Theobromine	280
Caffeine	280
(+)-Catechin	205
(-)-Epicatechin	280
Procyanidin B2	205
cyclo(Pro-Val)	205

Once the peak area related to UV absorbance at specific wavelengths was determined for each peak of each compound (Table 3.5), standard curves were plotted with known analyte concentration on the x-axis and peak size on the y-axis, and the line describing each curve derived as a mathematical equation allowing for the estimation of the concentration of each compound in each sample. Coefficients of determination (R^2) were also calculated for each curve, with values greater than 0.99 being deemed acceptable (tables 3.6 through 3.12).

3.2.2.2 Precision Study:

A precision study was conducted using the extraction method outlined herein by extracting 9 subsamples across two days of a sample of Ghanaian chocolate liquor prepared from roasted cacao (i.e., 3 extractions on the first day and 6 on the second). The extraction method was carried out as follows:

The day prior to sample extraction, approximately 20.0 g of Ghanaian chocolate liquor made from roasted nibs was massed into a whirling-blade grinder, which was then placed in a -20 °C freezer overnight. Extraction solvent (ES) was prepared (section 3.2.1) a day prior to analysis to allow it to equilibrate at room temperature. The next morning, the sample of liquor held overnight at -20 °C was ground for 50 seconds; once complete, the lid of the grinder remained affixed until the contents had warmed to approximately room temperature (i.e., greater than ca. 30 minutes). Once the grinder and contents had warmed, a subsample of approximately 1.0 g from the ground liquor sample was massed accurately from the grinder directly into each of three 50 mL centrifuge tubes.

Next, using a motorized pipette filler and volumetric pipette, 25 mL of ES was dispensed into each of the centrifuge tubes. Each tube was then vortexed until the ground liquor was wetted and homogenously mixed, for approximately 1-2 seconds per tube.

Each tube was then placed into an ultrasonic water bath, verified to contain deionized water up to the marked fill line of the bath. A custom copper cooling coil was then turned on, allowing maintenance of temperature in the ultrasonic bath at approximately 30 °C. The bath was then set for 30 minutes of sonication and turned on. The tubes were then vortexed consecutively at 10 min elapsed and 20 min elapsed for approximately 1 second each, before being replaced in the running sonicator, and being rotated one slot clockwise. Once the 30 minutes had elapsed, the tubes were centrifuged (Eppendorf model #5804, rotor model # F-34-6-38) for a duration of 2 minutes, at 11,000 rpm (15,557 x g) with acceleration setting of 9, and braking setting of

7. The supernatant from each tube was then decanted into fresh, labeled tubes. This entire process was then carried out once more with 25 mL of fresh ES for the residue of each subsample, resulting in a second extraction of the same material. The supernatant of the first and second extractions were then combined in the same collection tube, giving a prepared sample extracted by 50 mL ES in all cases. Each prepared subsample extraction in centrifuge tube with tightened cap was then gently inverted twice by hand, and finally vortexed for approximately 1 second to assure homogeneity of the contents. 1 mL of each prepared sample was then passed through a 0.45 µm nylon membrane filter into an amber-tinted 2 mL HPLC autosampler vial and immediately analyzed. The concentrations of each compound were then calculated using the standard curve described above (tables 3.6 through 3.12) and the Relative Standard Deviation (RSD) (a.k.a. percent coefficient of variation) was calculated for the values of Day 1 (Table 3.13). This method, as just described, was then repeated with 6 additional subsamples of the ground Ghana cacao liquor sample, and RSD values were similarly calculated, along with weighted average RSD values accounting for the results of both days (Table 3.13).

3.2.2.3 Recovery Study

3.2.2.3.1 Flavonoid Recovery Study

For this portion of the recovery study, Ghanaian chocolate liquor was pre-extracted, spiked with known concentrations of standard compounds and then extracted anew, analyzed, and quantified in order to compare the expected concentration with the

actual concentration. To do this, approximately 20.0 g Ghanaian liquor was measured into a whirling-blade grinder (see section 3.2.1.1), covered with the lid and placed in a freezer at -20 °C overnight. The next morning the liquor was ground for 60 seconds. After allowing the liquor to warm, still covered, to approximately room temperature (i.e., > ca. 30 minutes), approximately 1.0 g of the ground liquor was accurately massed into each of twelve 50 mL conical centrifuge tubes in all. The liquor in these tubes was extracted using 25 mL of HPLC water, by vortexing until wetted, and then sonicating for 30 minutes, with the temperature set to 50 °C. The tubes were then centrifuged (see section 3.2.1.1) for a duration of 5 minutes, at 11,000 rpm (15,557 x g) with acceleration setting of 9, and braking setting of 7. The supernatant was decanted, and this process was repeated thrice more so that the cacao liquor in each tube was extracted four times with 25 mL HPLC water each time for a total of 100 mL in all. The tubes were then transferred to a vacuum oven held at approximately 50 °C with vacuum until completely dried (i.e., approximately 24 hours). The compacted and dried pellet in each tube was then broken and crushed with a small spatula until it was reduced to at least 10 small pieces with associated residue dust. Three tubes were held as controls, and no standard solution was added. Nine additional tubes were spiked with known concentrations of the three flavonoids (i.e., (+)-catechin, (-)-epicatechin, procyanidin B2) made up in pure HPLC methanol, at 3 different concentration levels (i.e., 40%, 70%, and 100% expected concentration) with 3 tubes per each of the three levels. Expected concentration was based upon values in (Stark et al., 2006).

After spiking, all tubes were sonicated for 5 minutes at between 25 °C and 30 °C. All tubes were then placed in an unheated vacuum oven with vacuum switched on until all methanol had evaporated (i.e., approximately 20 hours). All tubes were then extracted. The extraction method consisted of the following:

Extraction solvent (ES) was prepared as described in Materials section 3.2.1 a day prior to analysis to allow it to equilibrate at room temperature. Next, using a motorized pipette filler and volumetric pipette, 25 mL of ES was dispensed into each of the centrifuge tubes. Each tube was then vortexed until the spiked liquor was wetted and homogenously mixed, for approximately 1-2 seconds per tube.

Each tube was then placed into an ultrasonic water bath, verified to contain deionized water up to the marked fill line. A custom copper cooling coil was then turned on, allowing maintenance of temperature in the ultrasonic bath at approximately 30 °C. The bath was then set for 30 minutes of sonication and turned on. The tubes were then vortexed consecutively at 10 min elapsed and 20 min elapsed for approximately 1 second each, before being replaced in the running sonicator, and being rotated one slot clockwise. Once the 30 minutes had elapsed, the tubes were centrifuged (section 3.2.1.1) for a duration of 8 minutes, at 11,000 rpm (15,557 x g) with acceleration setting of 9, and braking setting of 7. The supernatant from each tube was then decanted into fresh, labeled tubes, and this was repeated until all tubes had been centrifuged and decanted. This entire process was then carried out once more with 25 mL of fresh ES for the residue of each tube, resulting in a second extraction of the same material. The

supernatant of the first and second extractions were then combined in the same collection tube, giving a prepared sample extracted by 50 mL of ES in all cases. Each mixed subsample extraction with tightened cap was then gently inverted twice by hand, and finally vortexed for approximately 1 second to assure homogeneity of the contents. 1 mL of each prepared sample was then passed through a 0.45 µm nylon membrane filter into an amber-tinted 2 mL HPLC autosampler vial for immediate analysis.

3.2.2.3.2 Caffeine and Cyclo(Proline-Valine) Recovery Study

For this portion of the recovery study, Ghanaian chocolate liquor was pre-extracted, spiked with known concentrations of standard compounds and then extracted, analyzed, and quantified in order to compare the expected concentration with the actual concentration. To do this, approximately 20.0 g Ghanaian liquor was measured into a whirling-blade grinder, covered with the lid, and placed in a freezer at -20 °C overnight. The next morning the liquor was ground for 60 seconds. After allowing the liquor to warm, still covered, to approximately room temperature (i.e., > ca. 30 minutes), approximately 1.0 g of the ground liquor was accurately massed into nine 50 mL conical centrifuge tubes in all. The liquor in these tubes was extracted using 25 mL of HPLC water, by vortexing until wetted, and then sonicating for 30 minutes, with the temperature set to 50 °C. The tubes were then centrifuged (section 3.2.1.1) for a duration of 5 minutes, at 11,000 rpm (15,557 x g) with acceleration setting of 9, and braking setting of 7. The supernatant was decanted, and this process was repeated thrice more so that the cacao liquor in each tube was extracted four times with 25 mL

HPLC water each time for a total of 100 mL in all. The tubes were then transferred to a vacuum oven held at approximately 50 °C with vacuum until completely dried (i.e., approximately 24 hours). The compacted and dried pellet in each tube was then broken and crushed with a small spatula until it was reduced to at least 10 small pieces with associated residue dust. Because control tubes had already been analyzed, in this case, 9 tubes were prepared, being spiked with known concentrations of caffeine and cyclo(Pro-Val), made up in pure HPLC methanol, at 3 different concentration levels (i.e., 40%, 70%, and 100% expected concentration), with 3 tubes per each of the three levels. Expected concentration was based upon values in (Stark et al., 2006).

After spiking, all tubes were sonicated for 5 minutes at between 25 °C and 30 °C. All tubes were then placed in an unheated vacuum oven with vacuum switched on until all methanol had evaporated (i.e., approximately 20 hours). All tubes were then extracted. The extraction method consisted of the following:

Extraction solvent (ES) was prepared as described in Materials section 3.2.1 a day prior to analysis to allow it to equilibrate at room temperature. Next, using a motorized pipette filler and volumetric pipette, 25 mL of ES was dispensed into each of the centrifuge tubes. Each tube was then vortexed until the spiked liquor was wetted and homogenously mixed, for approximately 1-2 seconds per tube.

Each tube was then placed into an ultrasonic water bath, verified to contain deionized water up to the marked fill line. A custom copper cooling coil was then turned on, allowing maintenance of temperature in the ultrasonic bath at approximately 30 °C. The

bath was then set for 30 minutes of sonication and turned on. The tubes were then vortexed consecutively at 10 min elapsed and 20 min elapsed for approximately 1 second each, before being replaced in the running sonicator, and being rotated one slot clockwise. Once the 30 minutes had elapsed, the tubes were centrifuged (section 3.2.1.1) for a duration of 8 minutes, at 11,000 rpm (15,557 x g) with acceleration setting of 9, and braking setting of 7. The supernatant from each tube was then decanted into fresh, labeled tubes, and this was repeated until all tubes had been centrifuged and decanted. This entire process was then carried out once more with 25 mL of fresh ES for the residue of each tube, resulting in a second extraction of the same material. The supernatant of the first and second extractions were then were then combined in the same collection tube, giving a prepared sample extracted by 50 mL of ES in all cases. Each subsample extraction in a centrifuge tube with tightened cap was then gently inverted twice by hand, and finally vortexed for approximately 1 second to assure homogeneity of the contents. 1 mL of each prepared sample was then passed through a 0.45 µm nylon membrane filter into an amber-tinted 2 mL HPLC autosampler vial for immediate analysis.

3.2.2.3.3 Theobromine Recovery Study

Given the low solubility of theobromine (Stark et al., 2006), this portion of the recovery study was handled separately, as methanol, the solvent used for all other compounds in this study, results in a saturated solution of theobromine that is too dilute to be useful as a stock solution. Rather, the increased solubility of theobromine in alkaline solution

(Spiller, 1997), allowed the use of 0.5 M methanolic NaOH in this study. For this portion of the recovery study, Ghanaian chocolate liquor was pre-extracted, spiked with known concentrations of theobromine and then extracted, analyzed, and quantified in order to compare the expected concentration with the actual concentration. To do this, approximately 20.0 g Ghanaian liquor was measured into a whirling-blade grinder (section 3.2.1.1), covered with the lid and placed in a freezer at -20 °C overnight. The next morning the liquor was ground for 60 seconds. After allowing the liquor to warm, still covered, to approximately room temperature (i.e., > ca. 30 minutes), approximately 1.0 g of the ground liquor was accurately massed into nine 50 mL conical centrifuge tubes in all. The liquor in these tubes was extracted using 25 mL of HPLC water, by vortexing until wetted, and then sonicating for 30 minutes, with the temperature set to 50 °C. The tubes were then centrifuged for a duration of 5 minutes, at 11,000 rpm (15,557 x g) with acceleration setting of 9, and braking setting of 7. The supernatant was decanted, and this process was repeated thrice more so that the cacao liquor in each tube was extracted four times with 25 mL HPLC water each time for a total of 100 mL in all. The tubes were then transferred to a vacuum oven held at approximately 50 °C with vacuum until completely dried (i.e., approximately 24 hours). The compacted and dried pellet in each tube was then broken and crushed with a small spatula until it was reduced to at least 10 small pieces with associated residue dust. Because control tubes had already been analyzed, in this case, 9 tubes were prepared, being spiked with known concentrations of theobromine made up in 0.5 M NaOH in methanol, at 3 different concentration levels (i.e., 40%, 70%, and 100% expected concentration), with 3

tubes per each of the three levels. Expected concentration was based upon values in (Stark et al., 2006).

After spiking, all tubes were sonicated for 5 minutes at between 25 °C and 30 °C. The solvent in each tube was then neutralized with 0.5 M HCl, and the tubes were then placed in an unheated vacuum oven with vacuum switched on until all solvent had evaporated (i.e., approximately 20 hours). All tubes were then extracted. The extraction method consisted of the following:

Extraction solvent (ES) was prepared (section 3.2.1) a day prior to analysis to allow it to equilibrate at room temperature. Next, using a motorized pipette filler and volumetric pipette, 25 mL of ES was dispensed into each of the centrifuge tubes. Each tube was then vortexed until the spiked liquor was wetted and homogenously mixed, for approximately 1-2 seconds per tube.

Each tube was then placed into an ultrasonic water bath, verified to contain deionized water up to the marked fill line. A custom copper cooling coil was then turned on, allowing maintenance of temperature in the ultrasonic bath at approximately 30 °C. The bath was then set for 30 minutes of sonication and turned on. The tubes were then vortexed consecutively at 10 min elapsed and 20 min elapsed for approximately 1 second each, before being replaced in the running sonicator, and being rotated one slot clockwise. Once the 30 minutes had elapsed, the tubes were centrifuged (section 3.2.1.1) for a duration of 8 minutes, at 11,000 rpm (15,557 x g) with acceleration setting of 9, and braking setting of 7. The supernatant from each tube was then decanted into

fresh, labeled tubes, and this was repeated until all tubes had been centrifuged and decanted. This entire process was then carried out once more with 25 mL of fresh ES for the residue of each tube, resulting in a second extraction of the same material. The supernatant of the first and second extractions were then combined in the same collection tube, giving a prepared sample extracted by 50 mL of ES in all cases. Each mixed subsample extraction with tightened cap was then gently inverted twice by hand, and finally vortexed for approximately 1 second to assure homogeneity of the contents. 1 mL of each prepared sample was then passed through a 0.45 µm nylon membrane filter into an amber-tinted 2 mL HPLC autosampler vial for immediate analysis.

3.2.2.3.4 HPLC Analytical Method

Prepared spiked pre-extracted samples in amber-tinted 2 mL HPLC autosampler vials were analyzed via HPLC. Mobile phase (section 3.2.1) was purged with high-purity helium for 15 minutes. The system was then purged of mobile phase already in the lines for 5 minutes at 4 mL/min, to avoid potential issues with gas bubble formation. Inlet frits in the mobile phase containers were also visually checked at this time to be sure that no air bubbles were trapped behind them. Initial conditions for the HPLC method (i.e., 0.75 mL/min flow rate at 100% Channel A) were then run for at least 20 minutes until backpressure had stabilized. Finally, the HPLC method was run in full, consisting of a gradient (Table 3.4) with 10-minute post-time for re-equilibration of the columns at 100% Channel A conditions. A 0.75 mL/min flowrate was maintained throughout the run, with column temperature held at 30 °C. Injection volume was 10 µL, with the

needle being washed with methanol in between injections. All wavelengths between 190 nm and 300 nm were monitored at 1 nm increments. All compounds of interest were ultimately quantified using either 205 nm or 280 nm wavelengths (Table 3.5). An initial blank was run to assure that separation conditions would be exactly the same for each actual sample injection. Prepared spiked pre-extracted samples were run in order of lowest to highest concentration (i.e., non-spiked controls (0%), 40%, 70%, and then 100% spiked pre-extracted samples).

3.3 Results & Discussion

3.3.1 Standard Curve Results

Standard curve concentrations, area counts, equations, and coefficients of determination can be found below (Tables 3.6 through 3.12). The wavelengths used were chosen to balance maximum absorbance and selectivity for each compound, both in a standard mixture, and in samples of cacao extracts from multiple origins with multiple roast profiles. The HPLC method was optimized to achieve these aims as well. All coefficients of determination (R^2) greater than 0.99 were considered acceptable for quantification purposes.

Table 3. 6: Standard curve equations and R^2 values

Analyte	Standard Curve Equation	Coefficient of Determination (R^2)
Theobromine @ 280 nm	y = 34690x + 548.01	0.9952
Caffeine @ 280 nm	y = 32039x + 20.99	0.9996
(-)-Epicatechin @ 280 nm	y = 9489.7x - 1.5051	0.9999
(+)-Catechin @205 nm	y = 148382x + 276.06	0.9947
Procyanidin B2 @ 205 nm	y = 108918x + 40.658	0.9999
Cyclo(Pro-Val) @ 205 nm	y = 25110x + 14.485	0.9987

Y= Area count and X= analyte concentration (mg/mL)

Table 3. 7: Theobromine standard curve concentrations and area counts

Concentration mg/mL	Area (280 nm)
0.297	**10632.0**
0.249	**9339.0**
0.200	**7691.8**
0.152	**5803.2**
0.104	**4015.2**

Table 3. 8: Caffeine standard curve concentrations and area counts

Concentration mg/mL	Area (280 nm)
0.084	**2682.0**
0.067	**2204.1**
0.051	**1659.5**
0.035	**1145.9**
0.020	**650.68**

Table 3. 9: (-)-Epicatechin standard curve concentrations and area counts

Concentration mg/mL	Area (280 nm)
0.096	908.0
0.072	680.4
0.050	469.0
0.028	266.8
0.001	13.5

Table 3. 10: (+)-Catechin standard curve concentrations and area counts

Concentration mg/mL	Area (205 nm)
0.061	9027.2
0.046	7227.2
0.032	5225.5
0.018	3114.2
0.0005	72.5

Table 3. 11: Procyanidin B2 standard curve concentrations and area counts

Concentration mg/mL	Area (205 nm)
0.067	7368.5
0.051	5595.9
0.035	3849.5
0.020	2270.0
0.0010	114.7

Table 3. 12: cyclo(Proline-Valine) standard curve concentrations and area counts

Concentration mg/mL	Area (205 nm)
0.074	1867.0
0.056	1402.1
0.038	1026.0
0.022	561.2
0.0005	14.6

3.3.2 Precision Study Results

As mentioned previously (section 3.222), the precision study took place across two days. RSD values for day 1, day 2, and a weighted average were calculated (Table 3.13) given that 3 samples were analyzed on day 1 and 6 samples on day 2. In all cases, all RSD values were well below 5%, with the exception of cyclo(Pro-Val), which was found to be 7.65%. Low concentrations of this compound are expected to have resulted in this slightly high RSD. We deemed precision for extraction of all compounds to be acceptable under the circumstances.

Table 3. 13: Precision study relative standard deviation (RSD) for each compound

Compounds	Day 1 (3 samples)	Day 2 (6 Samples)	Weighted Average
Theobromine	0.97%	0.55%	0.69%
(+)-Catechin	2.63%	2.25%	2.38%
Procyanidin B2	2.11%	0.82%	1.25%
cyclo(Pro-Val)	5.52%	8.71%	7.65%
Caffeine	1.02%	0.52%	0.69%
(-)-Epicatechin	1.31%	0.83%	0.99%

3.3.3 Recovery Study Results

Recovery of all analytes of interest from pre-extracted cacao liquor, with spikes at 40%, 70% and 100% of expected concentrations (Stark et al., 2006) are included in Table 3.14. All values for all compounds and spike levels fell within a range of 75% and 115% with global averages across all spike levels across all three chemical classes falling within a range of 87% to 107% for the method used, and average recovery levels across all spike levels for individual compounds ranging from 90% to 106%. Given known difficulties

with theobromine solubility (Wadsworth, 1920), and flavan-3-ol stability (Zhu et al., 2002), these values were deemed to be acceptable.

Table 3. 14: Recovery Results

Analyte	40%	70%	100%	Average Per Compound Across All Levels
Theobromine @ 280 nm	81%	98%	106%	95%
Caffeine @ 280 nm	82%	106%	111%	100%
(-)-Epicatechin @ 280	90%	99%	100%	96%
(+)-Catechin @205 nm	75%	96%	100%	90%
Procyanidin B2 @ 205	102%	107%	109%	106%
Cyclo(Pro-Val) @ 205 nm	91%	115%	113%	106%
Average Across All Compounds By Level	**87%**	**103%**	**107%**	**99%**

3.4 Conclusions

The aqueous 70% N,N-Dimethylformamide solvent described above, as used for standard curve development and extraction of bitter compounds of interest from cacao, functioned quite well. It resulted in acceptable standard curves for all compounds, with good separation on our dual C18 column configuration. Additionally, the precision study suggests that extraction of all compounds was quite repeatable, with only slightly high (i.e., > 5% RSD), but still acceptable (i.e., < 10% RSD), precision for one compound (i.e., cyclo(Pro-Val)). Additionally, our recovery study showed the true benefit of this extraction method, as even without the use of boiling water, theobromine showed good recovery at levels close to those expected in cacao (Stark et al., 2006), as did caffeine, all flavan-3-ols analyzed, and cyclo(Pro-Val). For this reason, this method of extraction and analysis is suggested for use in the study of bitter compounds in cacao in the future, especially when efficient extraction and separation is a necessity.

CHAPTER 4

4 QUANTITATIVE ANALYSIS OF IMPORTANT BITTER COMPOUNDS IN CHOCOLATE MADE FROM THREE DIFFERENT ORIGINS OF CACAO USING EIGHT ROAST PROFILES

4.1 Introduction

Chocolate is made from the fermented, dried, and roasted seeds of the *Theobroma cacao* tree (Aprotosoaie et al., 2016a), which in addition to being an important agricultural food crop (Aprotosoaie et al., 2016a), is known to contain bioactive flavonoid polyphenols with beneficial health effects in humans (Andújar et al., 2012; Aprotosoaie et al., 2016b). Such effects include improvement of antioxidant status (Cooper et al., 2008), positive impacts on cardiovascular health (Andújar et al., 2012; Ellam et al., 2013) and endocrine system function (Andújar et al., 2012; Bowser et al., 2017; Strat et al., 2016), and association with cancer prevention (Ishaq et al., 2017), LDL cholesterol reduction (Ellam et al., 2013) and reduction of obesity and related conditions (Matsui et al., 2005; Min et al., 2013; Sun et al., 2015). However, despite strong evidence for the healthful qualities of cacao, products which have the highest levels of cacao flavonoids of all eating-chocolate, such as high-cacao-percentage dark chocolate (Langer et al., 2011; Miller et al., 2009), are known to be quite bitter, a taste modality that is not readily appreciated by many humans (Drewnowski et al., 2000; Fischer et al., 2005). Though the complex causes of bitterness in cacao are still not completely understood, it has long been known that two methylxanthines, theobromine and caffeine, impart bitterness to cacao and chocolate (Stark et al., 2006;

Woskresensky, 1842). Additionally, certain flavan-3-ols, sometimes called catechins, which are a class of the aforementioned healthy bioactive polyphenolic flavonoids, also found in tea (Scharbert et al., 2005), are likewise bitter and present in cacao at levels above the detection threshold (Stark et al., 2005a)(e.g., (-)-epicatechin (Stark et al., 2006)). Yet, what else is known of bitterness in cacao is sparse and can seem contradictory. Pickenhagen's (1975) work was the first to describe the importance in cocoa of the cyclic dipeptides called 2,5-diketopiperazines (DKPs), while suggesting some form of interaction between theobromine and DKPs, leading to the "typical bitterness" (Pickenhagen et al., 1975) of cacao. Yet Pickenhagen's assertions were only confirmed with mixed results by Bonvehí and Coll (2000), who ultimately determined that the incredible complexity of bitterness in roasted cacao, and the role of DKPs within it, required further evaluation. The most recent significant work on bitterness in cacao (Stark et al., 2006) suggested for the first time that a DKP called cyclo(L-Pro-L-Val) is the most important compound responsible for the bitterness in cacao. In this work (Stark et al., 2006), the contribution to bitterness in cacao of compounds in the methylxanthine, flavan-3-ol and 2,5-diketopiperazine classes was confirmed, though the effect of roasting itself was not specifically studied. Roasting, considered by some to be the most important step in cacao processing (Aprotosoaie et al., 2016a), results in the creation of the aforementioned bitter diketopiperazines (DKPs) from peptides (Rizzi, 1989; Ziegleder, 2017), and darker roasts, particularly at higher temperatures, appear to increase DKP levels the most (Bonvehí et al., 2000), whereas unroasted cocoa contains virtually no DKPs (Bonvehí et al., 2000). However, roasting is required to alter the harsh,

unpleasant flavor of raw cacao, and to develop characteristic cocoa aroma (Ziegleder, 2009). Additionally, roasting alters the concentrations of epicatechin, and its epimers and oligomers (Kothe et al., 2013; Stanley et al., 2018), compounds which are both bitter and astringent (Stark et al., 2006), sometimes in unexpected ways, but seemingly related to varietal (Kothe et al., 2013). Loss of epicatechin at temperatures over 70 °C occurs, and at a roasting temperature of 120°C catechin content has been seen to increase by approximately 650% in fermented cacao (Payne et al., 2010). This is probably due to epimerization of epicatechin to catechin (De Taeye et al., 2014; Oracz et al., 2015; Payne et al., 2010) and decomposition of procyanidins (Oracz et al., 2015), probably first to epicatechin followed by epimerization (Zhu et al., 2002). Finally, it has been previously noted that a small loss of methylxanthines from the cotyledons to the cacao shell may occur during roasting, apparently increasing with degree of roast (Wadsworth, 1922).

Given the apparent complexity of the changes and potential changes in important bitter compounds in cacao during roasting, a study designed specifically to analyze the changes in the most important of these compounds (Stark et al., 2006) due to variation in roast profile seemed a necessity. This is particularly the case since Stark's otherwise impressive work (Stark et al., 2006) on bitterness in cacao could say nothing about the impact of roasting on those compounds, as data were based upon a single cacao sample with an undefined roasting treatment. Additionally, previous work had already noted that differences in DKP formation are dependent upon roast profile but perhaps also

origin (Bonvehí et al., 2000; Rizzi, 1989). Yet Stark (2006) left us uncertain about the variation between bitter compounds in roasted cacao from different populations/origins, given that he focused on Ghana only. Therefore, fourteen years after Stark's (2006) work, much is still to be learned about the variation in bitter-compound composition in cacao within and between different cacao populations/origins and across different roast profiles, even considering that the significance of the three primary classes of bitter compounds—DKPs, flavan-3-ols, and methylxanthines—seems clear. Indeed, the continued dearth of work in the area of cacao bitterness over the last decade, while faced with a growing desire for healthy, functional versions of foods such as chocolate (Mintel, 2016; Pacyniak, 2014; 2016), makes research into the impact of cacao roasting on bitter compounds all the more timely. Our goal is, therefore, to study the impact of roasting (time and temperature) on samples of three different origins of cacao with multiple roast treatments chosen in accordance with Response Surface Methodology (RSM), given that the relationship of roast profile and various chemicals of interest has already been studied successfully using RSM in other agricultural products such as coffee (Madihah, Zaibunnisa, Norashikin, Rozita, & Misnawi, 2012; Mendes, de Menezes, Aparecida, & Da Silva, 2001), sesame seeds (Kahyaoglu & Kaya, 2006), pistachio nuts (Kahyaoglu, 2008), and even Indonesian cocoa (Farah & Zaibunnisa, 2012), though bitterness was not investigated. In each of these cases, Response Surface Methodology (RSM) was employed successfully as a guide for design and analysis of the experiment and elucidation of the complex changes that occur in seed foods during roasting, and should therefore allow us, too, to

model changes in chemical concentrations in the chocolate liquors based upon roasting conditions (time and temperature) and origin.

4.2 Materials & Methods

4.2.1 Materials

4.2.1.1 Cacao

All cacao from each of three origins (i.e., Madagascar, Ghana, Peru), was fermented and dried at origin before being packed into approximately 65 kg burlap sacks, themselves packed into steel shipping containers, and then transported to cacao warehouses in the United States by cargo ship.

4.2.1.1.1 Madagascar:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from three separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total (i.e., 7 kg per lot): Lots #2182432 (2018 harvest), #2672711 (2018 harvest), #1427425 (2018 harvest). This cacao was obtained from Guittard Chocolate. Upon receipt, all three samples of this origin were blended into a homogenous composite, and hand-sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.), unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro Supergrain Premium RT bags with high vapor and gas barrier characteristics until roasted.

4.2.1.1.2 Ghana:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from three separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total (i.e., 7 kg per lot): Lot #19003223 (2018/2019 main harvest), #482019 (2018/2019 main harvest), #3729 (2017/2018 main harvest). This cacao was obtained from Guittard Chocolate. Upon receipt, all three samples of this origin were blended into a homogenous composite, and hand-sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.), unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro Supergrain Premium RT bags with high vapor and gas barrier characteristics until roasted.

4.2.1.1.3 Peru:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from two separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total from two separate lots covering multiple months as follows: 7 kg was sampled from a lot consisting of the January through early March 2018 harvest, and 14 kg (7 kg x 2) was sampled from a lot consisting of the late March to late June 2018 harvest. This cacao was obtained from Marañon Cacao. Upon receipt, all three samples of this origin were blended into a homogenous composite, and hand-

sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.), unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro Supergrain Premium RT bags with high vapor and gas barrier characteristics until roasted.

4.2.1.1.4 Cacao Quality Parameters

Cacao quality parameters related to phenotype, ripeness, fermentation, and drying, of a representative portion (180 g) of each 21 kg composite are detailed in Table 4.1.

Table 4. 1: **Preliminary statistics on composite samples from each origin**

Origin:	**Ghana**	**Madagascar**	**Peru**
Bean-Count:	89 beans per 100 g	86 beans per 100 g	74 beans per 100g
Slaty Color:	1%	0%	0%
Fully Purple:	1%	0%	26%
Partly Purple:	17%	6%	18%
Total Slaty or Purple:	19%	6%	44%
Light Brown:	0%	27%	7%
Medium Brown:	58%	62%	27%
Dark Brown:	23%	5%	10%
Total Brown:	81%	94%	44%
White Beans:	0%	0%	12%
Moisture Content:	6.13 %	6.68%	5.83%
Internal Mold:	0%	0%	0%
Infested:	1%	0%	1%

4.2.1.2 Chemicals

The following solvents were obtained from Fisher Scientific: HPLC Water Lot# 189570, HPLC Methanol Lot# 187803, and 88% Formic Acid Lot# 180909. The following solvent was obtained from Millipore Sigma: N,N-Dimethylformamide Lot# SHBL1229.

Mobile phase mixtures used for HPLC analysis consisted of an aqueous phase (Channel A), containing 1140 µL 88% formic acid, and 50 mL methanol, made to volume in a 1 L volumetric flask with HPLC water, and an organic phase (Channel B), containing 1140 µL 88% formic acid, made to volume in a 1 L volumetric flask with HPLC methanol.

Cacao liquor extraction solvent (ES) consisted of 700 mL N,N-Dimethylformamide mixed with 340 mL HPLC water, which gives an approximate volume of 1000 mL depending upon room temperature, once the mixed solvent is at equilibrium with the room temperature.

4.2.1.3 Additional Materials and Analytical Equipment

HPLC analysis was performed with an Agilent 1200 series HPLC with Chemstation for LC Rev. B.02.01 [244] software, with Kinetex C18 and Zorbax Eclipse XDB-C18 columns, both 4.6 mm x 250 mm, run in sequence, with a photodiode array detector (DAD).

The roasting oven consisted of a 120 V Binder model # FD56 precision laboratory oven with forced air convection.

Roasting temperatures were monitored and recorded with a 4-channel Omega model #RDXL4SD datalogger with Type-K thermocouples (Omega part #5sc-tt-k-30-36).

Room temperature and humidity during roasting were monitored with a SensorPush #HT1 device with humidity calibrated with Boveda Low-RH One-Step 32% calibration kit model #CAL32-SM, and temperature and humidity verified with a Bacharach Sling Psychrometer model #0012-7012.

Preliminary cacao cooling was performed by a Lasker 20" box fan model #B20200, placed under ¼" stainless mesh bolted to Cambro food bin. Cacao cracking was performed with a custom 305 mm 3-roll CrankandStein Cracker, and winnowing was performed using a custom Food-Grade winnower with 4" clear FDA-approved for food contact PVC, with air flow powered by a Grizzly model #:G0441 3HP cyclone.

Chocolate liquor was ground using a Spectra 11 Stone Wet Grinder model #101 and strained through a Kitchenaid model # KES161OHOBW fine-mesh strainer. Cooled and solidified chocolate liquor was wrapped in aluminum foil and vacuum sealed in FoodSaver brand multi-layer vacuum bag with nylon vapor barrier.

Cacao liquor extraction was carried out using an ultrasonic water bath (Kendal model # HB-S-23DHT), vortexer (Vortex-Genie model # 12-812-v1), motorized pipette filler (RPI Corp model # 248646) paired with borosilicate 25 mL volumetric pipette, and a centrifuge with 6-slot fixed-angle rotor (Eppendorf model #5804, rotor model # F-34-6-38). 50 mL conical centrifuge tubes (Celltreat #229421) were also used. Subsamples for triplicate analysis were massed with an analytical balance (DeltaRange model # AG204).

4.2.2 Roasting Experimental Design

For the roasting experimental design (Figure 4.1 and Table 4.2), temperature range (i.e., 24°C to 171°C) and time range (i.e., 0 to 80 minutes) were chosen as reasonable ranges for potential modification of bitterness based upon literature regarding common roasting temperatures (Afoakwa et al., 2008; Ziegleder, 2017) and the advice of chocolate professionals; this includes a raw treatment at 24°C (approximate room temperature) for 0 minutes as a control. Combinations of time and temperature that are impossible, judged so extreme so as to be burnt, or that in essence repeat other combinations (e.g., 0 minutes and any temperature aside from 24°C, or 24°C and any time combination aside from 0 minutes), were excluded. The specific time and temperature combinations within these ranges were chosen as part of a Response Surface Methodology (RSM) approach, specifically using JMP 14.0.0 software and an I-optimal algorithm, which minimizes average variance of prediction for potential model coefficients (Jones & Goos, 2012; Myers, Montgomery, & Anderson-Cook, 2016; Oyejola & Nwanya, 2015), while at the same time seeking to minimize covariance of the potential coefficients of the model (Oyejola et al., 2015). The resulting design is an irregularly shaped non-rotatable design, akin to a central composite design (CCD), with a duplicate centerpoint to allow for pure error estimation. In addition to the previously mentioned reasons, this model was chosen because standard CCDs have already been commonly and successfully used in roasting optimization experiments (Farah et al., 2012; Kahyaoglu, 2008; Lee, Yoo, Lee, Kwon, & Pyun, 2001; Madihah et al., 2012; Mendes et al., 2001; Özdemir & Devres, 2000), while it has also been noted in the last

decade that computer-based optimal design algorithms are effective at helping to choose designs for complex problems that cannot easily be solved with standard design types, such as in the example of irregular experimental regions (Goos & Jones, 2011; Jones et al., 2012).

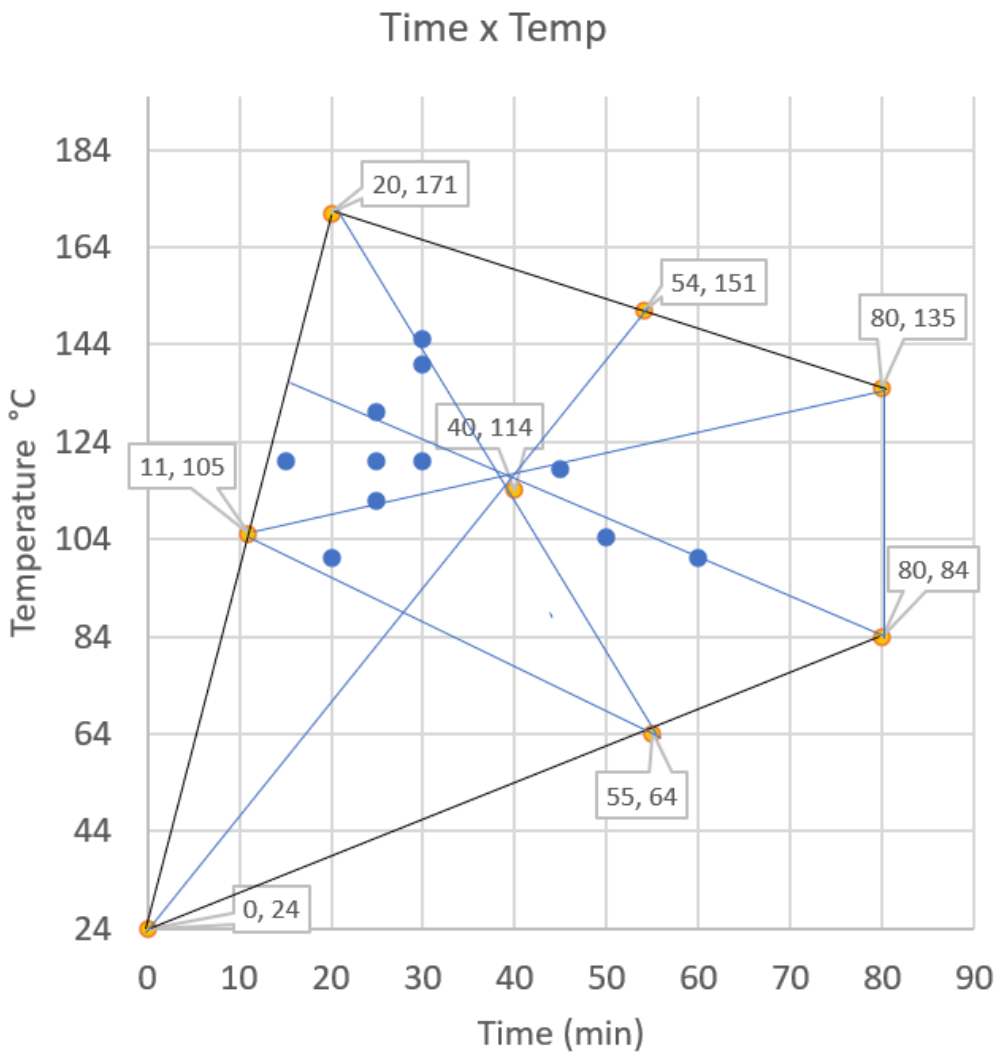

Figure 4. 1: Two-dimensional representation of roasting experimental design:

Golden points are in the experimental design, and blue points are based upon literature (Beckett et al., 2017; Whymper, 1921) and advice of chocolate professionals. All time and temperature combinations are repeated for each of the three origins (Madagascar, Ghana, and Peru).

Table 4. 2: Randomized modified I-optimal experimental design for roasting

Minutes	Temperature ˚C	Origin	Roast Order
40	114	Peru	1
40	114	Ghana	2
54	151	Mad	3
80	84	Peru	4
0	24	Peru	5
20	171	Ghana	6
80	84	Ghana	7
20	171	Mad	8
54	151	Peru	9
55	64	Peru	10
40	114	Peru	11
0	24	Ghana	12
40	114	Ghana	13
40	114	Mad	14
11	105	Peru	15
55	64	Mad	16
11	105	Mad	17
54	151	Ghana	18
11	105	Ghana	19
40	114	Mad	20
80	135	Peru	21
80	135	Ghana	22
0	24	Mad	23
80	84	Mad	24
55	64	Ghana	25
80	135	Mad	26
20	171	Peru	27

4.2.3 Roasting and Winnowing Method

410 g of cacao, origin specified by experimental design (Table 4.2) was weighed onto each of two stainless steel mesh roasting trays (820 g total). The cacao was spread into an evenly distributed single-bean layer on each tray. A thermocouple was placed in a hole made with a steel needle in a bean of approximate average size in the middle of each tray. The thermocouples were then connected to the datalogger, and the datalogger was set to record. The oven was preheated to the setpoint required in the roasting design (Table 4.2), with the convection fan on. Once the setpoint was reached, 10 additional minutes were allowed to elapse to assure preheating completion. Once pre-heating was completed, the timer was set to the required length for the roast in question (Table 4.2), and the two trays were loaded into the oven with one tray right above the fan and the other right below it. The oven door was immediately closed and the timer was started.

Once the roast was complete, the oven was turned off, and the door was immediately opened. Forced-air cooling of the roasted cacao was employed to minimize carry-over roasting through immediate winnowing, augmented with preliminary fan cooling for higher temperature roasts, to attain sufficient temperature decreases for all treatments (i.e., a drop to 50°C or less within 5 minutes). The end-result of this cooling and winnowing process, when augmented with manual removal of any remaining visible shell pieces, was room temperature cacao nibs, the main ingredient of chocolate liquor.

Each roast was completed in duplicate on the same day, and nibs from the roasted and winnowed duplicate treatments blended until homogeneous. Therefore, 1640 g of cacao was roasted in total for each treatment. Of this 1640 g, 1000 g of blended roasted cacao nibs were placed in a food-grade polypropylene bag, with air substantially removed, and labeled with all identifying information. The nibs were then stored in a room with temperature below 19.5 °C with relative humidity (RH) at approximately 40% or less. The nibs were turned into chocolate liquor within 48 hours in all cases.

4.2.4 Chocolate Liquor Production Method

For chocolate liquor production, the stone bowl and grinding stones of the Spectra 11 wet grinder were preheated to approximately 49 °C to 54 °C overnight in a warming cabinet. Once the stone grinder components were mounted to the machine, it was immediately turned on and 1000 g of nibs were added slowly, about 20 mL at a time until a thick paste began to form, with the lid then being replaced, until the paste became less viscous. Nibs were then once again added in this way until the paste began to appear glossy and less viscous, allowing for faster addition. All 1000 g of nibs were added to the grinder within approximately 20 minutes. Once all nibs were added, the internal elements of the wet grinder were scraped with a silicone food-grade spatula, and this scraping process was repeated at 30-minute intervals three additional times, for four scrapings total. The chocolate liquor was refined to a smooth texture for a total of 8 hours from the time that all nibs were added to each batch at a temperature of approximately 50 °C, maintained by production of frictional heat from the refining

process itself. The chocolate liquor was then poured through a fine-mesh culinary strainer, into a poly food-storage container using a silicone spatula to remove as much liquor as possible from the wet grinder. A yield of approximately 900 g was achieved in all cases. The chocolate liquor was then covered with an air-tight lid, placed in a storage room at or below 19.5 °C with RH% at approximately 40% or less. All chocolate liquor batches were allowed to solidify at storage temperature, then unmolded, separated into two approximately equivalent portions—one each for chemical and sensory analysis—with each wrapped individually with an aluminum sheet, vacuum sealed in a multi-layer vacuum bag with nylon vapor barrier, and stored at or below 19.5 °C with RH% at approximately 40% or less until they were prepared for chemical analysis.

4.2.5 Sample Preparation & Extraction Method

4.2.5.1 Preliminary Sample Preparation

The day prior to sample extraction, approximately 20.0 g of chocolate liquor from each sample to be analyzed was massed accurately into its own whirling-blade coffee grinder (Hamilton Beach model #80335R), which was then placed in a -20 °C freezer overnight. Additionally, approximately 40 g of each chocolate liquor sample was wrapped within an aluminum sheet, vacuum sealed in a multi-layer vacuum bag with nylon vapor barrier, and then sent for moisture and crude fat analysis. The remaining portion of the sample was then re-wrapped in aluminum foil, vacuum sealed once more, and then returned to storage at or below 19.5 °C with RH% at approximately 40% or less. Three samples were

prepared in this way for analysis each day, in the same order that the roasts were performed (Table 4.2).

4.2.5.2 Final Sample Preparation and Extraction

Extraction solvent (ES) was prepared as described above a day prior to analysis to allow it to come to room temperature. The samples of liquor held overnight at -20 °C were, one at a time, and in order (Table 4.2), each ground for 50 seconds; once complete, the lid of the grinder remained affixed until the contents had warmed to approximately room temperature (i.e., greater than 30 minutes). Once the grinders and contents had warmed, approximately 1.0 g of ground liquor from each sample was massed accurately from each grinder directly into 50 mL centrifuge tubes in triplicate, resulting in three subsamples for each of the three samples analyzed, or nine tubes total, in a given day.

Next, using the motorized pipette filler, 25 mL of ES was dispensed into each of the nine centrifuge tubes. Each tube was then vortexed until the ground liquor was wetted and homogenously mixed, for approximately 1-2 seconds per tube.

Each tube was then placed into an ultrasonic water bath, verified to contain deionized water up to the fill line. A custom copper cooling coil was then turned on, allowing maintenance of temperature in the ultrasonic bath at approximately 30 °C. The bath was then set for 30 minutes of sonication and turned on. The tubes were then vortexed consecutively at 10 min elapsed and 20 min elapsed for approximately 1 second each, before being replaced in the running sonicator, and being rotated one slot clockwise.

Once the 30 minutes had elapsed, the tubes were centrifuged for a duration of 2 minutes, at 11,000 rpm (15,557 x g) with acceleration setting of 9, and braking setting of 7 (Eppendorf model #5804, rotor model # F-34-6-38). The supernatant from each tube was then decanted into fresh, labeled tubes, and this was repeated until all tubes had been centrifuged and decanted. This entire process was then carried out once more with 25 mL of fresh ES for the residue of each subsample, resulting in a second extraction of the same material. The supernatant of the first and second extractions were then combined in the same collection tube, giving a prepared sample extracted by 50 mL starting ES in all cases. Each mixed subsample extraction with tightened cap was then gently inverted twice by hand, and finally vortexed for approximately 1 second to assure homogeneity of the contents. 1 mL of each prepared sample was then passed through a 0.45 µm nylon membrane filter into an amber-tinted 2 mL HPLC autosampler vial and immediately analyzed.

4.2.6 HPLC Analytical Method

Prepared samples in amber-tinted 2 mL HPLC autosampler vials were analyzed via HPLC. Mobile phase, prepared as detailed in section 4.2.1.2, was purged with high-purity helium for 15 minutes. The system was then purged of mobile phase already in the lines for 5 minutes at 4 mL/min, to avoid potential issues with gas bubble formation. Inlet frits in the mobile phase containers were also visually checked at this time to be sure that no air bubbles were trapped behind them. Initial conditions for the HPLC method (i.e., 0.75 mL/min flow rate at 100% Channel A) were then run for at least 20 minutes

until backpressure had stabilized. Finally, the HPLC method was run in full, consisting of a gradient as seen below (Table 4.3) with 10-minute post-time for re-equilibration of the columns at 100% Channel A conditions. A 0.75 mL/min flowrate was maintained throughout this cycle, with column temperature held at 30 °C. Injection volume was 10 µL, with the needle being washed with methanol in between injections. All wavelengths between 190 nm and 300 nm were monitored at 1 nm increments. All compounds of interest were ultimately quantified using either 205 nm or 280 nm wavelengths. An initial blank was run to assure that separation conditions would be exactly the same for each actual sample injection. Prepared samples were run in order of the experimental design (Table 4.2), with duplicate and triplicate prepared samples being run subsequently in the same sequential order, for a total of 9 analyses each day.

Table 4. 3: HPLC gradient for prepared cacao liquor extract analysis

Time (min)	**% Channel A**	**% Channel B**
0	100	0
25	70	30
50	50	50
55	0	100
Post (10 min)	100	0

4.2.7 Statistical Analysis of Chemical Data

4.2.7.1 Method:

Principal components analysis (PCA), and mixed-model linear regression were carried out on quantitative chemical data. Complete details of each analysis are included in the results and discussion sections below.

4.2.7.2 Software:

Analysis of the sensory data took place in RStudio v. 1.2.1334, build 1379(f1ac3452) running R version 3.6.0 Patched (2019-06-04 r76666). Packages used include lmerTest (v. 3.1-2: Kuznetsova, Brockhoff, and Christensen, 2017) for mixed-model analysis and ggplot2(v.3.3.2: Wickham, 2016) for PCA biplots and contour plots.

4.3 Results & Discussion

4.3.1 Quantitative Chemical Data

Chemical data collected with HPLC analysis for all roasted liquor samples in the experimental design (Figure 4.1 and Table 4.2 above) can be seen in Table 4.4.

Table 4. 4: Mean concentrations of bitter compounds for all roasting treatments

Origin	Time (min)	Temp (°C)	Theo (mg/g)	Caff (mg/g)	Epi (mg/g)	Cat (mg/g)	ProB2 (mg/g)	cPV (mg/g)	Moist %	Fat %
Ghana	0	24	11.60	1.40	1.87	0.01	0.89	0.00	2.59	52.8
Ghana	11	105	11.88	1.37	1.89	0.03	0.89	0.00	2.09	54.7
Ghana	20	171	12.12	1.53	0.97	0.45	0.35	0.15	1.31	54.7
Ghana	40	114	11.73	1.51	1.40	0.04	0.64	0.07	2.11	56.8
Ghana	40	114	11.49	1.41	1.31	0.06	0.72	0.18	2.13	53.0
Ghana	54	151	12.05	1.47	0.72	0.39	0.25	0.17	0.99	54.3
Ghana	55	64	12.00	1.37	1.77	0.01	0.87	0.00	2.30	54.4
Ghana	80	84	11.75	1.46	1.79	0.01	0.80	0.00	2.22	56.8
Ghana	80	135	12.02	1.44	0.83	0.24	0.38	0.17	1.02	55.4
Mad	0	24	7.94	1.86	1.85	0.08	0.81	0.00	2.67	48.5
Mad	11	105	8.17	1.90	1.86	0.08	0.79	0.00	2.20	49.0
Mad	20	171	8.18	1.92	1.00	0.36	0.27	0.19	1.48	51.8
Mad	40	114	8.34	1.96	1.25	0.07	0.58	0.13	1.96	52.9
Mad	40	114	7.85	1.95	1.29	0.07	0.56	0.13	1.82	51.1
Mad	54	151	8.12	1.92	0.65	0.28	0.11	0.16	1.56	52.5
Mad	55	64	8.13	1.86	1.82	0.07	0.77	0.00	2.56	50.8
Mad	80	84	8.17	1.84	1.76	0.07	0.72	0.00	2.49	51.7
Mad	80	135	8.08	1.92	0.76	0.19	0.22	0.19	1.44	55.5
Peru	0	24	10.91	3.07	3.12	0.09	1.29	0.00	2.53	55.4
Peru	11	105	11.37	3.03	2.95	0.10	1.42	0.00	2.47	57.9
Peru	20	171	11.58	3.10	1.49	0.78	0.57	0.18	1.33	57.1
Peru	40	114	11.18	3.38	2.10	0.15	1.00	0.07	2.59	56.2
Peru	40	114	11.37	3.20	2.21	0.19	1.20	0.15	2.29	54.6
Peru	54	151	11.65	3.24	1.16	0.63	0.35	0.18	1.30	58.3
Peru	55	64	11.49	3.16	3.01	0.11	1.45	0.00	2.57	54.6
Peru	80	84	11.31	3.19	2.86	0.08	1.16	0.00	2.54	54.6
Peru	80	135	11.50	3.34	1.37	0.45	0.56	0.17	1.45	57.4

Theo=Theobromine, Caff=Caffeine, Epi= Epicatechin, Cat=Catechin, ProB2=Procyanidin B2, cPV=cyclo(Pro-Val), Moist=Moisture, Fat=Crude Fat

4.3.2 Principal Components Analysis

Principal components analysis (PCA) of the chemical data was performed (i.e., bitter compounds, moisture, and crude fat) for all origins. The PCA biplot (Figure 4.2) shows substantial separation between the three origins, though a slight overlap does occur between Ghana and Madagascar, with roasting treatment 13 belonging to Ghana, and roasting treatment 26 belonging to Madagascar. However, no overlap occurs between Peru and either of the other two origins, suggesting that the difference between the Peruvian cacao and the other two origins, based upon the variation of all analyzed compounds, is significant. This is not unexpected, as the Peruvian cacao was sourced from Marañon farm, where the cacao had previously been genotyped by the USDA, and was found to be significantly different from varieties such as Amelonado (Fang et al., 2014), which is often found in West Africa (Edwin & Masters, 2005; Motamayor et al., 2008; Sukha, Umaharan, & Butler, 2017). This finding aligns well with earlier work (Motamayor et al., 2008) suggesting that Marañon is in a separate genetic cluster from all other known clusters of *Theobroma cacao* (Motamayor et al., 2008). It can also be seen in the biplot that procyanidin B2 and epicatechin are highly correlated, even across chocolate from multiple origins, with the likelihood of significant genetic variation, and with a variety of roast profiles, a detail which supports previous research stating as much (Cooper et al., 2007). Additionally, both catechin and cyclo(Pro-Val) increases are relatively highly correlated with increases in time and temperature, a characteristic that was investigated further in the linear regression section below (section 4.3.3).

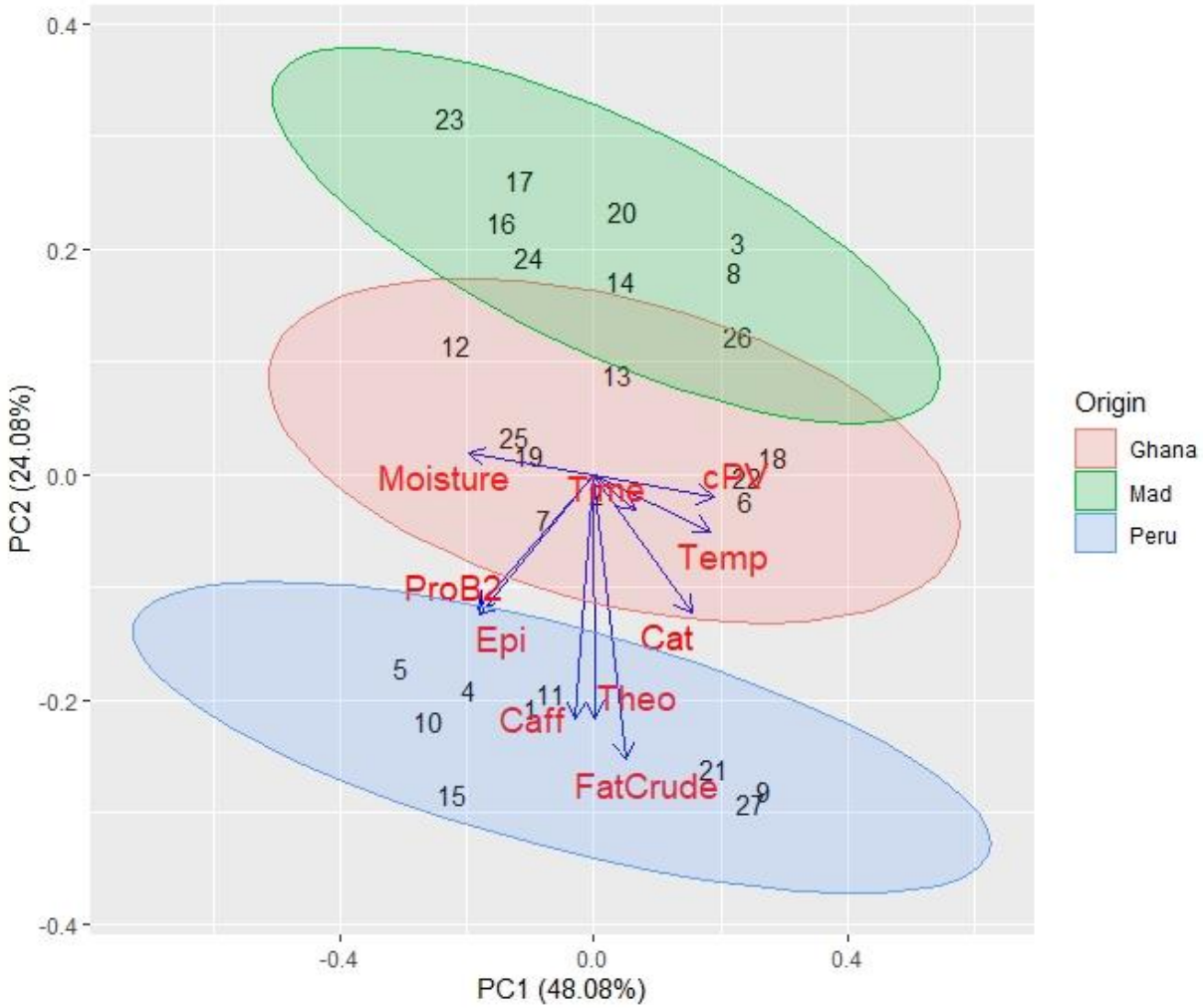

Figure 4. 2: PCA biplot showing 72.16% of the variance in the data between the three origins across all roasting treatments for the factors shown.

ProB2=Procyanidin B2, Epi= Epicatechin, Caff=Caffeine, Theo=Theobromine, Cat=Catechin, Temp=Temperature, cPV=cyclo(Pro-Val)

4.3.3 Linear Regression

First, data was mean-centered and scaled given that predictors (i.e., time and temperature) and response variables (i.e., chemical concentrations) are on quite different scales. Next, backward/forward stepwise selection was used to choose the model with the best (smallest) Bayesian Information Criterion (BIC) value starting with

the full potential model (i.e., all relevant first order and second-order effects and their interactions):

Single Bitter Compound ~ Origin + Time + Temperature + Time:Temperature + $Time^2$ + $Temperature^2$ + Time: $Temperature^2$ + Temperature: $Time^2$ + Origin:Time + Origin:Temperature + Origin:Time:Temperature + Origin:$Time^2$ + Origin:$Temperature^2$ + Origin:Time: $Temperature^2$ + Origin:Temperature: $Time^2$

If diagnostic plots or variance inflation factor (VIF) analysis suggested that the model was less than optimal, which generally was not the case, various transformations of the response were used to search for a better model, as well as investigation of the effect of adding or removing nominally insignificant terms. These details will be shared for each individual compound below when relevant. Each selected model was then fit, and estimates were obtained of p-values and coefficients for each factor in the model. After model selection was finalized, contour plots were prepared for each model to help visualize the predicted values of the selected models. All details can be seen below for each compound starting with section 4.3.3.1.

4.3.3.1 Catechin:

The BIC-selected model for best predicting Catechin is:

Sqrt(Catechin) ~ Origin + Time + Temperature + $Temperature^2$ + Origin:Temperature

Diagnostic plots showed some irregularities, but a square-root transformation of the response dramatically corrected these issues. Diagnostic plots of the ultimately selected

model show no apparent pattern in the residuals vs fitted values, an approximately normal distribution in the Normal Q-Q plot, relative homoscedasticity, and no high-leverage points.

The summary of calculated model statistics can be seen in tables 4.5 and 4.6

Table 4. 5: Summary of calculated model statistics, including effect estimates

```
Coefficients:

             Estimate Std. Error t value           Pr(>|t|)
(Intercept)   0.28755    0.01211   23.74 0.0000000000000014 ***
Origin1      -0.07870    0.01200   -6.56 0.0000028107528078 ***
Origin2      -0.02760    0.01200   -2.30             0.0330 *
Time          0.03941    0.01031    3.82             0.0012 **
Temp          0.19057    0.00922   20.66 0.0000000000000177 ***
I(Temp^2)     0.09906    0.00897   11.04 0.0000000010486044 ***
Origin1:Temp  0.03139    0.01223    2.57             0.0189 *
Origin2:Temp -0.06027    0.01223   -4.93 0.0000934724630154 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0441 on 19 degrees of freedom
Multiple R-squared:  0.968,    Adjusted R-squared:  0.956
F-statistic: 81.6 on 7 and 19 DF,  p-value: 0.000000000000789
```

Table 4. 6: ANOVA Table (Type III tests)

```
Response: sqrt(Catechin)

            Sum Sq Df F value              Pr(>F)
(Intercept)  1.096  1   563.5 0.0000000000000014 ***
Origin       0.164  2    42.2 0.0000001017322305 ***
Time         0.028  1    14.6              0.0012 **
Temp         0.830  1   426.8 0.0000000000000177 ***
I(Temp^2)    0.237  1   121.8 0.0000000010486044 ***
Origin:Temp  0.047  2    12.2              0.0004 ***
Residuals    0.037 19

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the adjusted R-squared is 0.956, suggesting that the predictors describe approximately 95.6% of the variance in the square root of catechin, and the p-value for the model (i.e., p-value= 0.000000000000789) is statistically highly significant (i.e., p-value ≤0.001). Additionally, all model terms (Table 4.6) are statistically highly significant (i.e., p-value ≤0.001), with the exception of Time, which is significant (i.e., p-value≤0.05). VIF values for all terms were below 2.5.

What can be seen in Table 4.5 above, and in Figures 4.3, 4.4, and 4.5 below, is that in this roasting experiment, temperature is the most important single predictor of catechin, with unroasted cacao having the lowest concentration of catechin and the roasting treatment with the highest temperature (i.e., 171 °C) having the highest concentration of catechin. Still, the picture is slightly more complex, as time does also play a role in catechin increase. It can be said that as temperature and time increase across our experimental region, so does generally the concentration of catechin (Table 4.5 and Figures 4.3, 4.4, 4.5), and there is some quadratic curvature in the predicted response dependent upon temperature, most relevant in the experimental region at lower temperatures (Figures 4.3, 4.4, 4.5). But the estimated effect of temperature is still much larger than any other single effect. In fact, effect estimates (Table 4.5) suggest that an increase in temperature of one standard deviation (i.e., 43 °C degrees) yields 4.8 times the increase in the square root of catechin concentration (i.e., 0.19057/0.03941) as does a one standard deviation increase in roasting time (i.e., approximately 27 minutes), and given the much greater statistical significance of temperature (i.e., p-

value =0.0000000000000177) as opposed to time (i.e., p-value=0.0012), the importance of temperature as a predictor of catechin concentration for these particular origins and given this specific experimental region, is reinforced. This overall trend is in agreement with previous work that showed increases of catechin during roasting for some origins as high as 6.5-fold (Payne et al., 2010).

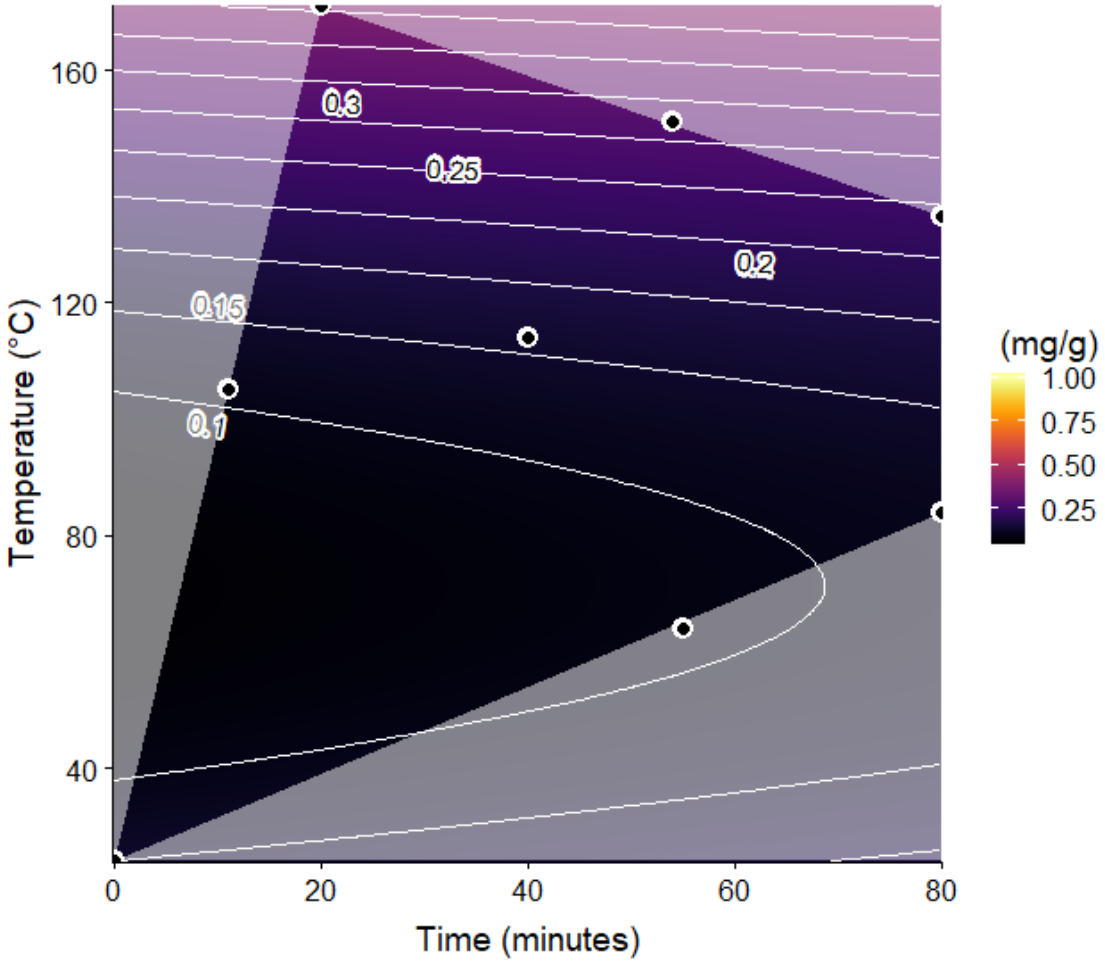

Figure 4. 3: MADAGASCAR: Contour plot for predicted Catechin concentration for roasted Madagascan chocolate liquor treatments across the experimental region.

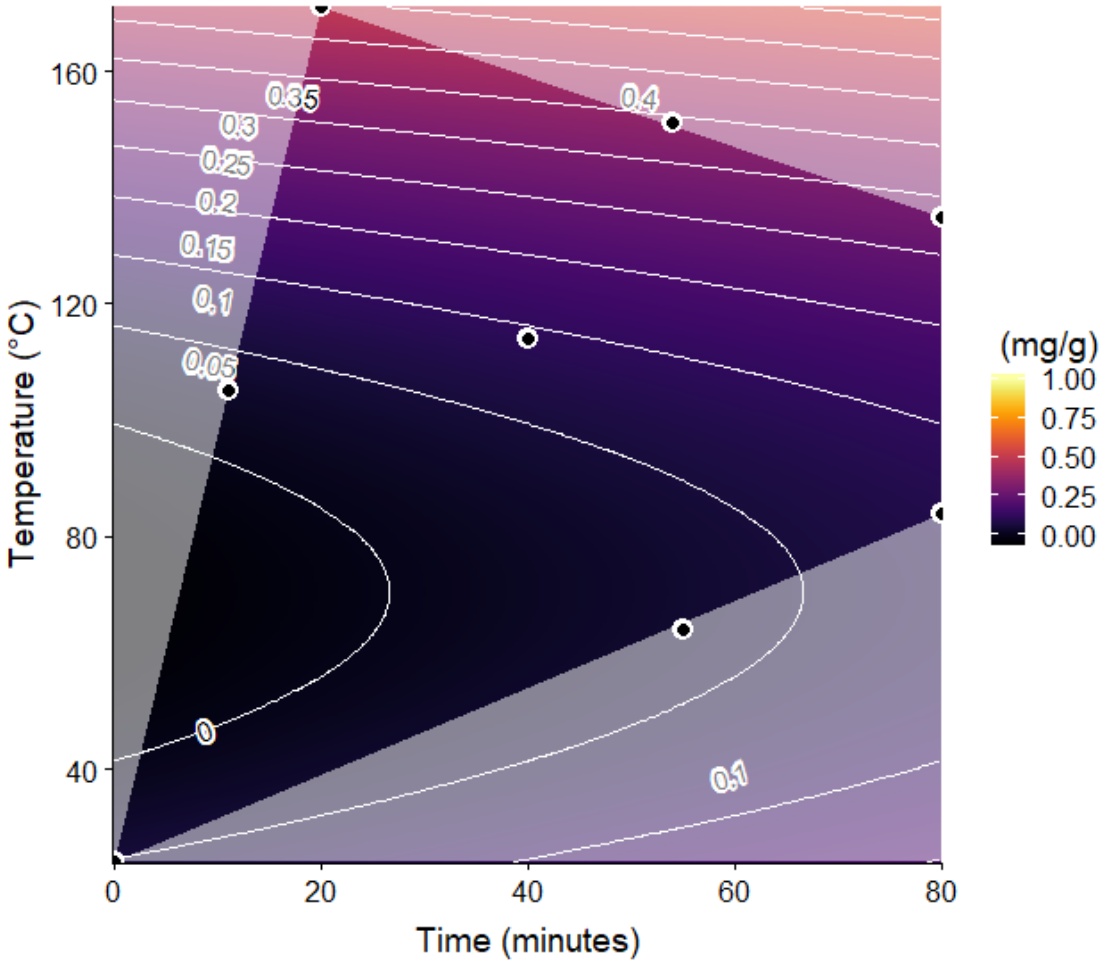

Figure 4. 4: GHANA: Contour plot for predicted Catechin concentration for roasted chocolate liquor treatments across the experimental region.

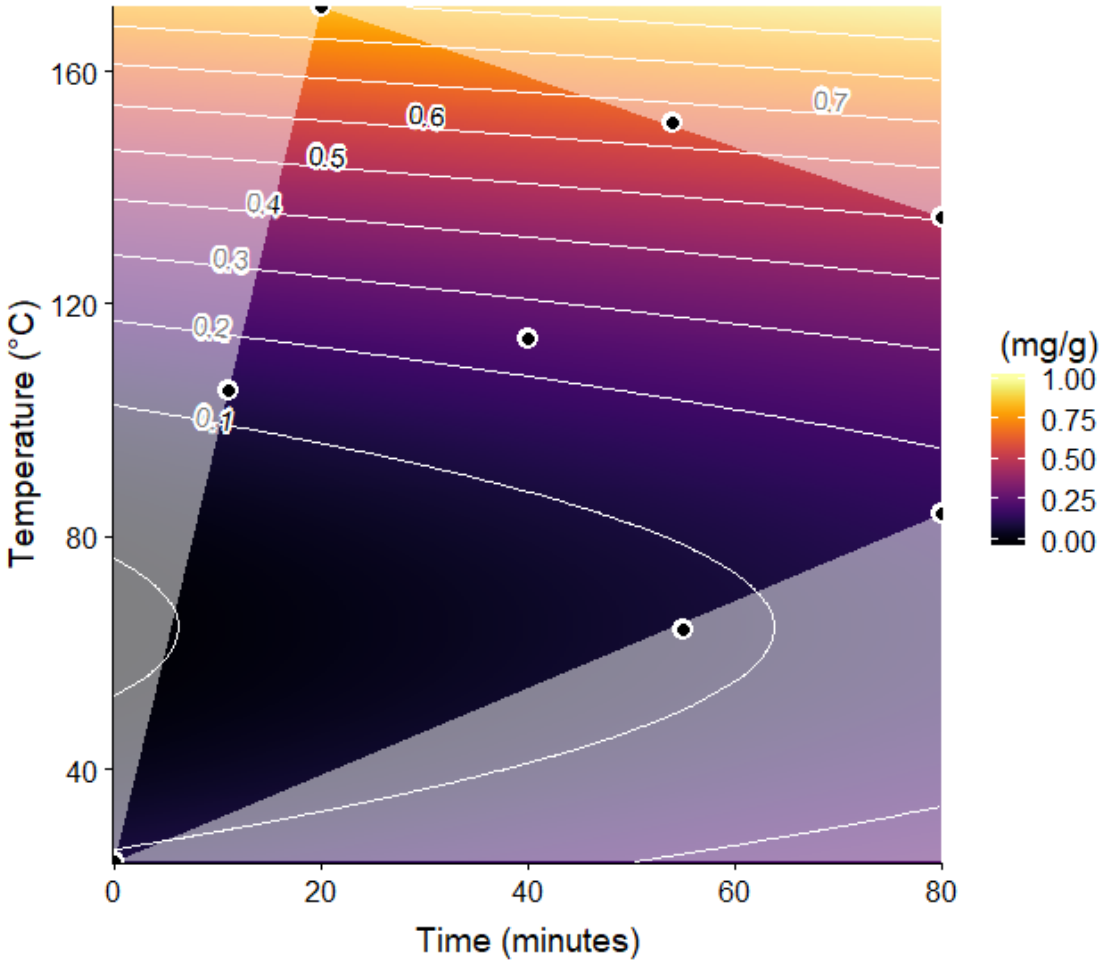

Figure 4. 5: PERU: Contour plot for predicted Catechin concentration for roasted Peruvian chocolate liquor treatments across the experimental region.

As for origin-based differences, it is interesting to note that though a similar trend exists with increasing catechin concentration as roast temperature increases for all three origins, Peru has, by far, the highest concentration of catechin at the roasting treatment with the highest temperature (i.e., 171 °C) (Table 4.4 and Figure 4.5), which is approximately double or more that of either Ghana or Madagascar (Table 4.4 and Figures 4.3 and 4.4).

4.3.3.2 Epicatechin

The BIC-selected model for best predicting Epicatechin was:

$$\text{Epicatechin} \sim \text{Origin} + \text{Time} + \text{Temperature} + \text{Time:Temp} + \text{Temperature}^2 + \text{Time}^2 + \text{Origin:Time} + \text{Origin:Temp} + \text{Origin:Time:Temp} + \text{Origin:Temperature}^2$$

However, diagnostic plots showed a clear and curved pattern in the plot of standardized residuals versus fitted values, and four very high-leverage points. It was determined that by removing the least significant term, the Origin:Temperature^2 interaction (p-value = 0.259), these issues were largely resolved, with the curvature in the residual plot almost entirely disappearing, the number of high-leverage points being reduced to one, and the distribution of residuals approximating a normal distribution. Additionally, the single remaining high-leverage point was carefully examined, and the value was found to be reasonable given the treatment conditions. Further, removing the interaction term maintained essentially the same BIC value (-101.4 instead of -101.5), and only slightly increased the residual sums of squares (RSS) for the model from 0.09 to 0.11. Therefore, the final selected model is:

$$\text{Epicatechin} \sim \text{Origin} + \text{Time} + \text{Temperature} + \text{Time:Temp} + \text{Temperature}^2 + \text{Time}^2 + \text{Origin:Time} + \text{Origin:Temp} + \text{Origin:Time:Temp}$$

The summary of calculated model statistics can be seen in tables 4.7 and 4.8.

Table 4. 7: Summary of calculated model statistics, including effect estimates

```
Coefficients:

                  Estimate Std. Error t value          Pr(>|t|)
(Intercept)        0.04613    0.03149    1.46           0.16677
Origin1           -0.40150    0.02622  -15.31 0.000000001070052 ***
Origin2           -0.44458    0.02622  -16.96 0.000000000301691 ***
Time              -0.34084    0.02304  -14.79 0.000000001635885 ***
Temp              -0.81309    0.02273  -35.77 0.000000000000023 ***
I(Temp^2)         -0.19083    0.02146   -8.89 0.000000689163469 ***
I(Time^2)          0.22519    0.02593    8.68 0.000000900837928 ***
Time:Temp         -0.36909    0.02272  -16.25 0.000000000512899 ***
Origin1:Time       0.02911    0.02773    1.05           0.31302
Origin2:Time       0.00962    0.02773    0.35           0.73421
Origin1:Temp       0.17007    0.03138    5.42           0.00012 ***
Origin2:Temp       0.13582    0.03138    4.33           0.00082 ***
Origin1:Time:Temp  0.06154    0.02808    2.19           0.04722 *
Origin2:Time:Temp  0.02895    0.02808    1.03           0.32139

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0938 on 13 degrees of freedom
Multiple R-squared:  0.996,    Adjusted R-squared:  0.991
F-statistic:  227 on 13 and 13 DF,  p-value: 0.000000000000416
```

Table 4. 8: ANOVA Table (Type III tests)

```
Response: Epicatechin

                Sum Sq Df F value            Pr(>F)
(Intercept)       0.02  1    2.15             0.167
Origin            9.16  2  521.08 0.000000000000388 ***
Time              1.92  1  218.89 0.000000001635885 ***
Temp             11.24  1 1279.14 0.000000000000023 ***
I(Temp^2)         0.70  1   79.09 0.000000689163469 ***
I(Time^2)         0.66  1   75.41 0.000000900837928 ***
Time:Temp         2.32  1  264.01 0.000000000512899 ***
Origin:Time       0.02  2    1.06             0.375
Origin:Temp       0.84  2   47.71 0.000001028572085 ***
Origin:Time:Temp  0.10  2    5.42             0.019 *
Residuals         0.11 13

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the adjusted R-squared is 0.991, suggesting that the predictors describe approximately 99.1% of the variance in the response, and the p-value for the model (i.e., p-value= p-value: 0.000000000000416) is statistically highly significant (i.e., p-value≤0.001). Additionally, all model terms (Table 4.8) are statistically highly significant (i.e., p-value ≤0.001), with the exception of Origin:Time:Temperature which is significant (p-value≤0.05), and Origin:Time which is not statistically significant, though it is important for this term to remain in the model due to heredity. VIF values for all main effects were below 2.5, and below 2.7 for all terms.

What can be seen in Table 4.8 above, and Figures 4.6, 4.7, and 4.8 below, is that in this experiment, though temperature is the most important single predictor of epicatechin, time and also time:temperature interactions are quite important. For this reason, though raw (i.e., 0 minutes at 24°C) and very lightly roasted (i.e., 11 minutes at 105°C) cacao have the highest concentration of epicatechin, it is not the highest temperature roast (i.e., 171 °C for 20 minutes) that resulted in the lowest epicatechin concentrations. Rather, the slightly longer roasts with slightly lower temperatures (i.e., 54 minutes at 151 °C) had the lowest epicatechin concentrations respectively for each origin. Still, it can be said that as temperature and time both increase across our experimental region, so does generally the concentration of epicatechin decrease (Table 4.8 and Figures 4.6, 4.7, 4.8), and there is some curvature in the predicted response dependent upon squared terms of both temperature and time (Figures 4.6, 4.7, 4.8). But the estimated main effect of temperature is still considerably larger than any other

single effect. In fact, effect estimates (Table 4.8) suggest that an increase in temperature of one standard deviation (i.e., 43 °C degrees) yields 2.4 times the decrease in epicatechin concentration (i.e., -0.81309/-0.34084) as does a one standard deviation increase in roasting time (i.e., approximately 27 minutes). This pattern of general epicatechin decrease with increased roast temperatures is in agreement with previous work (Payne et al., 2010).

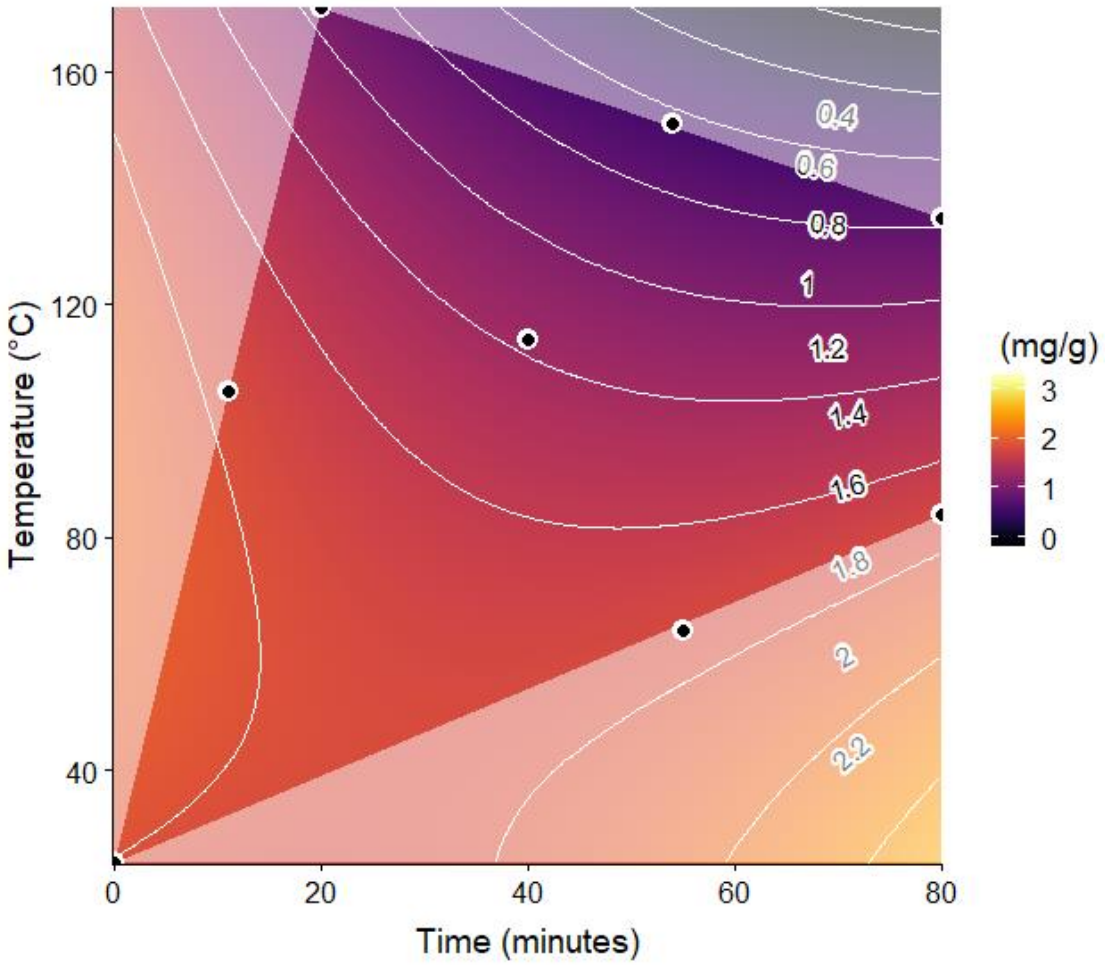

Figure 4. 6: MADAGASCAR: Contour plot for predicted Epicatechin concentration for roasted Madagascan chocolate liquor treatments across the experimental region.

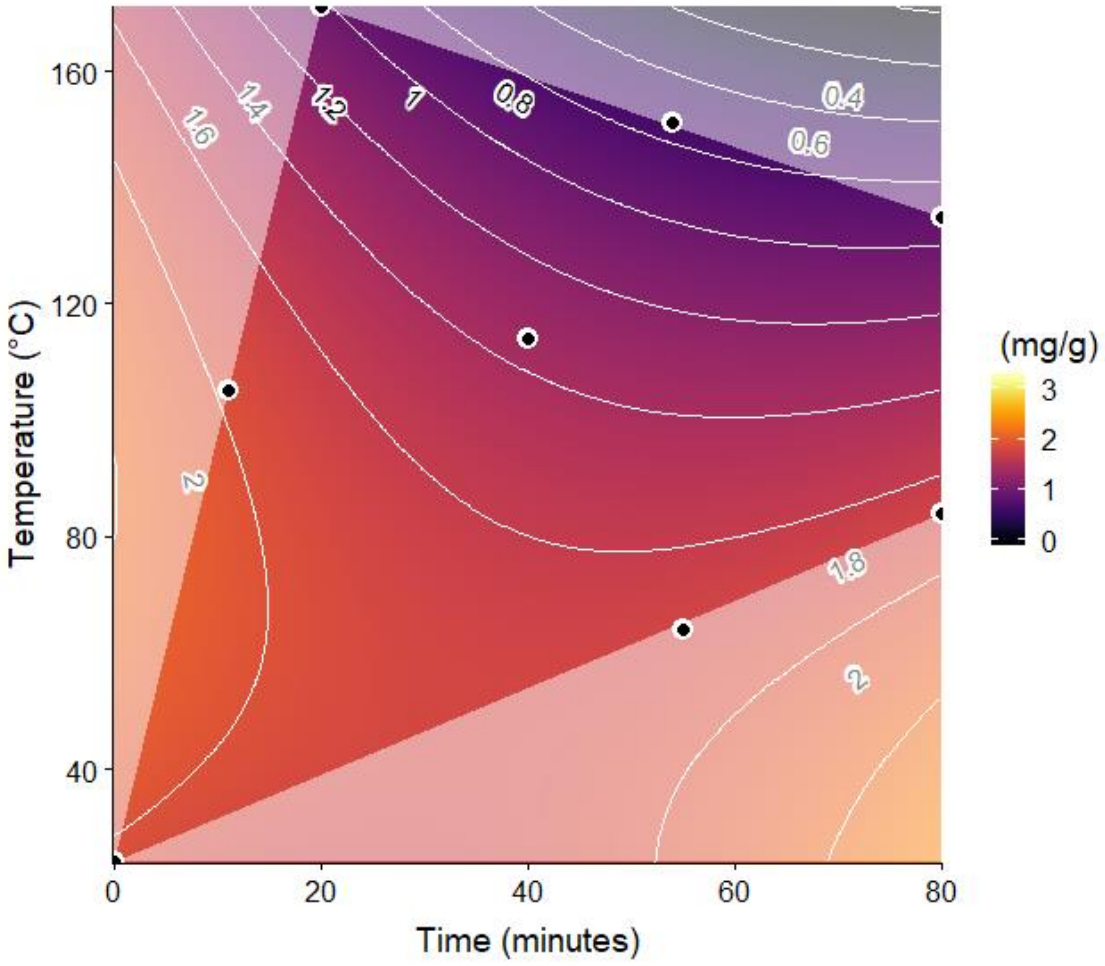

Figure 4. 7: GHANA: Contour plot for predicted Epicatechin concentration for roasted Ghanaian chocolate liquor treatments across the experimental region.

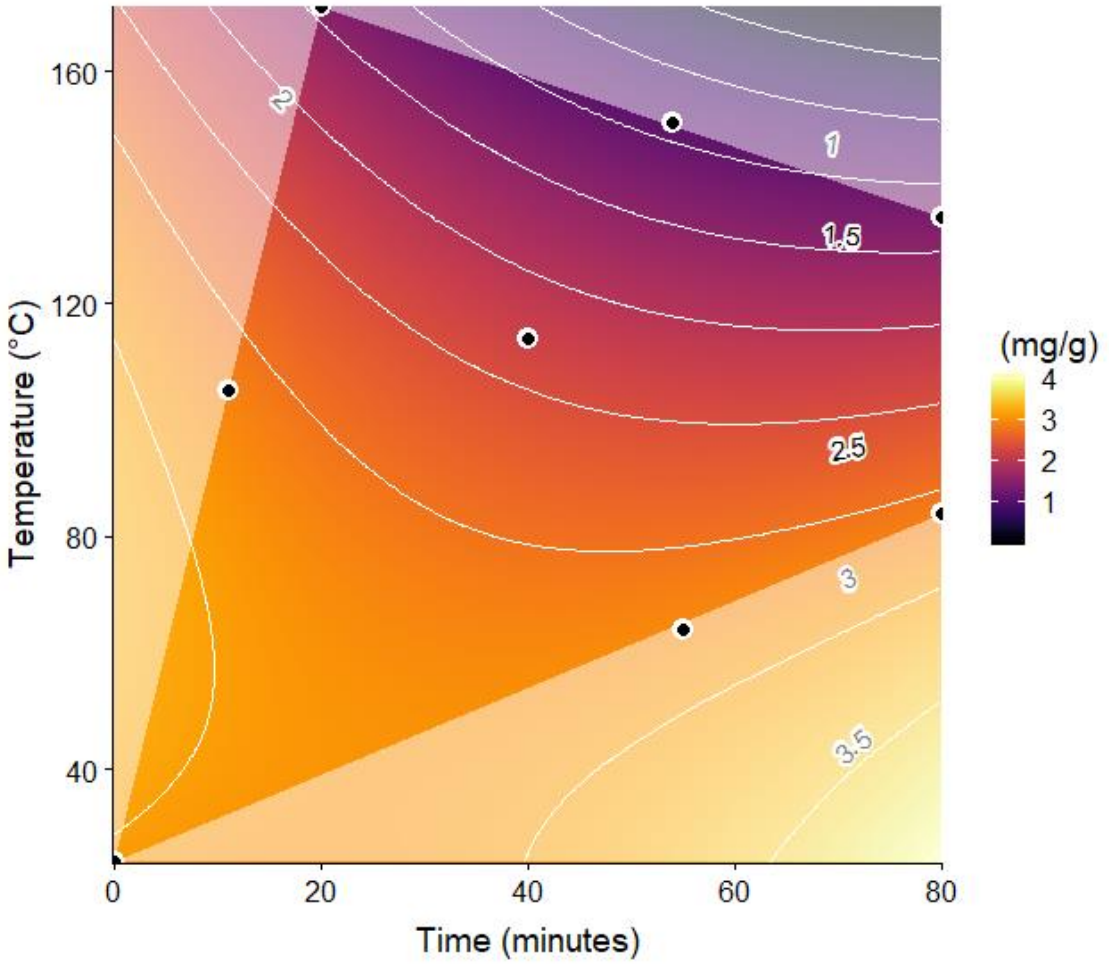

Figure 4. 8: PERU: Contour plot for predicted Epicatechin concentration for roasted Peruvian chocolate liquor treatments across the experimental region.

As for origin-based differences, it is interesting to note that though a similar trend exists with decreasing epicatechin concentration as roast temperature and time increase across all three origins, Peru has the highest raw and roasted concentrations of epicatechin (Table 4.4 and Figures 4.6, 4.7, and 4.8) across the entire experimental region, which is approximately 1.6 to 1.7 times as much as either Madagascar or Ghana respectively (Table 4.4 and Figures 4.6, 4.7, and 4.8).

4.3.3.3 Procyanidin B2

The BIC-selected model for best predicting Procyanidin B2 is:

Procyanidin B2 ~ Origin + Time + Temp + Time:Temp + Temperature^2 + Time^2 + Temperature:Time^2 + Origin:Time + Origin:Temp + Origin:Time^2 + Origin:Temperature:Time^2

However, there were some concerns with two of the diagnostic plots, as the Normal Q-Q plot (i.e., standardized residuals by theoretical quantiles) did not show a primarily linear trend, meaning the residual distribution of the model was not approximating a normal distribution. Additionally, a curved pattern arguably existed in the residuals in the Scale Location plot (i.e., standardized residuals by fitted values). VIF values were also higher than expected. While investigating these potential issues with manual single-term additions and deletions, as well as response transformations, it was ultimately determined that the model previously selected for Epicatechin, which has an R=0.97 correlation with Procyanidin B2, produced a predictive model that greatly improved all diagnostic plots and showed VIF values for all terms below 2.7. Additionally, though the BIC for this smaller model was slightly worse (i.e., an increase from 12.1 to 19.9), ANOVA showed a difference between the models that was not quite significant (p-value=0.078). Therefore, the final selected model became:

Procyanidin B2 ~ Origin + Time + Temp + Time:Temp + Temperature^2 + Time^2 + Origin:Time + Origin:Temp + Origin:Time:Temperature

The summary of calculated model statistics can be seen in tables 4.9 and 4.10.

Table 4. 9: Summary of calculated model statistics, including effect estimates

```
Coefficients:

                  Estimate Std. Error t value      Pr(>|t|)
(Intercept)         0.2450     0.0677    3.62       0.00312 **
Origin1            -0.2511     0.0564   -4.45       0.00065 ***
Origin2            -0.5273     0.0564   -9.35 0.00000038899 ***
Time               -0.3857     0.0495   -7.79 0.00000300797 ***
Temp               -0.8545     0.0489  -17.48 0.00000000021 ***
I(Temp^2)          -0.3056     0.0461   -6.62 0.00001652888 ***
I(Time^2)           0.1183     0.0558    2.12       0.05358 .
Time:Temp          -0.3013     0.0488   -6.17 0.00003382656 ***
Origin1:Time        0.0628     0.0596    1.05       0.31149
Origin2:Time        0.0176     0.0596    0.30       0.77265
Origin1:Temp        0.1577     0.0675    2.34       0.03608 *
Origin2:Temp        0.0981     0.0675    1.45       0.16971
Origin1:Time:Temp   0.1019     0.0604    1.69       0.11542
Origin2:Time:Temp   0.0516     0.0604    0.85       0.40847

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.202 on 13 degrees of freedom
Multiple R-squared:  0.98,    Adjusted R-squared:  0.959
F-statistic: 48.2 on 13 and 13 DF,  p-value: 0.0000000081
```

Table 4. 10: ANOVA Table (Type III tests)

```
Response: Procyanidin B2

                  Sum Sq Df F value        Pr(>F)
(Intercept)         0.53  1   13.09        0.0031 **
Origin              8.07  2   99.31 0.00000001332 ***
Time                2.46  1   60.61 0.00000300797 ***
Temp               12.42  1  305.53 0.00000000021 ***
I(Temp^2)           1.78  1   43.88 0.00001652888 ***
I(Time^2)           0.18  1    4.50        0.0536 .
Time:Temp           1.55  1   38.06 0.00003382656 ***
Origin:Time         0.08  2    1.00        0.3930
Origin:Temp         0.59  2    7.32        0.0074 **
Origin:Time:Temp    0.27  2    3.34        0.0673 .
Residuals           0.53 13

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the adjusted R-squared is 0.959, suggesting that the predictors describe approximately 95.9% of the variance in the response, and the p-value for the model

(i.e., p-value= 0.0000000081) is statistically highly significant (i.e., p-value≤0.001). Additionally, the following model terms (Table 4.10) are statistically highly significant (i.e., p-value≤0.001): Origin, Time, Temperature, Temperature^2 and Time:Temperature. Also, Origin:Temperature is statistically significant (i.e., p-value ≤0.05), and Origin:Temperature:Time and Time^2 are close to statistical significance (i.e., 0.05 < p-value ≤0.1). Finally, Origin:Time is not at all statistically significant, but remains due to reasons of heredity given the presence in the model of Origin:Temperature:Time, which results in a model with a better (lower) BIC than if it were removed.

What can be seen in Table 4.9 above, and Figures 4.9, 4.10, and 4.11 below, is that in this experiment, though temperature is the most important single predictor of procyanidin B2, time and also time:temperature interactions are quite important as well. For this reason, though raw (i.e., 0 minutes at 24°C) and very lightly roasted (i.e., 11 minutes at 105°C, or 55 minutes at 64°C) cacao have the highest concentrations of procyanidin B2, it is not the highest temperature roast (i.e., 171 °C for 20 minutes) that resulted in the lowest procyanidin B2 concentration. Rather, the slightly longer roasts with slightly lower temperatures (i.e., 54 minutes at 151 °C) had the lowest procyanidin B2 concentrations for each respective origin. Still, it can be said that as temperature and time increase across our experimental region, so does generally the concentration of procyanidin B2 decrease (Table 4.4 and Figures 4.9, 4.10, 4.11), though there is some curvature in the predicted response dependent upon squared terms of both temperature and time (Figures 4.9, 4.10, 4.11), especially at lower end of the

temperature and time ranges. But the estimated main effect of temperature is still considerably larger than any other single effect. In fact, effect estimates (Table 4.9) suggest that an increase in temperature of one standard deviation (i.e., 43 °C degrees) yields 2.2 times the decrease in procyanidin B2 concentration (i.e., -0.8545/-0.3857) as does a one standard deviation increase in roasting time (i.e., approximately 27 minutes). This general behavior, and the magnitude of its change correlates almost exactly to that of epicatechin R=0.97), though starting concentration values are very different. This makes sense given that previous work has shown that these two compounds are highly correlated, even across chocolate from multiple origins with various roast profiles (Cooper et al., 2007).

As for origin-based differences, it is interesting to note that though a similar trend exists with decreasing procyanidin B2 concentration as roast temperatures and times increase in the experimental region for all three origins, Peru has the highest raw and roasted concentrations of procyanidin B2 (Table 4.4 and Figures 4.9, 4.10, and 4.11), averaging 1.6 to 1.9 times as much as either Ghana or Madagascar respectively across the experimental region (Table 4.4 and Figures 4.9, 4.10, and 4.11).

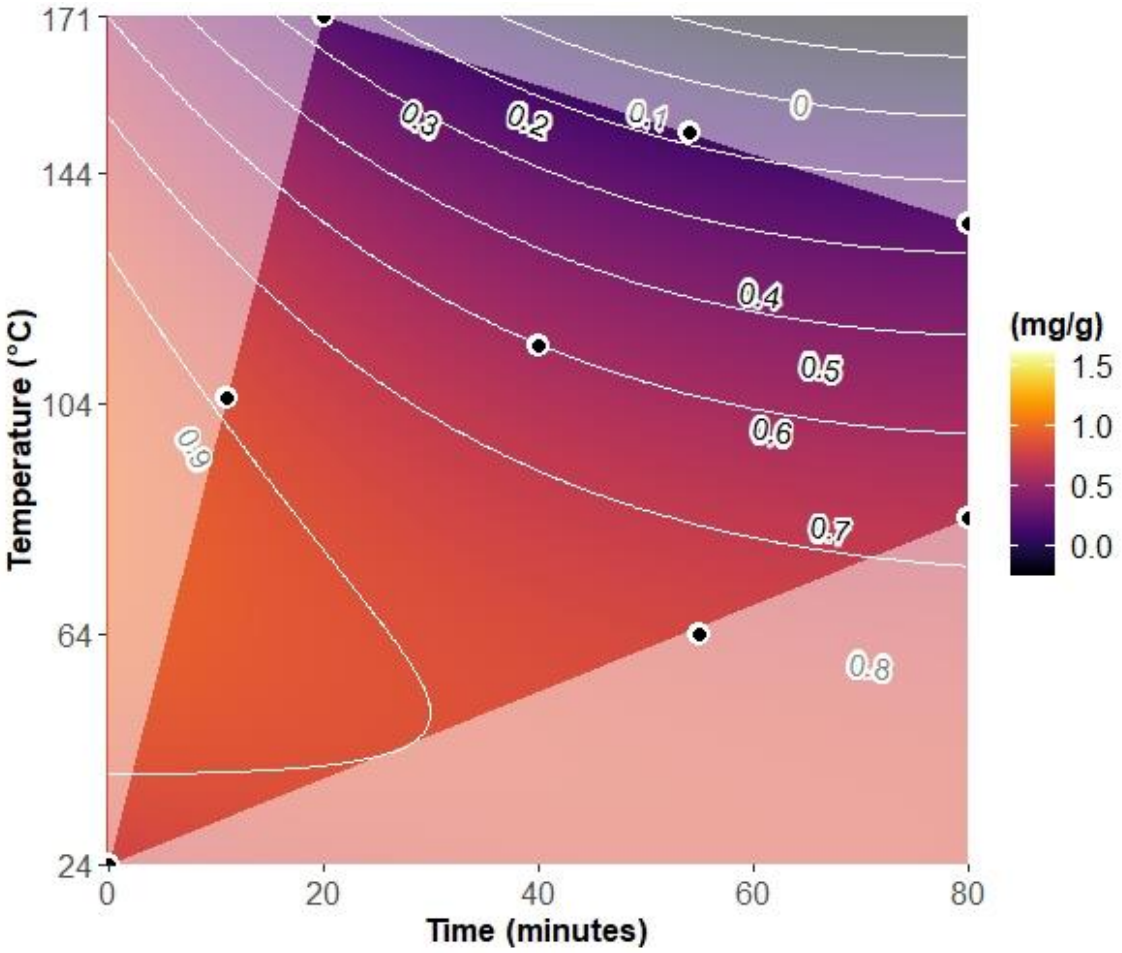

Figure 4. 9: MADAGASCAR: Contour plot for predicted Procyanidin B2 concentration for roasted Madagascan chocolate liquor treatments across the experimental region.

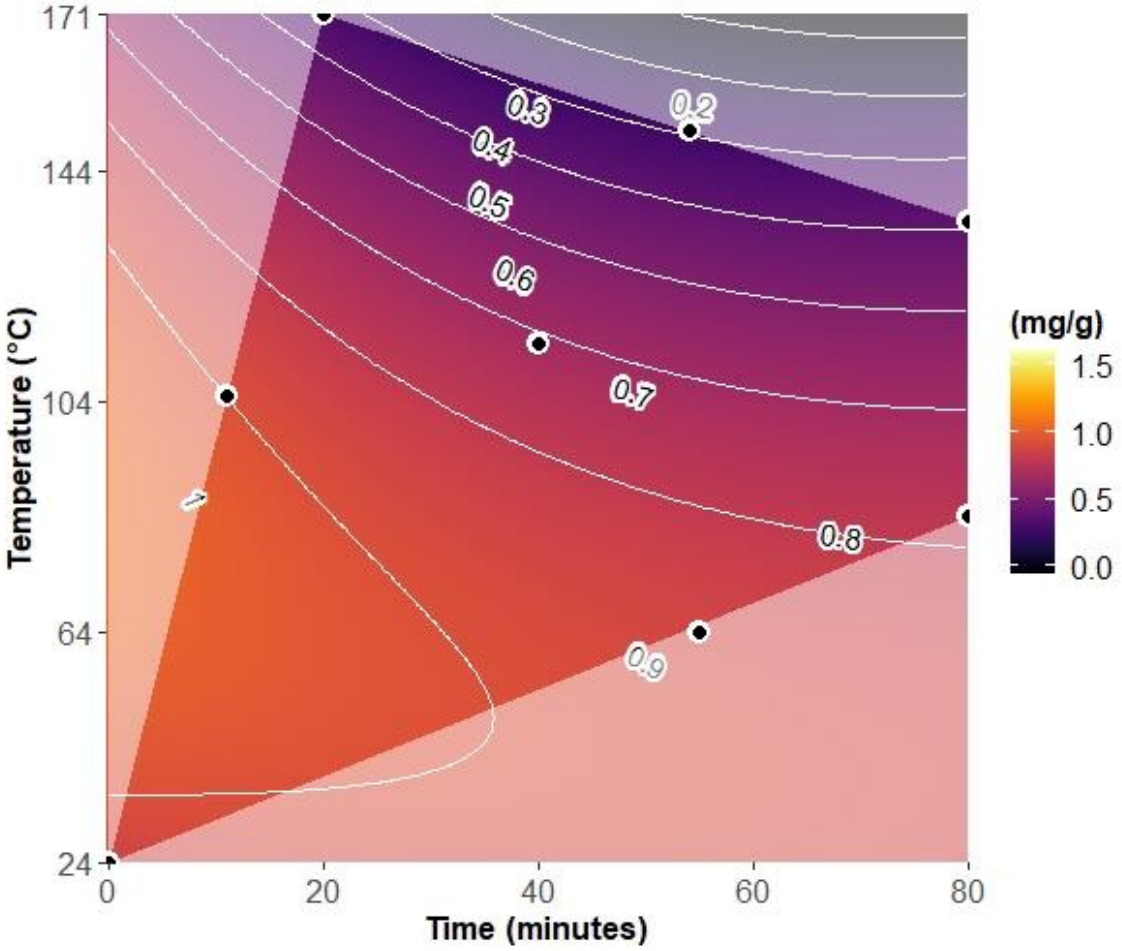

Figure 4. 10: GHANA: Contour plot for predicted Procyanidin B2 concentration for roasted chocolate liquor treatments across the experimental region.

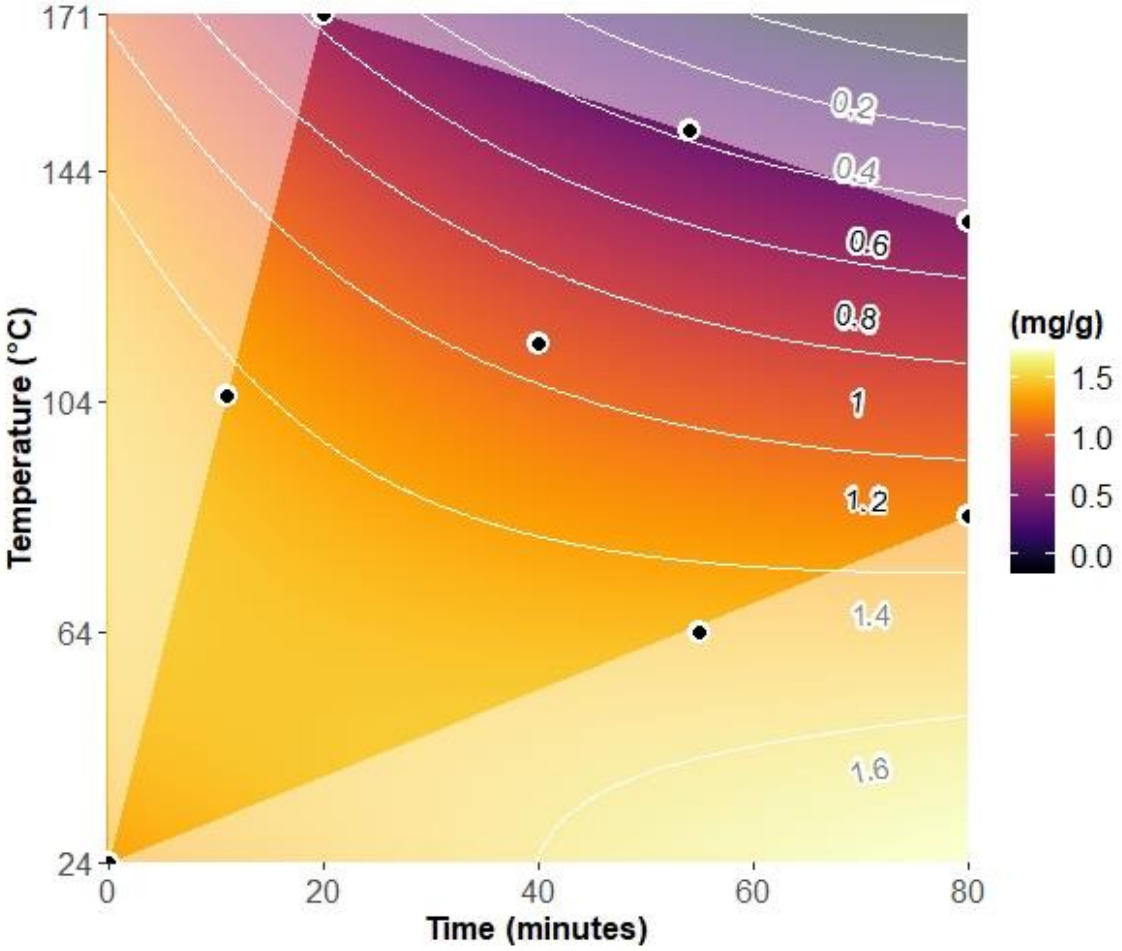

Figure 4. 11: PERU: Contour plot for predicted Procyanidin B2 concentration for roasted Peruvian chocolate liquor treatments across the experimental region.

4.3.3.4 Cyclo(Proline-Valine)

The BIC-selected model for best predicting cyclo(Proline-Valine) is:

cyclo(Proline-Valine) ~ Time + Temp + Time:Temp + Temperature^2 + Time^2

Despite the relatively small number of observations available, imperfect estimates of at least 3 roast treatments for each origin (i.e., set to 0 mg/mL) due to small concentrations of cyclo(Pro-Val) close to the limit of detection, combined with

unavoidable chromatographic interferences, diagnostic plots appeared quite reasonable. Normal Q-Q plot showed approximately normal residual distribution, no high-leverage points were found, and only a slight curvature was apparent in the residuals vs fitted plot only because of three observations. However, no response transformation or adjustment of the model corrected this issue. Therefore, the stepwise-selected model was maintained. VIF values were all below 2.5.

The summary of calculated model statistics can be seen in tables 4.11 and 4.12.

Table 4. 11: Summary of calculated model statistics, including effect estimates

```
Coefficients:

            Estimate Std. Error t value      Pr(>|t|)
(Intercept)   0.0286     0.1296    0.22       0.82754
Time          0.2912     0.0948    3.07       0.00578 **
Temp          0.9954     0.0935   10.64 0.00000000065 ***
I(Temp^2)     0.1927     0.0883    2.18       0.04055 *
I(Time^2)    -0.3209     0.1067   -3.01       0.00670 **
Time:Temp     0.4420     0.0935    4.73       0.00011 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.386 on 21 degrees of freedom
Multiple R-squared:  0.88,    Adjusted R-squared:  0.851
F-statistic: 30.7 on 5 and 21 DF,  p-value: 0.00000000556
```

Table 4. 12: ANOVA Table (Type III tests)

```
Response: cyclo(Pro-Val)

            Sum Sq Df F value        Pr(>F)
(Intercept)   0.01  1    0.05       0.82754
Time          1.40  1    9.44       0.00578 **
Temp         16.85  1  113.23 0.00000000065 ***
I(Temp^2)     0.71  1    4.76       0.04055 *
I(Time^2)     1.35  1    9.05       0.00670 **
Time:Temp     3.33  1   22.36       0.00011 ***
Residuals     3.13 21

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the adjusted R-squared is 0.851, suggesting that the predictors describe approximately 85.1% of the variance in the response, and the p-value for the model (i.e., p-value= p-value: 0.00000000556) is statistically highly significant (i.e., p-value ≤0.001). Additionally, the following model terms are statistically highly significant (i.e., p-value ≤0.001): Temperature and Time:Temperature, while Time, Temperature^2, and Time^2 are all statistically significant (i.e., p-value ≤0.05). No terms in the selected model are insignificant.

What can be seen in Table 4.11 above, and Figure 4.12 below, is that in this experiment, temperature is the most important single predictor of cyclo(Pro-Val). Still, it can be said that as temperature and time increase across our experimental region, so does generally the concentration of cyclo(Pro-Val) (Table 4.4 and Figure 4.12), and there is some curvature in the predicted response dependent upon squared terms of both temperature and time (Figures 4.11). But the estimated main effect of temperature is still considerably larger than any other single effect. In fact, effect estimates (Table 4.11) suggest that an increase in temperature of one standard deviation (i.e., 43 °C degrees) yields 3.4 times the increase in cyclo(Pro-Val) concentration (i.e., 0.9954/0.2912) as does a one standard deviation increase in roasting time (i.e., approximately 27 minutes). These findings are reasonable given that roasting has previously been found to result in the creation of other diketopiperazines (DKPs) from peptides (Rizzi, 1989; Ziegleder, 2017), and darker roasts, particularly at higher temperatures, appear to increase DKP

levels the most (Bonvehí et al., 2000), whereas unroasted cocoa contains virtually no DKPs (Bonvehí et al., 2000).

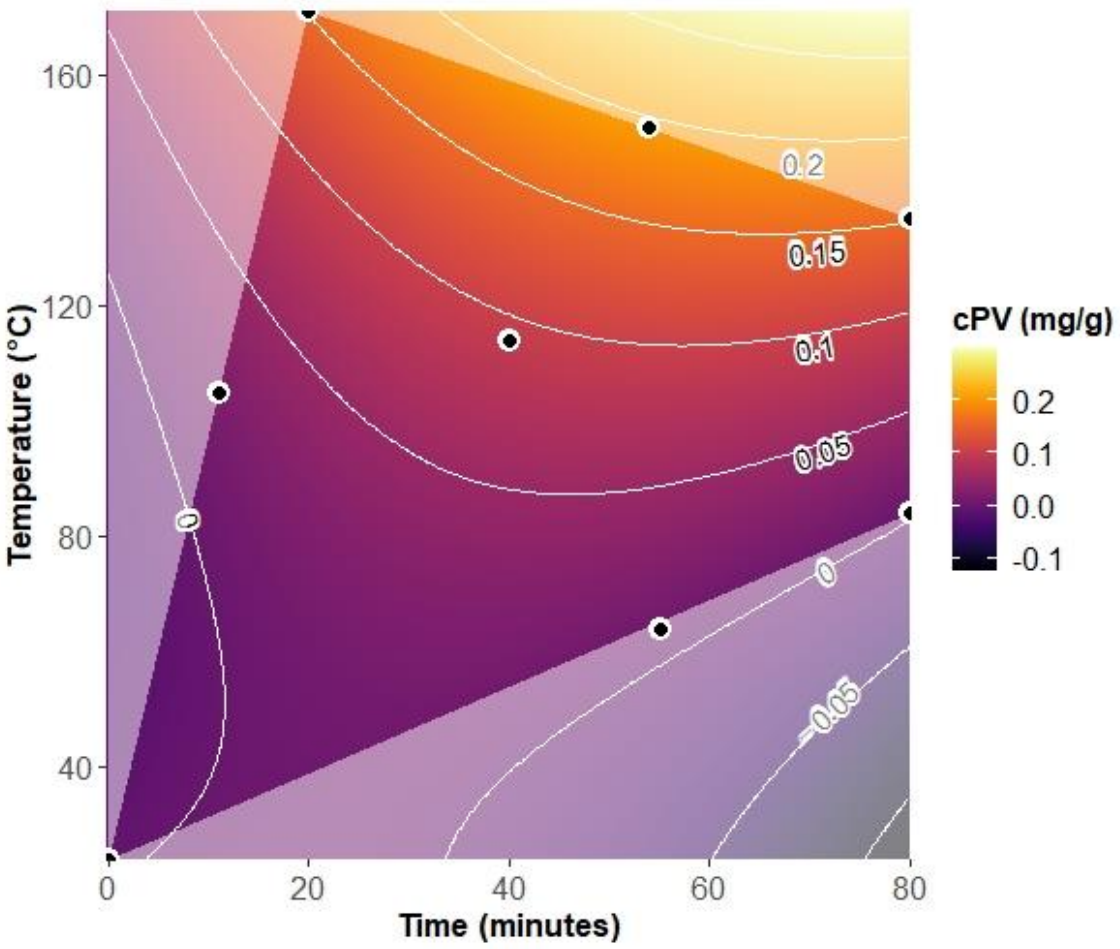

Figure 4. 12: Contour plot for predicted cyclo(Pro-Val) concentration for roasted chocolate liquor treatments across the entire experimental region for Madagascar, Ghana, and Peru.

As for origin-based differences, origin was not selected as an important predictor for cyclo(Pro-Val) concentration. This is likely because the range of concentrations for each

origin was quite similar across the roasting region (Table 4.4), with mean values also being quite similar (Table 4.13), meaning that cyclo(Pro-Val) can be relatively accurately predicted for any of the origins investigated without an origin effect term in the equation.

Table 4. 13: Region-wide mean concentrations of cyclo(Pro-Val)

Origin	Mean (mg/g) in chocolate liquor
Ghana	0.082
Madagascar	0.089
Peru	0.081
Overall	**0.084**

4.3.3.5 Caffeine

The BIC-selected model for best predicting Caffeine was:

$$\text{Caffeine} \sim \text{Origin} + \text{Time} + \text{Temp} + \text{Time:Temp} + \text{Origin:Time} + \text{Time}^2$$

However, diagnostic plots showed that the residual distribution was exhibiting some non-normality. After performing manual single-term additions and deletions, it was determined that removal of the Time^2 and Time:Temp terms dramatically improved the residual distribution, and other diagnostic plots showed no obvious patterns in the residuals, there were also no high-leverage points. Finally, the BIC value for the new model was actually slightly better (i.e., -38.7 instead of -37.8), and ANOVA showed that

there was not a significant difference between the two models (p-value=0.15), making the new model preferred. Therefore, the newly selected model became:

Caffeine ~ Origin + Time + Temp + Origin:Time

The summary of calculated model statistics can be seen in tables 4.14 and 4.15.

Table 4. 14: Summary of calculated model statistics, including effect estimates

```
Coefficients:
            Estimate  Std. Error     t value             Pr(>|t|)
(Intercept)   0.0000  0.016224453    0.00                1.000
Origin1      -0.9739  0.022944842  -42.45 < 0.0000000000000002 ***
Origin2      -0.3608  0.022944842  -15.73       0.000000000001 ***
Temp          0.0366  0.016960031    2.16                0.043 *
Time          0.0270  0.016960031    1.60                0.126
Origin1:Time -0.0231  0.023381926   -0.99                0.334
Origin2:Time -0.0384  0.023381926   -1.64                0.116

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0843 on 20 degrees of freedom
Multiple R-squared:  0.995,    Adjusted R-squared:  0.993
F-statistic:  606 on 6 and 20 DF,  p-value: <0.0000000000000002
```

Table 4. 15: ANOVA Table (Type III tests)

```
Response: Caffeine

            Sum Sq Df F value              Pr(>F)
(Intercept)   0.00  1    0.00               1.000
Origin       25.74  2 1810.96 <0.0000000000000002 ***
Temp          0.03  1    4.68               0.043 *
Time          0.02  1    2.55               0.126
Origin:Time   0.05  2    3.54               0.048 *
Residuals     0.14 20

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the adjusted R-squared is 0.993, suggesting that the predictors describe approximately 99.3% of the variance in the response, and the p-value for the model

(i.e., p-value<0.0000000000000002) is approaching 0 and therefore is statistically highly significant (i.e., p-value≤0.001). Interestingly, of the model selected (Table 4.15), only Origin is statistically highly significant (i.e., p-value ≤0.001), while Temp, and Origin:Time are statistically significant (i.e., p-value ≤0.05), and Time is not significant, but is maintained due to heredity. VIF values for all terms are below 2.5.

What can be seen in Table 4.14 above, and Figures 4.13, 4.14, and 4.15 below, is that in this experiment, little change occurs in caffeine concentration with temperature increase and time progression. This is clear for each of the origins (Figures 4.13, 4.14, and 4.15). What change does occur is likely due, at least in part, to the loss of moisture during the roast which can be seen in Figure 4.16. This makes sense, as moisture lost from cacao during roasting would lead to a slight concentration of any non-volatile compounds, such as caffeine, that do not degrade during the roasting process. When a model is fit to predict caffeine with moisture %, controlling for origin (i.e., Caffeine ~ Origin + Moisture), step-wise selection does not select the moisture term for removal, given that it has the best BIC, and moisture is relatively close to significant, with a p-value of 0.094, producing a model with an adjusted R^2 value of 0.99.

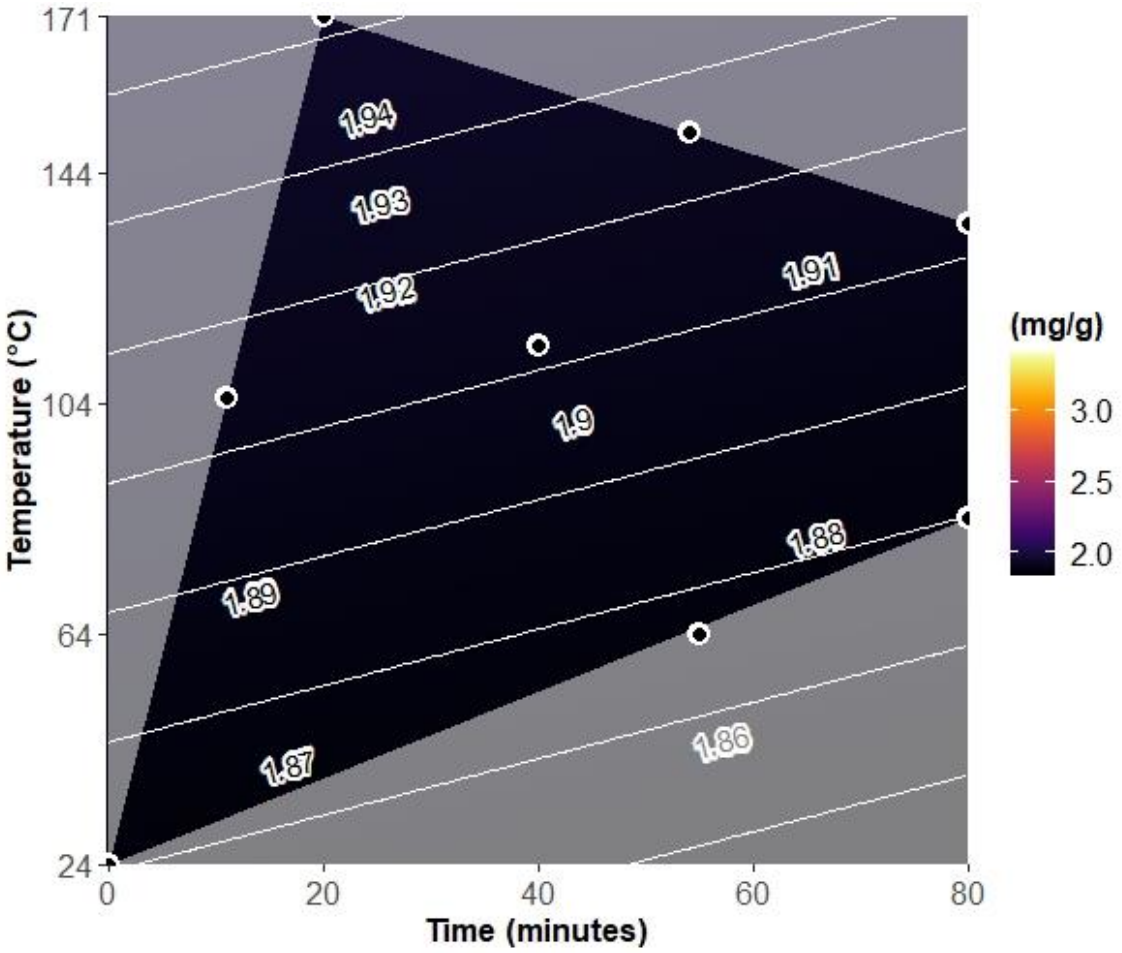

Figure 4. 13: MADAGASCAR: Contour plot for predicted caffeine concentration for roasted Madagascan chocolate liquor treatments across the experimental region.

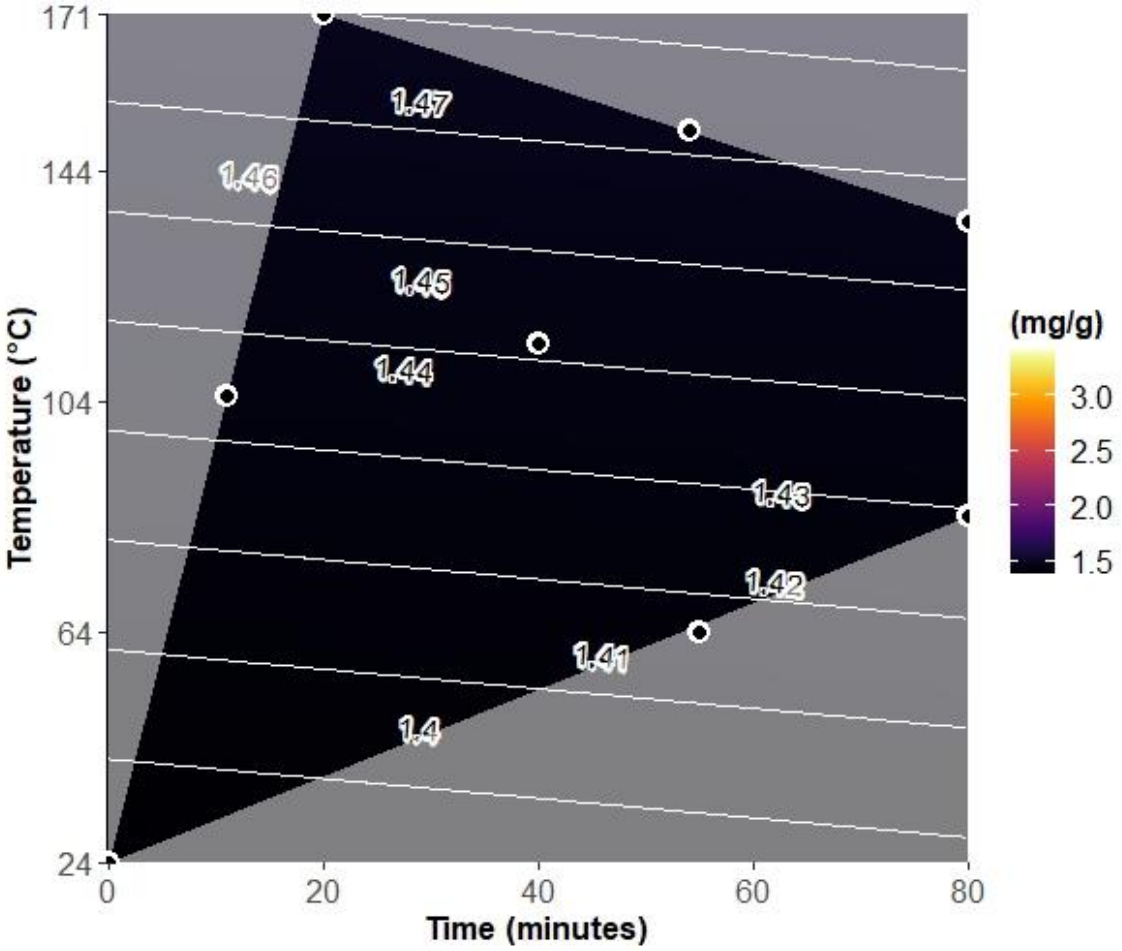

Figure 4. 14: GHANA: Contour plot for predicted caffeine concentration for roasted chocolate liquor treatments across the experimental region.

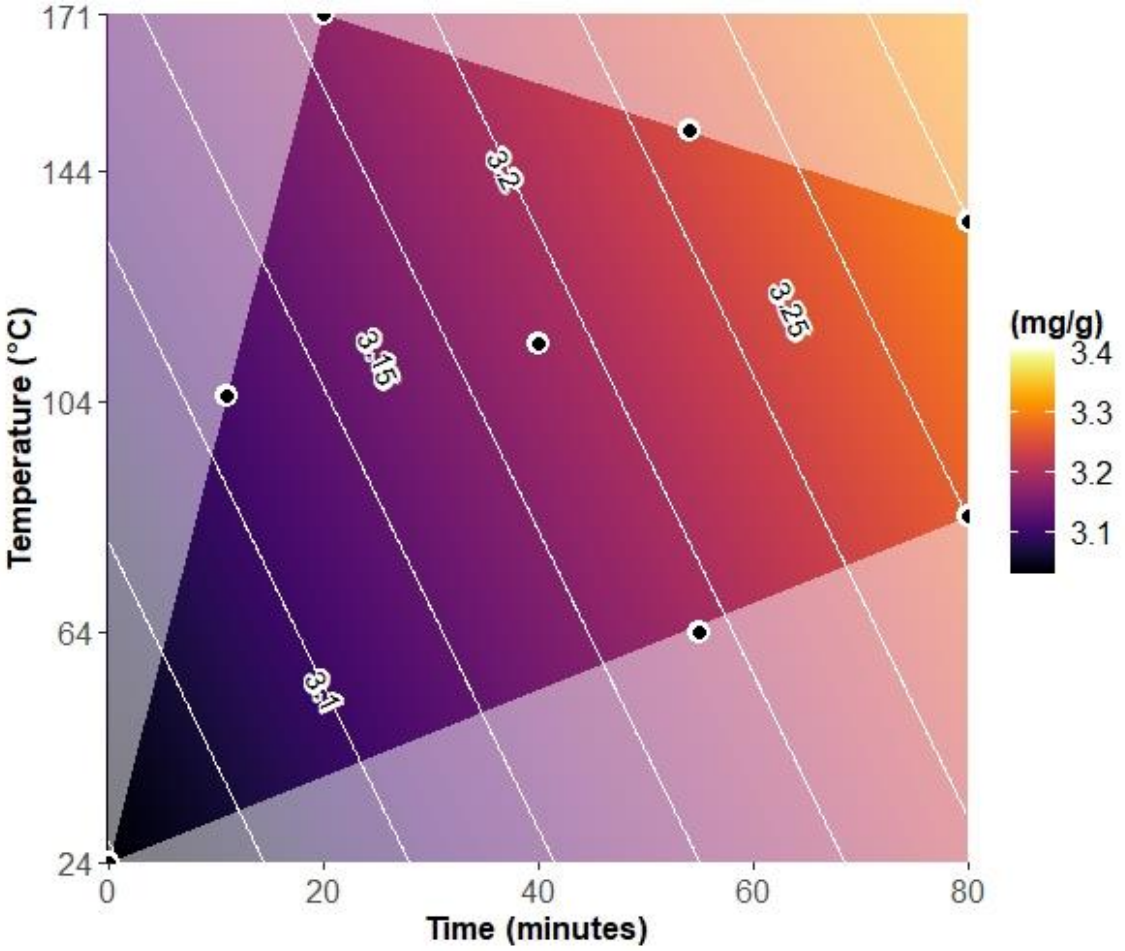

Figure 4. 15: PERU: Contour plot for predicted caffeine concentration for roasted Peruvian chocolate liquor treatments across the experimental region.

As for origin-based differences, origin is a particularly important predictor of caffeine in this experiment, with origin effects being large and statistically highly significant (i.e., <0.001). The Peruvian chocolate liquor had more caffeine than the other two origins, and Madagascar had relatively more than Ghana. Therefore, it can be said that roast-based differences in caffeine in chocolate are much smaller than origin-based differences in our experiment.

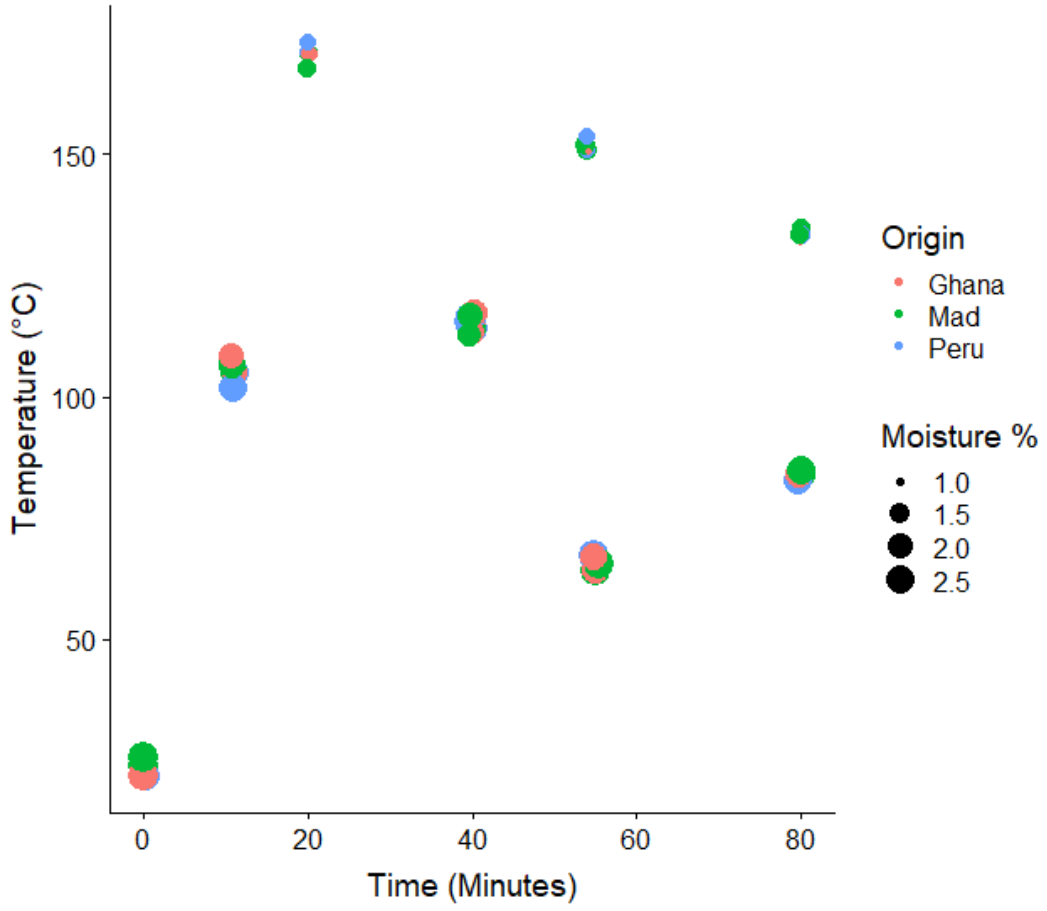

Figure 4. 16: Moisture % differences by roast treatment and origin, showing decrease in moisture with roasting. Slight jittering applied to overlapping points for increased visibility.

4.3.3.6 Theobromine

The BIC-selected model for best predicting Theobromine is:

Theobromine ~ Origin + Temp

Diagnostic plots do not show any obvious patterns in the residuals, there are no high-leverage points, and distribution of the residuals reasonably approximates a normal distribution given the strength of the p-value estimates of the model and model terms.

The summary of calculated model statistics can be seen in tables 4.16 and 4.17.

Table 4. 16: Summary of calculated model statistics, including effect estimates

```
Coefficients:
                          Estimate  Std. Error t value            Pr(>|t|)
(Intercept) -0.000000000000000316  0.01884    0.00               1.000
Origin1      0.824610586568035098  0.02665   30.94 <0.0000000000000002 ***
Origin2     -1.369556427024738454  0.02665  -51.39 <0.0000000000000002 ***
Temp         0.063775214461340865  0.01920    3.32               0.003 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0979 on 23 degrees of freedom
Multiple R-squared:  0.992,    Adjusted R-squared:  0.99
F-statistic:  896 on 3 and 23 DF,  p-value: <0.0000000000000002
```

Table 4. 17: ANOVA Table (Type III tests)

```
Response: Theobromine

            Sum Sq Df F value              Pr(>F)
(Intercept)   0.00  1       0               1.000
Origin       25.67  2    1339 <0.0000000000000002 ***
Temp          0.11  1      11               0.003 **
Residuals     0.22 23
---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the adjusted R-squared is 0.99, suggesting that the predictors describe approximately 99% of the variance in the response, and the p-value for the model (i.e., p-value <0.0000000000000002) is approaching 0, making it statistically highly significant (i.e., p-value ≤0.001). The Origin term is statistically highly significant (Table 4.17) (i.e., p-value ≤0.001), and the Temperature term is statistically significant (i.e., p-value ≤0.05). VIF values for all terms were below 2.5.

What can be seen in Table 4.16 above, and Figures 4.17, 4.18, and 4.19 below, is that in this experiment, very little change occurs in the Theobromine concentration across the experimental region within each origin, but what does occur is a slight increase in concentration related to temperature. This is clear for each of the origins (Figures 4.17, 4.18, and 4.19), and, as with caffeine, is likely due to moisture loss during roast (Figure 4.16). In fact, when a model is fit to predict theobromine with moisture %, controlling for origin, the moisture term is statistically significant (p-value≤0.05), with a p-value of 0.0026, and yielding a model with an adjusted R^2 of 0.991.

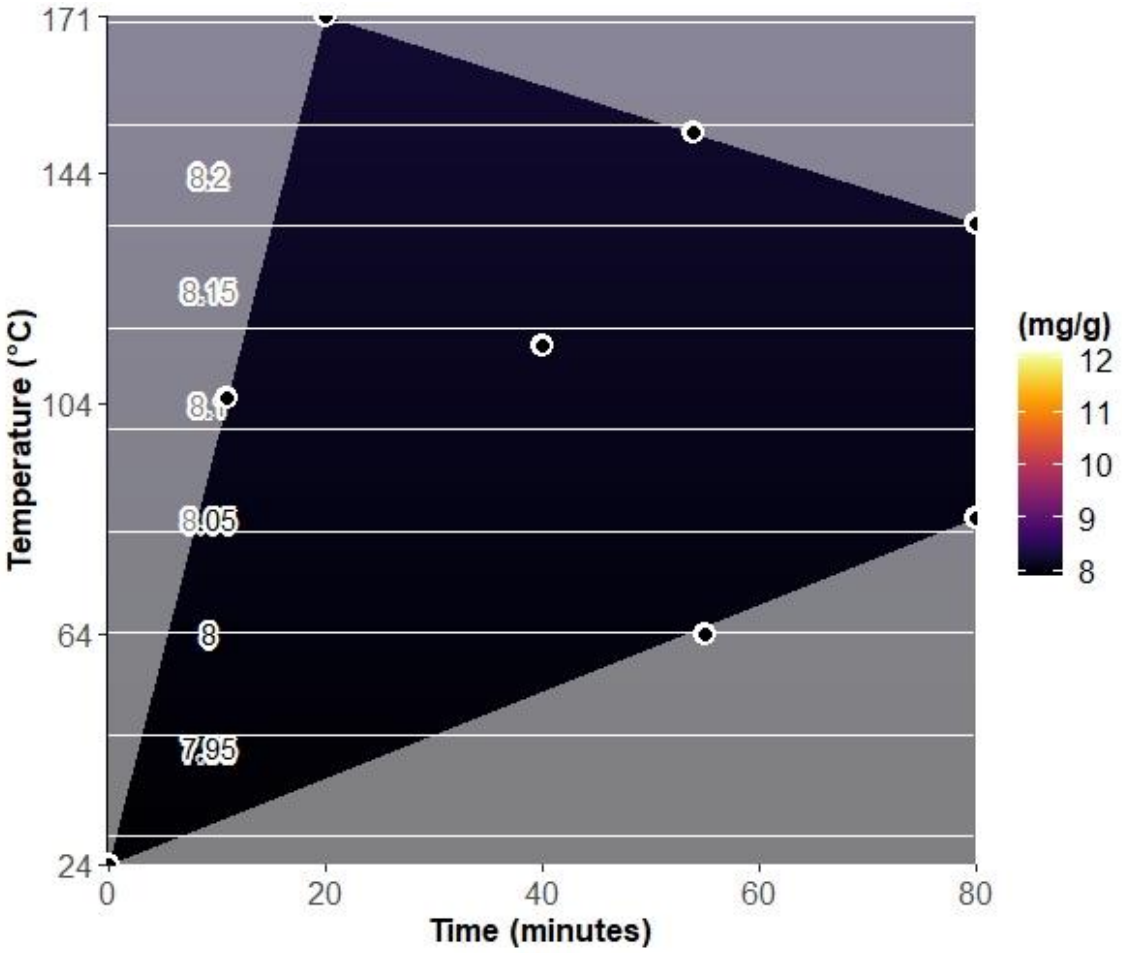

Figure 4. 17: MADAGASCAR: Contour plot for predicted theobromine concentration for roasted Madagascan chocolate liquor treatments across the experimental region.

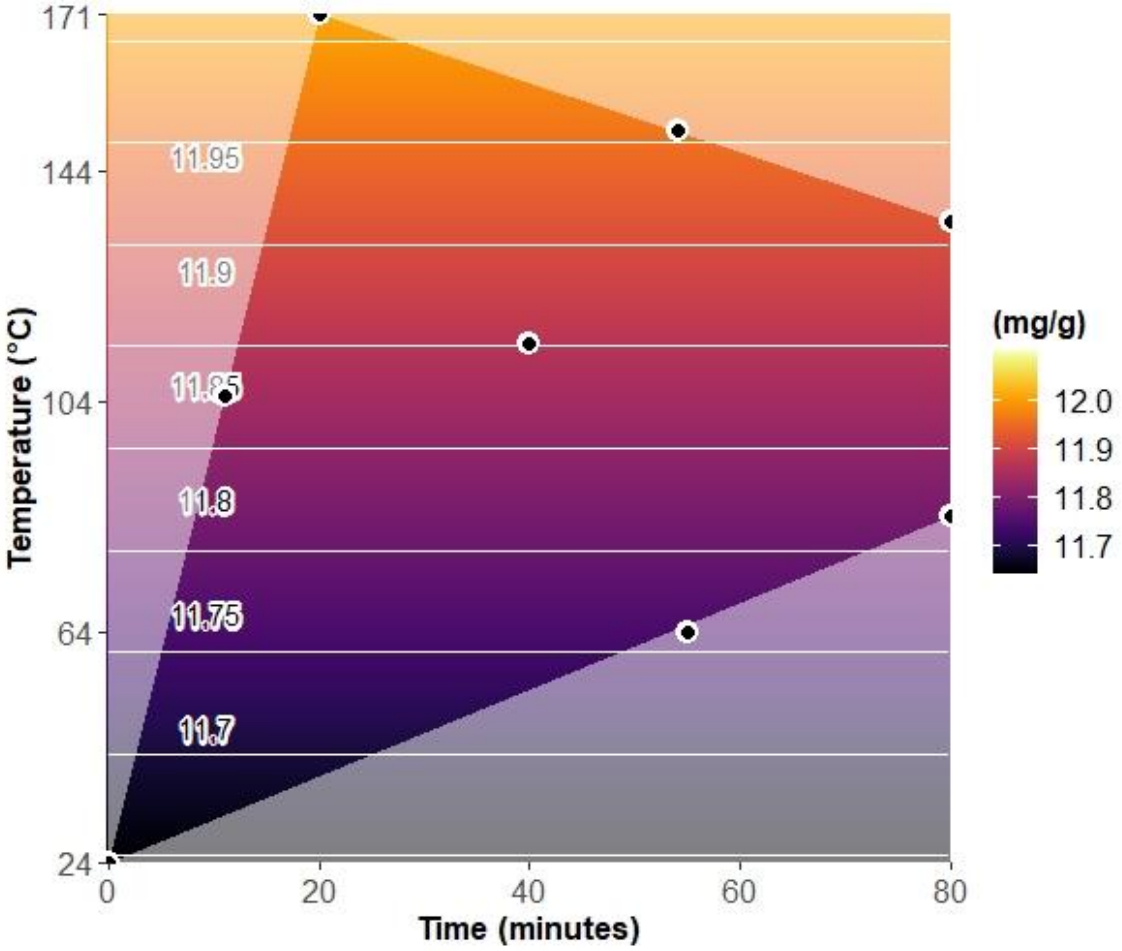

Figure 4. 18: GHANA: Contour plot for predicted theobromine concentration for roasted chocolate liquor treatments across the experimental region.

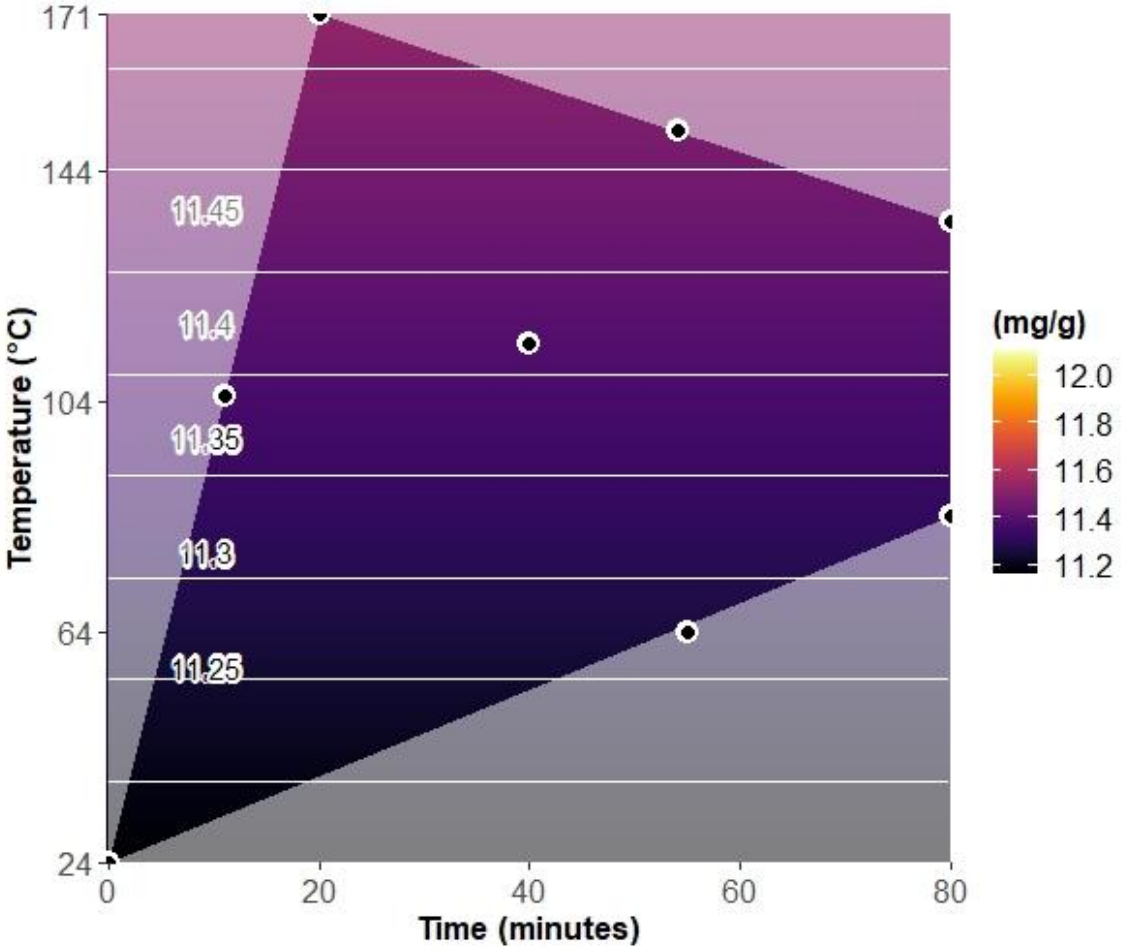

Figure 4. 19: PERU: Contour plot for predicted theobromine concentration for roasted Peruvian chocolate liquor treatments across the experimental region.

As for origin-based differences, origin is a particularly important predictor of theobromine in this experiment, with origin being statistically highly significant (i.e., <0.001) with a p-value <0.0000000000000002. This is most obvious when comparing figures 4.17, 4.18 and 4.19, as the change across the experimental region for each origin is quite moderate compared to the difference in concentrations between the origins, especially when Madagascar is compared to Ghana and Peru.

4.3.3.7 Overall Discussion

The above findings are relevant and interesting not just for the reasons already outlined, but because the most recent and commonly cited work on bitter compounds in cacao (Stark et al., 2006), was performed on a single sample of cacao from a single origin—Ghana—and with no available information on the quality of the prepared raw beans or roasting protocol. Yet, the publication claimed that the research reported within was "aimed at identifying the nonvolatile, key taste compounds responsible for the typical taste of roasted cocoa nibs" (Stark et al., 2006), and in this case, cyclo(Pro-Val) was for the first time claimed to be the most important bitter compound found in roasted cocoa nibs (Stark et al., 2006). However, though concentrations of certain bitter compounds determined in our experiment are sometimes comparable to Stark's findings (Stark et al., 2006), at least for certain treatments for specific origins, as seen in table 4.18 below, for example in the case of theobromine concentration for Peru, caffeine concentration for Ghana, and epicatechin concentration for Peru, in no case do the concentrations in any of our treatments approach that reported by Stark for cyclo(Pro-Val). In fact, even the highest value determined across the entire experimental region for any of the three origins (i.e. Ghana, Madagascar, and Peru), from chocolate liquor made of Madagascan cacao roasted for 20 minutes at 171°C, resulted in a concentration that was only about 11% as much as that reported by Stark (2006). This raises the question whether there may have been something abnormal in the ripeness, fermentation, or roast profile of the single sample analyzed for which data was reported by Stark. This many years later it is impossible to be certain, but given that we did not find a similar concentration in

any treatments from any origin, it seems well within the realm of possibility that the sample analyzed by Stark (2006) could have been abnormal in some respect. And if so, it might not necessarily be the case that cyclo(Pro-Val) is an important bitter compound in cacao, let alone the most important bitter compound as claimed (Stark et al., 2006). At the least, our work suggests much more study of DKPs in cacao, and cyclo(Pro-Val) in particular should be done by future researchers.

Table 4. 18: Comparisons between concentrations of important bitter compounds in previous research and this current research.

Closest values to those reported by Stark (2006) from all possible raw/roasted chocolate liquor treatments analyzed are reported for comparison purposes. All concentrations are in mg/g of ground cocoa/chocolate liquor. Mad=Madagascar

Time (minutes)	unknown	55	55	40	54	40	20
Temp (°C)	unknown	64	64	114	151	114	171
	Ghana (Stark)	**Peru**	**Ghana**	**Peru**	**Peru**	**Peru**	**Mad**
Theobromine	11.45	11.49					
Caffeine	1.01		1.37				
Epicatechin	2.50			2.21			
Catechin	0.69				0.63		
Procyanidin B2	1.21					1.20	
Cyclo(-Pro-Val)	1.74						*0.19

***Not similar. Only 11% of Stark's reported concentration

In summary, the Response Surface roasting experiment detailed herein was useful in confirming and elucidating previously determined findings regarding changes in important bitter compounds in cacao during roasting. Regarding the elements that may most obviously be controlled in determining a roast profile (i.e., time and temperature), as predictors, they were both significant for describing the extent of changes in bitter compounds in cacao in most cases, with theobromine being a noted exception in the case of time. Yet, even where time was an important term in a given model, the effect of temperature was consistently responsible for the largest change in concentration of bitter compounds within our experimental region, which we consider to include time and temperature combinations that are analogous to a range that may reasonably be chosen in commercial cacao roasting of whole cacao beans (Beckett et al., 2017; Whymper, 1921), as well as including raw treatments that are applicable to the production of raw chocolate. The specific changes within the roasting region were highlighted above in various figures (4.3 – 4.19), but in general, the trends of the effects of roasting on the bitter compounds of interest are as follows:

Roasting increased concentrations of the DKP cyclo(Pro-Val), as well as the flavan-3-ol catechin, while decreasing concentrations of both the flavan-3-ols epicatechin and its dimer procyanidin B2. Additionally, minor increases in caffeine and theobromine occurred, which may reasonably be explained by moisture loss as described above. However, time and temperature were not the only relevant terms in our models. An origin effect was present for each compound but one. By sampling from multiple lots of

several origins where, at least in the case of Marañon from Peru, it has previously been documented to be part of a unique genetic cluster, and by documenting the quality of the cacao prior to analysis, we are able to reasonably confirm that differences in bitter compounds may at least sometimes be expected based upon origin. How much variation exists within certain origins and across origins remains to be further determined, but in our study the Marañon Peruvian cacao was found to have the largest amounts of most of the compounds (i.e., catechin, epicatechin, procyanidin B2, caffeine) relative to roast profile. While Ghana and Peruvian Marañon had similar concentrations of Theobromine relative to roast profile. Madagascar had the lowest levels of theobromine, catechin, and procyanidin B2, and comparable levels of epicatechin as Ghana, with slightly higher amounts of caffeine than did Ghana. Interestingly, the only compound that exhibited no significant origin effect in its predictive model was cyclo(Pro-Val), though an increase in the compound was seen due to roast, with both time and temperature being significant, as expected based upon previous literature (Bonvehí et al., 2000; Rizzi, 1989; Ziegleder, 2017).

4.4 Conclusion

This study supports previous work as regards trends in general concentration changes during roasting of compounds in the three classes of bitter compounds studied (i.e., methylxanthines, 2,5-diketopiperazines, and flavan-3-ols). However, given that the previously stated importance of cyclo(Pro-Val) for cacao bitterness (Stark et al., 2006) was in large part dependent upon its concentration in a single analyzed sample with unknown fermentation or roasting protocol, and given that our quantitative chemical data does not support this previous finding, more research is required. Specifically, analysis of cacao from a variety of origins, with varied roast profiles, and in conjunction with rigorous sensory analysis should be undertaken before it can be certain whether any specific DKP, or the class as a whole, are important for cacao bitterness.

CHAPTER 5

5 CONSUMER SENSORY ANALYSIS OF CHOCOLATE MADE FROM THREE DIFFERENT ORIGINS OF CACAO USING EIGHT ROAST PROFILES

5.1 Introduction

5.1.1 Cacao and Chocolate

5.1.1.1 Cacao

Also known as cocoa, cacao consists of the fermented and dried seeds of the fruit of the tropical *Theobroma cacao* tree in the *Malvaceae* family (Aprotosoaie et al., 2016a). Cacao is a significant food commodity, with annual global consumption reaching approximately 4.6 million metric tons as of 2018 with an increase in demand of 3.9% over 2017 (Barchart, 2019), and plays an important role as the primary ingredient in the manufacture of chocolate (Aprotosoaie et al., 2016a).

5.1.1.2 Chocolate

Chocolate is a usually sweetened, uniquely flavored, solid paste that melts smoothly at human body temperature due to the presence and unique fatty acid composition of cacao fat, called cocoa butter (Aprotosoaie et al., 2016a). The aforementioned unique chocolate aroma is due mostly to the presence of cacao, required in American-made chocolate by FDA Title 21 (2019). Prior to transformation into chocolate, cacao is roasted to obtain a more complex flavor and character that is generally preferred by consumers over that of raw cacao (Aprotosoaie et al., 2016a).

5.1.2 Bitterness and its Sensation and Perception

5.1.2.1 Bitterness

Bitterness is one of the five taste modalities (i.e., salty, sweet, sour, bitter, umami) sensed by the tongue (Gaudette et al., 2013; Keast et al., 2003b). It has been noted that toxic compounds are often bitter (Keast et al., 2003b), and the ability of humans to taste bitter substances is likely to have evolved as a form of toxin detection (Keast et al., 2003a). This may explain why bitterness is generally disliked by humans (Drewnowski et al., 2000; Fischer et al., 2005) and even rejected in most food (Gaudette et al., 2013), despite famous exceptions such as coffee, beer, red wine, and dark chocolate (Gaudette et al., 2013; Keast et al., 2003b; Roy, 1997) that highlight the sometimes complex nature of human food choices (Gaudette et al., 2013).

5.1.2.2 Bitterness Sensation

Bitter sensation begins as a ligand/taste molecule contacts a bitter receptor cell, and bitter taste transduction results, with nerve impulses being sent to the brain (Drewnowski, 2001; Lawless et al., 2010b). Bitter taste receptor cells are clustered as part of taste buds, located on the papillae across the tongue, with more located on the palate, and in the throat (Drewnowski, 2001; Lawless et al., 2010b). Of the basic tastes (i.e., bitter, sweet, sour, salty, umami), bitter is the most complex (Drewnowski, 2001), and it is now known that in humans there are approximately 25 subtypes of G-protein-coupled receptors called TAS2Rs which are responsible for the transduction of bitter

taste from many thousands of compounds (Dagan-Wiener et al., 2018; Maehashi et al., 2009). Each bitter taste sensation is complex and unique, with four properties to define it: quality, intensity, temporal patterns and spatial patterns (Keast et al., 2003b).

5.1.2.3 Bitterness Perception

Bitterness perception starts with the sensation of bitter compounds, but also includes processing by the brain of incoming signals from other sensory modalities (i.e., other tastes, aromas, and somatosensory, aural, and visual inputs) (Lawless et al., 2010a). In short, bitterness perception is a combination of sensation and central cognitive effects resulting from other concurrently processed sensory information (Keast et al., 2003b). For example, aural stimulation (i.e. music) (Carvalho et al., 2017) can even affect bitterness perception.

5.1.2.4 Variation in Bitterness Sensation/Perception

There is a great deal of variation across human individuals regarding sensation of bitterness (Drewnowski, 2001; Mennella et al., 2005; Negri et al., 2012), and the ability to sense certain bitter compounds at all, which can be inherited (e.g., phenylthiocarbamide (PTC) and 6-n-propylthiouracil (PROP) (Drewnowski, 2001)). Genotype therefore leads to a large variance between individuals regarding sensitivity to bitter compounds, where approximately 1/3 of people are unable to sense certain bitter compounds, while others, so-called supertasters, show hypersensitivity to bitter compounds, due in part to larger numbers of taste buds, with all other individuals falling upon a spectrum between these extremes (Lawless et al., 2010b). Additionally, children

avoid bitter foods to a larger extent than do adults (Negri et al., 2012), and sex-based differences have also been described (Bartoshuk et al., 1994).

Variation in bitterness perception depends on more than genotype, age, and sex, however. It is an interesting case that perception of overall bitterness intensity of a mixture of bitter compounds at known concentrations is less than that of the sum of intensities of the individual compounds at the same concentrations (Keast et al., 2003b). It is also the case that bitterness may be suppressed by certain compounds with sweet, salty, and umami tastes, and enhanced by those with sour ones (Calviño et al., 1993; Drewnowski, 2001; Fischer et al., 1994). Short-term variance in perception also exists, when exposure to bitter compounds can lead to adaptation (i.e., decreased responsiveness) to bitterness (Lawless et al., 2010b), and bitterness can even be altered by the concentration of calcium ions in an individual's saliva (Neyraud et al., 2004).

5.1.3 Causes of Bitterness Variation in Cacao

Factors that impact bitterness in cacao, and therefore chocolate, are numerous, including varietal, growing conditions (e.g., hours of sunlight and rainfall), ripeness at harvest, and post-harvest processing (Kongor et al., 2016). Geographical location, even within a single country, appears to play a role in overall methylxanthine concentration and theobromine to caffeine ratio, probably due to altitude of the specimens from which the cacao is derived (Carrillo et al., 2014). However, cacao varietal, ripeness at harvest, and the multi-stage post-harvest processing (i.e., fermentation and drying)

have perhaps been most commonly noted as playing important roles (Afoakwa et al., 2008; Aprotosoaie et al., 2016a; Beckett et al., 2017; Kongor et al., 2016).

5.1.3.1 Varietal

Regarding cacao varietal, while ongoing research has led to a better understanding of the cacao genome and genetic variation in cacao generally (Johnson et al., 2009; Motamayor et al., 2008; Takrama et al., 2014; Zhang et al., 2011; Zhang et al., 2012), the sensory characteristics of the 10 identified genetic clusters of cacao (Motamayor et al., 2008) have yet to be thoroughly characterized, though an ongoing sensory evaluation project at Penn State will likely result in useful data on this topic (Hopfer & Brown, 2018, personal correspondence). Certainly, it is already clear that specific varietals show variation in bitterness (Kongor et al., 2016), with theobromine concentration (Bonvehí et al., 2000) and ratios of theobromine to caffeine being significantly different across origins/varietals/populations (Timbie et al., 1978; Trognitz et al., 2013; Zoumas et al., 1980). Additionally, there appear to be differences in the diketopiperazine (DKP) concentrations of the roasted cacao prepared from fermented and dried cacao of different origins (Rizzi, 1989).

5.1.3.2 Ripeness

Ripeness of the cacao fruit at the time of harvest is known to play an important role in cacao flavor in general (Afoakwa et al., 2008) and in bitter methylxanthine concentration in particular (Bonvehí et al., 2000; Timbie et al., 1978) as caffeine and theobromine both increase substantially between 12 weeks from flowering and fruit

harvest at 5-6 months (Senanayake et al., 1971). Cacao flavonoid profile, including flavan-3-ols, also depends upon ripeness (Nazaruddin et al., 2006; Rusconi et al., 2010), as unripe cacao has 29% less epicatechin than ripe cacao (Payne et al., 2010).

5.1.3.3 Post-Harvest Processing

5.1.3.3.1 Fermentation

While post-harvest processing, which consists of fermentation and drying, is a well-known cause of bitterness reduction in cacao (Afoakwa et al., 2008; Aprotosoaie et al., 2016a; Beckett et al., 2017), it is a complex one with many unknown factors of its own (John et al., 2019). It is known that during cacao fermentation in the presence of oxygen and polyphenol oxidase, flavonoid oxidation and polymerization occur, leading to browning, and reduction of bitterness and astringency (Ziegleder, 2017). Additionally, fermentation may result in an up to 40% reduction in bitter methylxanthines due to their diffusion through the cacao seed coat and loss in fermentation runoff/waste (Bonvehí et al., 2000), and flavonoids are also likely to be lost in runoff (Cooper et al., 2007). It has also been noted that the ratio of catechin to epicatechin shifts during fermentation, increasing during the process (Porter et al., 1991), likely due to increases in (-)-catechin as (-)-epicatechin concentration decreases (Cooper et al., 2007; Payne et al., 2010). Overall, there is a >80% decrease in total catechin monomers (i.e., epicatechin and catechin) during fermentation (Payne et al., 2010).

5.1.3.3.2 Drying

Well-managed drying practices result in further reductions in bitterness via additional flavonoid oxidation (Barišić et al., 2019; Ziegleder, 2009). Interestingly, and perhaps unexpectedly, very low concentrations of certain DKPs have been found in fermented and dried, but *unroasted*, cacao (Rizzi, 1989).

5.1.3.4 Roasting

Roasting, considered by some to be the most important step in processing cacao (Aprotosoaie et al., 2016a), results in the creation of bitter diketopiperazines (DKPs) from peptides (Rizzi, 1989; Ziegleder, 2017), and darker roasts, particularly at higher temperatures, appear to increase DKP levels the most (Bonvehí et al., 2000), whereas unroasted cocoa contains virtually no DKPs (Bonvehí et al., 2000). However, roasting is required to alter the harsh, unpleasant flavor of raw cacao, and to develop characteristic cocoa aroma (Ziegleder, 2009). Additionally, roasting alters the concentrations of epicatechin, and its epimers and oligomers (Kothe et al., 2013; Stanley et al., 2018), compounds which are both bitter and astringent (Stark et al., 2006), sometimes in unexpected ways seemingly related to varietal (Kothe et al., 2013). Loss of epicatechin at temperatures over 70°C occurs, and at a roasting temperature of 120°C catechin content has been seen to increase by approximately 650% in fermented cacao (Payne et al., 2010). This is probably due to epimerization of epicatechin to catechin (De Taeye et al., 2014; Oracz et al., 2015; Payne et al., 2010) and decomposition of procyanidins (Oracz et al., 2015), probably first to epicatechin followed by epimerization (Zhu et al., 2002). For this reason, the epicatechin/catechin ratio helps to understand the previous

processing of cacao (Payne et al., 2010). Interestingly, roasting also leads to production of high molecular weight melanoidins that bind to polyphenols such as catechin, epicatechin, and procyanidin B2 (Oracz et al., 2019). As for the epimerization of epicatechin during roasting, it appears to be from (-)-epicatechin to (-)-catechin, even though the naturally present form of catechin is (+)-catechin (De Taeye et al., 2014; Hu et al., 2016; Hurst et al., 2011; Payne et al., 2010). Cooper et al. (2007), was the first to show this (-)-epicatechin to (-)-catechin epimerization due to cacao processing, using a chiral column with UPLC (Cooper et al., 2007), and this was later confirmed by Kothe et al. (2013). Cooper et al. (2007) also showed that there is a strong linear relationship between epicatechin and procyanidins, as well as (+)-catechin, so that they can all be predicted from the epicatechin concentration, while (-)-catechin, however, cannot be predicted in the same fashion (Cooper et al., 2007). Finally, it has been known for decades that a small loss of methylxanthines from the cotyledons to the cacao shell is expected during roasting, apparently increasing with degree of roast (Wadsworth, 1922).

5.1.4 Further Context

While this research project began due to a significant interest in understanding the relationship between roasting (i.e., application of heat at specific temperatures for specific lengths of time) of cacao and consumer perception of chocolate made therefrom, especially as it pertains to Bitterness and Liking, we asked also about other characteristics in order to minimize “dumping” (Kappes, Schmidt, & Lee, 2006; Lawless et al., 2010b). This is a well-known situation where sensory evaluators have no option to

rate on the ballot a particular characteristic that they are sensing, and so they dump it into ratings for another characteristic, which is likely to be one of interest to the researcher, thereby adding regular but varying bias to the rating of the characteristic of interest (Kappes et al., 2006; Lawless et al., 2010b). In our case that might consist of consumer participants rating Sourness as part of Bitterness. So, we asked specifically for ratings of other known chocolate characteristics, such as Sourness, Astringency, etc. Upon completion of the study, this left us with a wealth of data based upon the 4970 observations of 8 sensory and hedonic characteristics of those 8 treatments per origin, across three origins. Though Bitterness and Liking were the main characteristics of interest, it seemed a shame to relegate the rest of the data to the dustbin of dumping avoidance, and so we decided to model much of it just the same. We ask for the reader's forgiveness at the length added to this chapter due to such a decision, but we found it to be of great interest.

As for our analytical approach to this chapter, it consists of several methods, with details further described in the Results & Discussion section of this chapter: PCA-based methods, and Mixed-Model Linear Regression. These methods are quite a common part of the modern sensory analysis toolbox (Lê & Worch, 2014) with mixed-model linear regression and type III ANOVA, specifically, being part of the "classic bread and butter technique" (Lawless, 2013; Næs, Brockhoff, & Tomić, 2010).

5.2 Materials & Methods

5.2.1 Materials

5.2.1.1 Cacao

All cacao from each of three origins (i.e., Madagascar, Ghana, Peru), was fermented and dried at origin before being packed into approximately 65 kg burlap sacks, themselves packed into steel shipping containers, and then transported to cacao warehouses in the United States by cargo ship.

5.2.1.1.1 Madagascar:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from three separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total (i.e., 7 kg per lot): Lots #2182432 (2018 harvest), #2672711 (2018 harvest), #1427425 (2018 harvest). This cacao was obtained from Guittard Chocolate (Burlingame, CA) Upon receipt, all three lots of this origin were blended into a homogenous composite, and hand-sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.), unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro (Concord, MA) Supergrain Premium RT bags with high vapor and gas barrier characteristics at <65% RH and <27 °C until roasted (approximately one month or less).

5.2.1.1.2 Ghana:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from three separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total (i.e., 7 kg per lot): Lot #19003223 (2018/2019 main harvest), #482019 (2018/2019 main harvest), #3729 (2017/2018 main harvest). This cacao was obtained from Guittard Chocolate (Burlingame, CA). Upon receipt, all three samples of this origin were blended into a homogenous composite, and hand-sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.), unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro (Concord, MA) Supergrain Premium RT bags with high vapor and gas barrier characteristics at <65% RH and <27 °C until roasted (approximately one month or less).

5.2.1.1.3 Peru:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from two separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total from two separate lots covering multiple months as follows: 7 kg was sampled from a lot consisting of the January through early March 2018 harvest, and 14 kg (7 kg x 2) was sampled from a lot consisting of the late March to late June 2018 harvest. This cacao was obtained from Marañon Cacao (San Diego, CA). Upon receipt, all three samples of this origin were blended into a homogenous composite, and

hand-sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.), unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro (Concord, MA) Supergrain Premium RT bags with high vapor and gas barrier characteristics at <65% RH and <27 °C until roasted (approximately one month or less).

5.2.1.1.4 Cacao Quality Parameters

Cacao quality parameters related to phenotype, ripeness, fermentation, and drying, of a representative portion (180 g) of each 21 kg composite are detailed in Table 5.1. Measurements were made just prior to conducting all roasting treatments. All three samples fall within acceptable parameters for good quality cacao according to the International Cocoa Organization (ICCO) (2015).

Table 5. 1: **Preliminary statistics on composite samples from each origin**

Origin:	Ghana	Madagascar	Peru
Bean-Count:	89 beans per 100 g	86 beans per 100 g	74 beans per 100 g
Slaty Color:	1%	0%	0%
Fully Purple:	1%	0%	26%
Partly Purple:	17%	6%	18%
Total Slaty or Purple:	19%	6%	44%
Light Brown:	0%	27%	7%
Medium Brown:	58%	62%	27%
Dark Brown:	23%	5%	10%
Total Brown:	81%	94%	44%
White Beans:	0%	0%	12%
Moisture Content:	6.13 %	6.68%	5.83%
Internal Mold:	0%	0%	0%
Infested:	1%	0%	1%

5.2.1.2 Additional Materials and Analytical Equipment

The roasting oven consisted of a Binder GmbH (Tuttlingen, Germany) model # FD56 120 V precision laboratory oven with forced air convection. This model of oven was also used for holding chocolate liquor in a just-melted state prior to the sensory analysis experiment. Roasting temperatures were monitored and recorded with an Omega (Norwalk, CT) model #RDXL4SD 4-channel datalogger with Type-K thermocouples (part #5sc-tt-k-30-36). Room temperature and humidity during roasting were monitored with a SensorPush (Brooklyn, NY) #HT1 device with humidity calibrated with Boveda (Minnetonka, MN) Low-RH One-Step 32% calibration kit model #CAL32-SM, and temperature and humidity verified with a Bacharach (New Kensington, PA) Sling Psychrometer model #0012-7012.

Preliminary cacao cooling was performed by a Lasko (West Chester, PA) 20” box fan model #B20200, placed under ¼” stainless mesh bolted to Cambro (Huntington Beach, CA) food bin. Cacao cracking was performed with a custom CrankandStein (Atlanta, GA) 305 mm 3-roll cocoa cracker, and winnowing was performed using a custom food-grade winnower with 4” clear FDA-approved for food contact PVC, with air flow powered by a Grizzly Industrial, Inc (Springfield, MO) model #:G0441 3HP cyclone.

Chocolate liquor was ground using a Spectra 11 (Tamil Nadu, India) Stone Wet Grinder model #101, and strained through a Kitchenaid (Benton Harbor, MI) model # KES161OHOBW fine-mesh strainer. Cooled and solidified chocolate liquor was wrapped in aluminum foil and vacuum sealed in FoodSaver (Oklahoma City, OK) brand multi-layer vacuum bag with nylon vapor barrier. Portioning of chocolate liquor for the sensory evaluation experiment was carried out using positive displacement pipettes (Eppendorf (Hamburg, Germany) Repeater M4 pipette and Eppendorf Repeater E3 pipette) with disposable pipette tips (Eppendorf Combitips Advanced Lot# I183098J).

Food-grade ingredients were used for training calibration solutions prior to chocolate liquor assessments, consisting of tannic acid obtained from Spectrum Chemicals (New Brunswick, NJ), citric acid and caffeine obtained from Sigma-Aldrich (St. Louis, MO), and C&H (Crockett, CA) Pure Cane Sugar (sucrose) obtained from a local grocery store.

5.2.2 Methods

5.2.2.1 Roasting Experimental Design

For the roasting experimental design (Table 5.2), temperature range (i.e., 24°C to 171°C) and time range (i.e., 0 to 80 minutes) were chosen as reasonable ranges for potential modification of bitterness based upon literature regarding common roasting temperatures (Afoakwa et al., 2008; Ziegleder, 2017) and the advice of chocolate professionals; this includes a raw treatment at 24°C (approximate room temperature) for 0 minutes as a control. Combinations of time and temperature that are impossible, judged so extreme so as to be burnt, or that in essence are repeats of other combinations (e.g., 0 minutes and any temperature aside from 24°C, or 24°C and any time combination aside from 0 minutes), were excluded. All time and temperature combinations are repeated for each of the three origins (Madagascar, Ghana, and Peru). Roast order was randomized. The specific time and temperature combinations within these ranges were chosen as part of a Response Surface Methodology (RSM) approach, using JMP 14.0.0 software (Cary, NC) and an I-Optimal algorithm, which minimizes average variance of prediction for potential model coefficients (Jones et al., 2012; Myers et al., 2016; Oyejola et al., 2015), while at the same time seeking to minimize covariance of model coefficients (Oyejola et al., 2015). The resulting design is an irregularly shaped non-rotatable design, akin to a central composite design (CCD), with a duplicate centerpoint to allow for pure error estimation. In addition to the previously mentioned reasons, this model was chosen because standard CCDs have already been commonly and successfully used in roasting optimization experiments (Farah et al., 2012;

Kahyaoglu, 2008; Lee et al., 2001; Madihah et al., 2012; Mendes et al., 2001; Özdemir et al., 2000), while it has also been noted that computer-based optimal design algorithms are effective at helping to choose designs for complex problems that cannot easily be solved with standard design types, such as in the example of irregular experimental regions (Goos et al., 2011; Jones et al., 2012).

Table 5. 2: Randomized modified I-optimal experimental design for roasting

Minutes	Temperature °C	Origin	Randomized Roast Order
40	114	Peru	1
40	114	Ghana	2
54	151	Madagascar	3
80	84	Peru	4
0	24	Peru	5
20	171	Ghana	6
80	84	Ghana	7
20	171	Madagascar	8
54	151	Peru	9
55	64	Peru	10
40	114	Peru	11
0	24	Ghana	12
40	114	Ghana	13
40	114	Madagascar	14
11	105	Peru	15
55	64	Madagascar	16
11	105	Madagascar	17
54	151	Ghana	18
11	105	Ghana	19
40	114	Madagascar	20
80	135	Peru	21
80	135	Ghana	22
0	24	Madagascar	23
80	84	Madagascar	24
55	64	Ghana	25
80	135	Madagascar	26
20	171	Peru	27

5.2.2.2 Roasting and Winnowing Method

For roasting and winnowing, 410 g of each cacao origin, specified by the experimental design roast order (Table 5.2), was weighed into each of two stainless steel mesh roasting trays (820 g total). The cacao was spread into an evenly distributed single-bean

layer on each tray. A thermocouple was placed in a hole made with a steel needle in a bean of approximate average size in the middle of each tray. The thermocouples were then connected to the datalogger, and the datalogger was set to record. The oven was preheated to the setpoint required in the roasting design (Table 5.2), with the convection fan on. Once the setpoint was reached, 10 additional minutes were allowed to elapse to assure preheating completion. Once pre-heating was completed, the timer was set to the required roasting time (Table 5.2), and the two trays were loaded into the oven with one tray right above the fan and the other right below it. The oven door was immediately closed (i.e., in less than approximately 5 seconds) and the timer was started.

Once the roast was complete, the oven was turned off, and the door was immediately opened. Forced-air cooling of the roasted cacao was employed to minimize carry-over roasting through immediate winnowing (i.e., breakage and removal of cocoa shell), augmented with preliminary fan cooling for higher temperature roasts (i.e., roasts of greater than 114°C in Table 5.2), to attain sufficient temperature decreases for all treatments (i.e., a drop to 50°C or less within 5 minutes). The end-result of this cooling and winnowing process, when augmented with manual removal of any remaining visible shell pieces, was room temperature cacao nibs, the main ingredient of chocolate liquor.

Each roast was completed in duplicate on the same day, and nibs from the roasted and winnowed duplicate treatments blended until homogeneous. Therefore, 1640 g of cacao was roasted in total for each treatment. Of this 1640 g, 1000 g of blended roasted

cacao nibs were allocated for chocolate liquor production and placed in a food-grade polypropylene bag, with air substantially removed, and labeled with all identifying information. The remaining 640 g were saved in case of future need. Nibs were stored in a room with temperature below 19.5 °C with relative humidity (RH) at approximately 40% or less and turned into chocolate liquor within 48 hours.

5.2.2.3 Chocolate Liquor Production Method

For chocolate liquor production, the stone bowl and grinding stones of the Spectra 11 wet grinder were preheated to approximately 49°C to 54°C overnight in a warming cabinet. Once the stone grinder components were mounted to the machine, it was immediately turned on and 1000 g of nibs were added slowly, over a period of 20 minutes, one large spoon (20 ml) at a time until a thick paste began to form with the lid then being replaced, until the paste became less viscous. Nibs were added in this way until the paste began to appear glossy and less viscous, allowing for faster addition. Once all nibs were added, the internal elements of the wet grinder were scraped with a silicone food-grade spatula, and this scraping process was repeated at 30-minute intervals three additional times, for four scrapings total. The chocolate liquor was refined to a smooth texture for a total of 8 hours from the time that all nibs were added to each batch at a temperature of approximately 50°C, maintained by production of frictional heat from the refining process itself. The final chocolate liquor was then poured through a fine-mesh culinary strainer, into a poly food-storage container using a silicone spatula to remove as much liquor as possible from the wet grinder. A yield of

approximately 900 g was achieved in all cases. The chocolate liquor was then covered with an air-tight lid, placed in a storage room at or below 19.5 °C with RH% at approximately 40% or less. All chocolate liquor batches were allowed to solidify at storage temperature, then unmolded, separated into two approximately equivalent portions—one each for chemical and sensory analysis—with each wrapped individually with an aluminum sheet, vacuum sealed in a multi-layer vacuum bag with nylon vapor barrier, and stored at or below 19.5 °C with RH% at approximately 40% or less until they were prepared for further analysis (i.e., within approximately 60 days for the chemical analysis, and within approximately 90 days for the sensory analysis).

5.2.2.4 Sensory Analysis Method

5.2.2.4.1 Chocolate Liquor Sample Sensory Preparation Method

Chocolate liquor samples were shipped on October 21st, 2019 during cool weather from Columbia, Missouri to State College, Pennsylvania. The transit took three days, and the intact package was received immediately upon delivery, the morning of October 24th, 2019 by Sensory Evaluation Center staff in the Food Science department at The Pennsylvania State University.

The samples were stored unopened at a cool room temperature below approximately 20°C for six days until October 30th, 2019, two days prior to final preparation of the sensory experiment, when the chocolate liquor samples were transferred to labeled, wide-mouth quart glass jars, sealed by covering the mouth of the jars first with aluminum foil followed by canning jar lids with rings, which were tightened to disallow

air exchange (Figure 5.2). Then, on the same day, all sealed jars were placed in a warming oven and held at 44°C with convection for 24 hours to melt, before having the temperature lowered to 39°C on Thursday, October 31, 2019 for another 24 hours to maintain the melted state of the liquor while minimizing any changes of flavor characteristics.

Figure 5. 1: Warming oven filled with jars of solidified chocolate liquor with tight lids

On Friday, 11/01/2019, the day that final sample preparation began, the work room was at 44% relative humidity (RH) and approximately 22.8°C as measured by a Bacharach

(New Kensington, PA) red-spirit sling psychrometer model #12-7012. As the viscosity of the chocolate liquor appeared somewhat high, which would make portioning consumer samples most difficult, the holding temperature of the oven was increased to 42°C in the morning of that same day. Portioning of chocolate liquor was carried out using precision positive displacement pipettes (Eppendorf Repeater M4 pipette and Eppendorf Repeater E3 pipette), with disposable pipette tips (Eppendorf Combitips Advanced Lot# i183098j), with pipettes set to 300 µL in order to obtain equal sample quantities of approximately 0.3 g in the shape of small chocolate disks, the mass of which was verified each time a new pipette tip was used (i.e., each time a new chocolate liquor sample was sampled). These 0.3 g quantities were deposited on parchment paper on a metal tray, and then chilled in a commercial refrigerator (TRUE Manufacturing, Inc (O'Fallon, MO) Refrigerator serial# 1-4596848) with thermostat set to approximately 4°C for approximately 1 hour until the chocolate liquor disks had solidified. The chocolate disks were then transferred to labeled stainless steel food storage containers covered with double-thick heavy-duty aluminum foil prior to being returned to the refrigerator to maintain the texture of the disks until testing in order to have a consistent mouthfeel and appearance (figures 5.3 and 5.4).

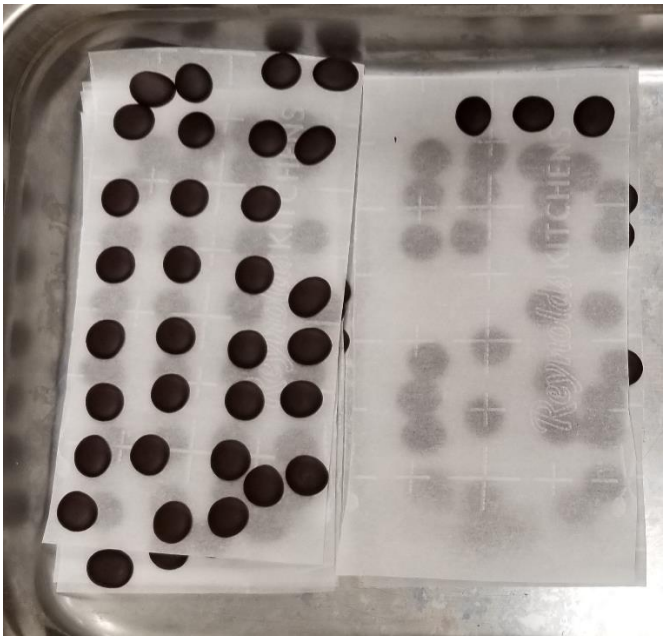

Figure 5. 2: Hotel pan filled with chocolate liquor disks

Figure 5. 3: Labeled hotel pans filled with chocolate liquor disks in a refrigerator

All 27 samples to be analyzed were prepared in this way over the course of approximately 36 hours from Friday, 11/1/2019 through 11/2/2019. As the first section of the study was fully recruited according to the experimental design (see Table 5.3 below) with 160 consumers, an excess of disks totaling approximately 200 for each of the 27 samples were prepared. After all disks were prepared for the first section of the sensory study, the warming ovens were turned off and all jars of liquor were cooled back to room temperature. Just prior to consumer analysis on the morning of Monday, November 4th, 2019, first day of the trial, approximately 8 disks of prepared chocolate liquor were transferred into each of 4 small, labeled plastic soufflé cups with lids (Figure 5.5) per sample, for all 27 samples. These cups were kept in the refrigerator until needed, and only one cup of each sample # was removed for immediate use. These cups were prepared fresh daily to avoid excessive scalping of aroma.

Figure 5. 4: Soufflé cups with pre-portioned chocolate-liquor disks for each day

The chocolate liquor sample preparation method described above was repeated identically for Section 2 of the sensory study to prepare sufficient samples to carry out the relevant experimental design in full (Table 5.4), likewise starting chocolate liquor disk preparation 3 days prior to the start of the trial, with only the total number of chocolate disks prepared being adjusted according to the design (i.e., an excess of 240 disks for each sample, as opposed to 200 for Section 1).

5.2.2.4.2 Consumer Evaluation Method

The sensory evaluation experiment took place in two sections. Section 1 consisted of a consumer evaluation study where all 27 chocolate liquor samples (i.e., all treatments across all 3 origins of cacao) were presented across 5 days without noseclips, allowing all consumers to taste and smell the samples in a way that is analogous to normal food consumption. Section 2, which took part 14 days later, required consumers to asses a subset of 9 of the 27 samples (i.e., all of the roasting treatments for the Ghanaian cacao only) across 3 days while wearing noseclips to block sensory input related to volatile aroma compounds, so that the interaction between aroma and all other sensory characteristics that were being measured could potentially be estimated for the Ghanaian samples. 145 consumers (aged 18-65, 38 males) were recruited for Section 1 (160 originally recruited minus 15 participants who withdrew prior to the first day of the trial), and 100 consumers (aged 18-65, 25 males) were recruited for section 2 (108 minus 8 participants who withdrew prior to the first day of the trial) from an in-house database of interested consumers. All consumer assessment took place in the Sensory Evaluation Center (SEC) within the Food Science Department of The Pennsylvania State

University in University Park, PA. The SEC consists of a commercial-grade food preparation area, used for sample assembly, a meeting area, used to check in consumers, and a testing room, used for product assessments. The testing room consists of 12 individual tasting booths equipped with computerized data collection via Compusense Cloud software (Academic Consortium, Guelph, ONT), and sample serving hatches through which all samples are passed to the participants. Lighting is adjustable, and red light was used for both sections of the study to mask color differences between the samples.

Consumers were scheduled for their sample evaluation in blocks of 12 at a time, 20-min apart. Each testing day, consumers evaluated 5 different samples and also underwent a brief sensory training on the first testing day for both Sections 1 and 2, prior to completing assessments. Small disks (0.3 g) of chocolate liquor (see section 5.2.5.1 above for preparation details) were served on small plastic tasting spoons on placemats showing the sample code and order of assessment (see figures 5.8 and 5.9 below). Research procedures were exempted from institutional review board review by the Penn State Office of Research Protections under the wholesome foods exemption in 45 CFR 46.101(b) (protocol number 33164). Consumer participants were compensated for their time according to the IRB protocol, and in Section 1 were given $5 incentives on days 1 and 2, $10 incentives on days 3 and 4, and $15 on day 5 for completing the entire trial. Consumers in Section 2 were given $5 incentives on days 1 and 2, and a $10 incentive on day 3 for completing the entire trial. Samples were assessed by consumers

according to the sensory evaluation designs developed for Sections 1 and 2 (see Tables 5.3 and 5.4 below).

5.2.2.4.3 Consumer Evaluation Experimental Design

5.2.2.4.3.1 Consumer Evaluation Experimental Design

Section 1

For Section 1, a Williams 27-present-27 design was created, consisting of 27 rows and 27 columns, assuring that each of the 27 samples would be presented in the design exactly 27 times, resulting in a design with positional and pairwise balance, which was important for achieving a near-balanced design for first-order carryover to the extent possible given various constraints. Because 160 consumers were initially recruited to taste 5 samples per day for each of 5 days, 4000 slots (160 x 5 x 5) were needed in the design. Since the Williams 27 design has only 729 slots (27 x 27) the design was repeated as needed to fill up the additional slots, therefore the full design was repeated approximately 5.5 times (4000 / 729), meaning that in the full 5-day, 160-consumer study, each sample would be tasted an average of 148 times (5.5 x 27) by an average of 104 consumers, with an average of 35 consumers tasting each sample two or more times. Summary statistics of this design, including the exact number of times each sample appears, and the number of consumers intended to taste each sample is shown in Table 5.3.

Table 5. 3: Consumer sensory design Section 1 summary statistics

Sample #	# Times Sample to be Tasted	# People to Taste 1 Time Only	# People to Taste 2 Times	# People to Taste 3 Times	# People to Taste 1 or more Times
1	151	65	25	12	102
2	151	69	26	10	105
3	151	64	27	11	102
4	151	63	29	10	102
5	150	68	29	8	105
6	148	72	23	10	105
7	148	76	21	10	107
8	148	77	19	11	107
9	148	75	23	9	107
10	148	73	27	7	107
11	148	77	25	7	109
12	148	80	25	6	111
13	148	80	25	6	111
14	148	75	26	7	108
15	148	70	27	8	105
16	147	69	28	6	104
17	146	70	27	6	104
18	145	69	26	8	103
19	145	66	26	9	101
20	145	71	25	8	104
21	145	71	25	8	104
22	146	68	27	8	103
23	147	68	26	9	103
24	147	62	24	11	98
25	149	59	25	12	97
26	151	59	25	14	98
27	153	58	28	13	99
Averages	**148**	**69**	**26**	**9**	**104**
Std Dev	**2**	**6**	**2**	**2**	**4**

5.2.2.4.3.2 Consumer Evaluation Experimental Design Section 2

For Section 2, a Williams 9-present-9 design was created, consisting of 9 rows and 9 columns, assuring that each of the 9 samples would be present in the design exactly 9 times, resulting in a design with positional and pairwise balance, which was important for achieving a near-balanced design for first-order carryover to the extent possible given various constraints. Because we recruited 108 consumers to taste 5 samples per day for each of 3 days, 1620 slots (108 x 5 x 3) were needed in the design. Since the Williams 9 design has only 81 slots (9 x 9) the design was repeated as needed to fill up the additional slots, therefore the full design was repeated approximately 20 times (1620 / 81), meaning that in the full 3-day, 108-consumer study, each sample would be tasted 180 times (20 x 9) by a total of 108 consumers, with 72 consumers tasting each sample twice. The exact number of times each sample appears in the design, and the number of consumers intended to taste them both appear in Table 5.4.

Table 5. 4: Consumer sensory design Section 2 summary statistics

Sample #	# Times Sample to be Tasted	# People to Taste 1 Time Only	# People to Taste 2 Times	# People to Taste 1 or more Times
1	180	36	72	108
2	180	36	72	108
3	180	36	72	108
4	180	36	72	108
5	180	36	72	108
6	180	36	72	108
7	180	36	72	108
8	180	36	72	108
9	180	36	72	108
Averages	**180**	**36**	**72**	**108**

It may be noted that the average number of people who were to taste each sample in Section 1 is 104 (Table 5.3), which is quite close to the number of people who were to taste each sample in Section 2 (i.e., 108) (Table 5.4). This was intentional, so that estimates of the effect of noseclips in Section 2 would be based upon ratings from approximately the same number of people as in Section 1 where samples were rated without noseclips.

5.2.2.4.4 Recruitment of Consumers for the Study

The SEC at The Pennsylvania State University, where the sensory portion of this study was conducted, maintains a secure electronic list of previous and potential participants for consumer and other sensory trials. This meant that many consumers receiving the recruitment email had previously been part of other consumer sensory trials, and even chocolate-related ones, and were familiar with the process of tasting samples and responding to questions using a computerized system (Compusense Cloud, Academic Consortium, Guelph, ONT). For recruitment, a screener was sent out to the SEC database to screen for eligible participants with regards to gender, age, food allergies, existence of taste defects, current medication, pregnancy or breastfeeding status, whether they smoked, had swallowing issues, or mouth piercings. Additionally, the screener included questions about chocolate preferences and frequency of consumption for several types of chocolate or chocolate-containing products. Table 5.5 shows the responses that were required for a consumer participant to be accepted into the study.

Table 5. 5: Recruitment requirements for this consumer study

Qualifying Characteristic	Requirement
Gender	Any
Age	18-65 years
Allergies	No allergies noted
Taste Defects	No taste defects
Medications	No medications taken
Pregnant	Not pregnant
Breast Feeding	Not breast feeding
Smoker	Non-smoker
Swallowing Problems	No swallowing problems
Mouth Piercings	No mouth piercings
Frequency of Milk Chocolate Consumption	Once a month or more frequently
Frequency of Dark Chocolate Consumption	Once a month or more frequently
Frequency of Chocolate Praline Consumption	Once every 3 months or more
Chocolate Preference	Either milk or dark chocolate accepted

In this way, 160 consumers were recruited for Section 1 of the study, and 108 consumers were recruited for Section 2 of the study.

5.2.2.4.5 Consumer Assessment Daily Process and Ballot

5.2.2.4.5.1 Section 1

On day one of the Section 1, 5-day study, consumers signed in at the front desk of the SEC and were given a tag with a number designating their sample set number, from 1 to 160. This number was noted in a sign-in manual and this specific code was not changed for the remainder of the 5-day trial. Once the consumer participant had received their numbered tag, they were shown into the assessment room, and picked one of the available empty individual tasting booths, with red lighting turned on. Once the

consumer had signed into the Compusense software, they were asked to turn off and put away their cellphone, to be courteous to fellow participants, and to pass their numbered tag through the hatch in front of them, whereupon they digitally provided Implied Informed Consent, and received an explanation about the number of tests in the section (i.e., one per day for 5 days), compensation amounts (i.e., consumers in section 1 were given $5 incentives on days 1 and 2, $10 incentives on days 3 and 4, and $15 on day 5 for completing the entire trial), that they would evaluate 5 dark chocolate samples per day on each of the 5 days, and that they would first receive training on rating-scale usage and attribute-rating calibration (i.e., sweet, sour, bitter, astringent, chocolate flavor). Participants were then asked to read the instructions on each page and respond to the best of their ability. The training then began.

For the training, participants were first given an explanation of how to use the Generalized Labeled Magnitude Scale (gLMS) in this study as follows:

> "The scale you will use today starts at *'no sensation'* (NS) on the left and ends at the *'strongest sensation of any kind'* on the right.
>
> Think of what the strongest sensation is for you and remember it throughout this session. The top of the scale should not change, regardless of the sensation or quality you are rating.
>
> Adjectives are placed along the scale. You should use these to help make your ratings, but feel free to click anywhere along the scale.
>
> When you are using the scale be sure to separate how *intense* something is from how much you *like or dislike* it. For example, if something is

weakly bitter and you dislike bitter, don't be tempted to rate it as more bitter because you don't like it.

Finally, remember that *the top is the strongest sensation of any kind which represents the most intense sensation you might experience across any type of sensation*. What you perceive to be the strongest sensation of any kind should be consistent across items. It is very important that the same sensation is at the top of the scale for each sensation you rate."

Next, the scale training began, with questions asked as seen in figure 5.6.

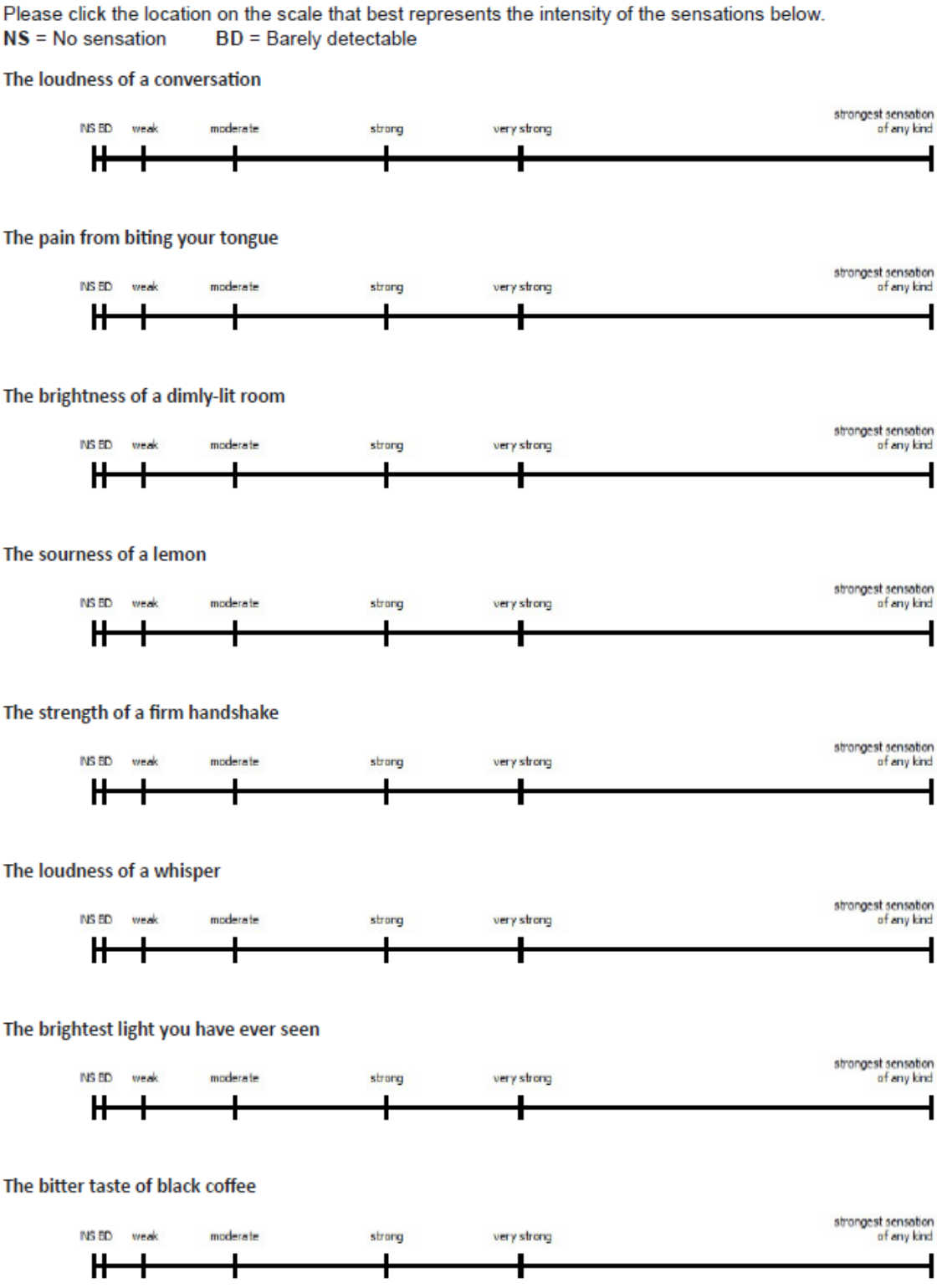

Figure 5. 5: Scale training questions with the generalized Labeled Magnitude Scale (gLMS).

Next in the training, a set of calibration samples was served to the participants to familiarize them with the attributes that they would be rating in the chocolate liquor samples. These attributes were, in order: astringent, sour, bitter, sweet, and Cocoa/Dark Chocolate. Each training solution was provided in a small, coded medicine cup (Figure 5.7), except for the cocoa/dark chocolate attribute which was described verbally only. Concentrations and descriptions of the calibration solutions are found in Table 5.6.

Figure 5. 6: Training calibration solutions

Table 5. 6: Training calibration solutions used in the sensory evaluation.

Calibration Name	Concentration and Preparation	Description
Astringent	41.5 g/L tannic acid in DI water	astringent (puckering, drying, roughing) sensations
Sour	1.5 g/L citric acid in DI water	sour sensations
Bitter	1.0 g/L caffeine in DI water	bitter sensations
Sweet	30 g/L sucrose in DI water	sweet sensations
Cocoa/Dark Chocolate	Verbal description only. No solution.	The intensity and richness of deep dark chocolate and cocoa flavors. For example, a piece of dark chocolate, or the smell of freshly baked chocolate brownies

Participants were asked to taste a solution monadically, swish it in the mouth for 5 seconds, expectorate, and consider that it characterized a particular named sensation, for example, “astringency (puckering, drying, roughing sensation),” and then they were asked to cleanse their palate with water, expectorate once more, and then move on to the next calibration sample. Each consumer panelist proceeded in this way through each characteristic until all solutions were tasted. The final part of this training was for the “cocoa/dark chocolate” attribute. Rather than tasting a sample, the participants were asked to think of the “intensity and richness of deep dark chocolate and cocoa flavors. For example, “a piece of dark chocolate, or the smell of freshly baked chocolate brownies.” At this point, the participants passed back their training samples, were reminded once more how to use the rating scale for which they had already been trained, and then they were passed a cafeteria tray containing the five coded samples that they would taste for the day’s session (figures 5.8 and 5.9)

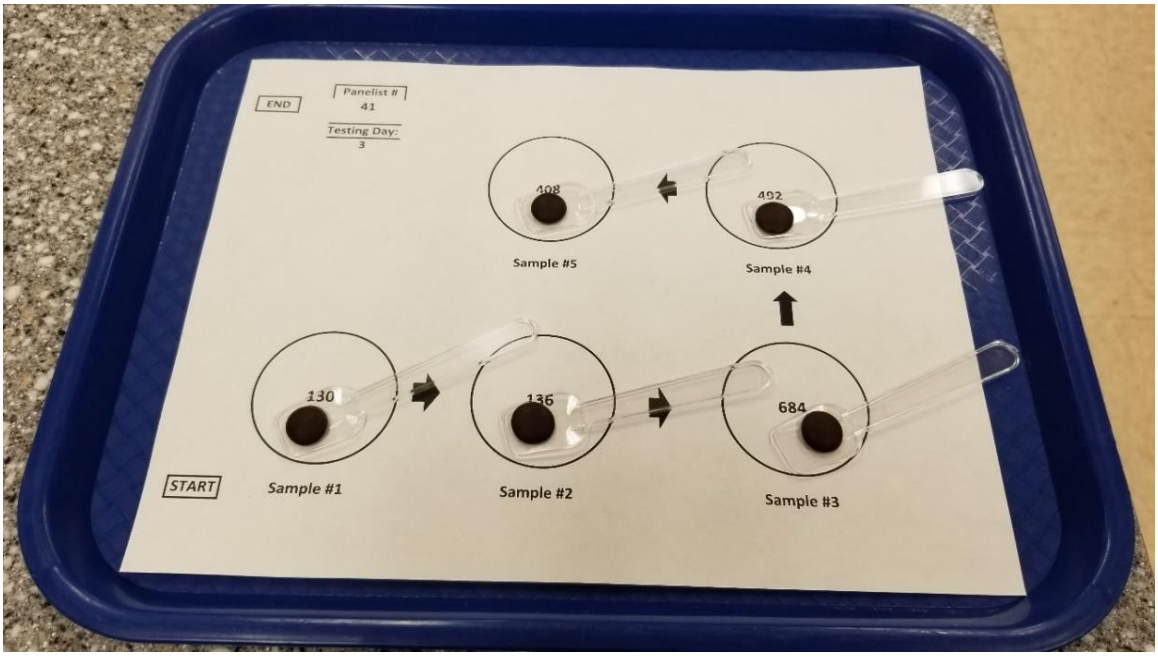

Figure 5. 7: Cafeteria tray with sample-serving placemat, plastic spoons, and samples.

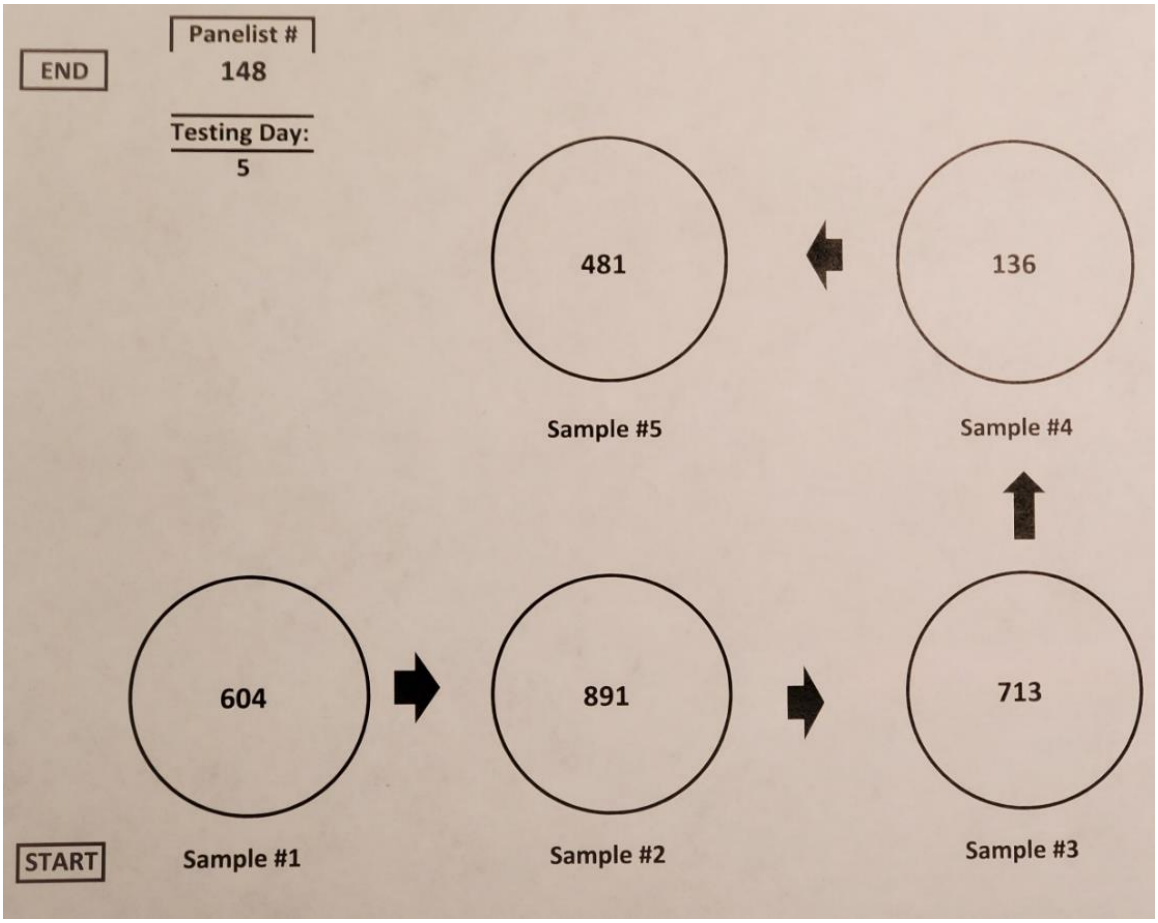

Figure 5. 8: Sample-serving placemat showing order with start and end locations, sample coding for each of the five samples, panelist #, and testing day

Once they received their samples, consumer participants were told to click “next” on the screen, and then to taste Sample #1, which had a three-digit code matching the one seen on the placemat (see example in Figure 5.9). Participants were asked to place the entire 0.3 g disk of chocolate liquor in their mouth, and to immediately rate how much they liked it, using a hedonic 9-point category scale (Figure 5.10).

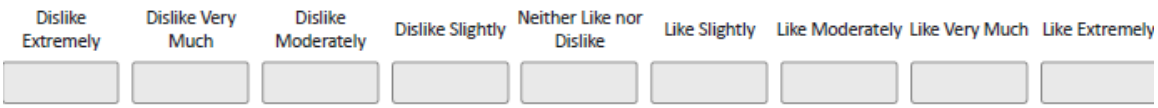

Figure 5. 9: The 9-point hedonic category scale used in this study.

Consumer participants were then immediately asked to rate, while the sample was still in their mouth, the intensities of the other characteristics (i.e., sourness, bitterness, astringency, sweetness, cocoa flavor, and other flavors and textures) (Figure 5.11). Once all attributes had been rated, participants had a mandatory 2-minute break where they were asked to rinse their mouth with room temperature water, swish, swirl and expectorate into a provided disposable polystyrene cup, before they proceeded to the remaining samples in the same fashion, one at a time, again with a 2-minute break and palate cleansing between each sample assessment. On the first tasting day of each session, after the last assessment, participants were asked about their chocolate preference (Figure 5.12), chocolate consumption frequency (Figure 5.13), gender, age, and ethnicity (Figure 5.14). For both sessions 1 and 2, participants returned for 4 or 2 more days respectively to evaluate the remaining samples in the same fashion.

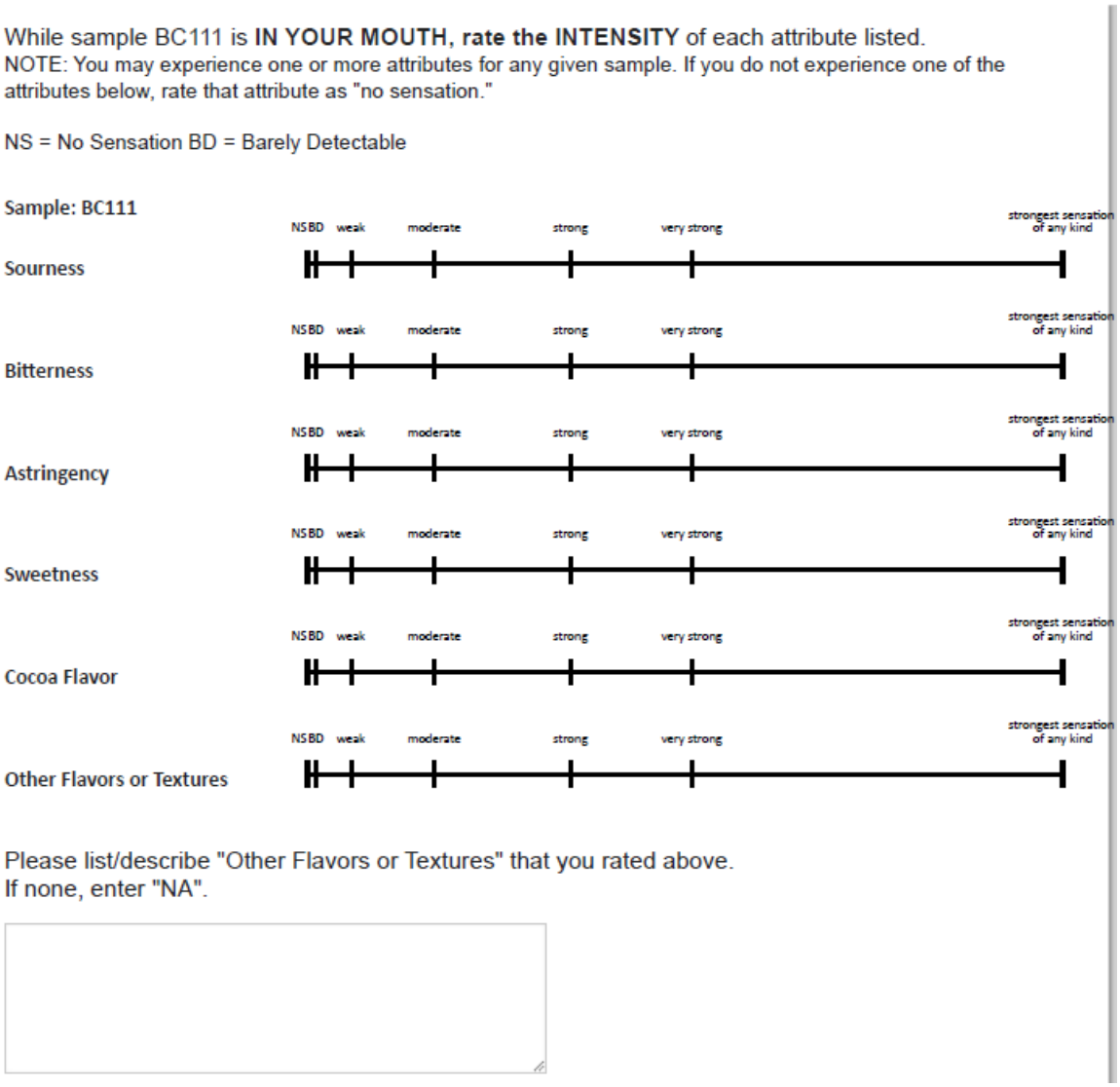

Figure 5. 10: Ballot for all chocolate liquor samples used in both sensory sessions.

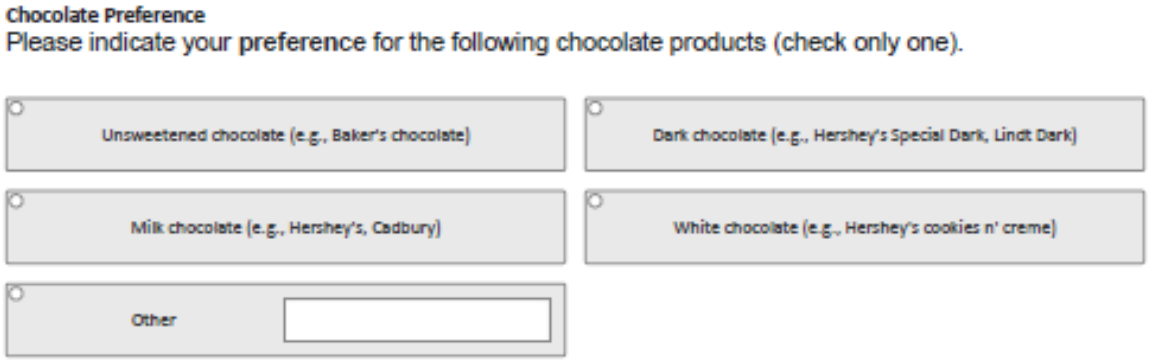

Figure 5. 11: Chocolate preference question asked on day 1 in both sessions.

Figure 5. 12: Chocolate consumption frequency questions

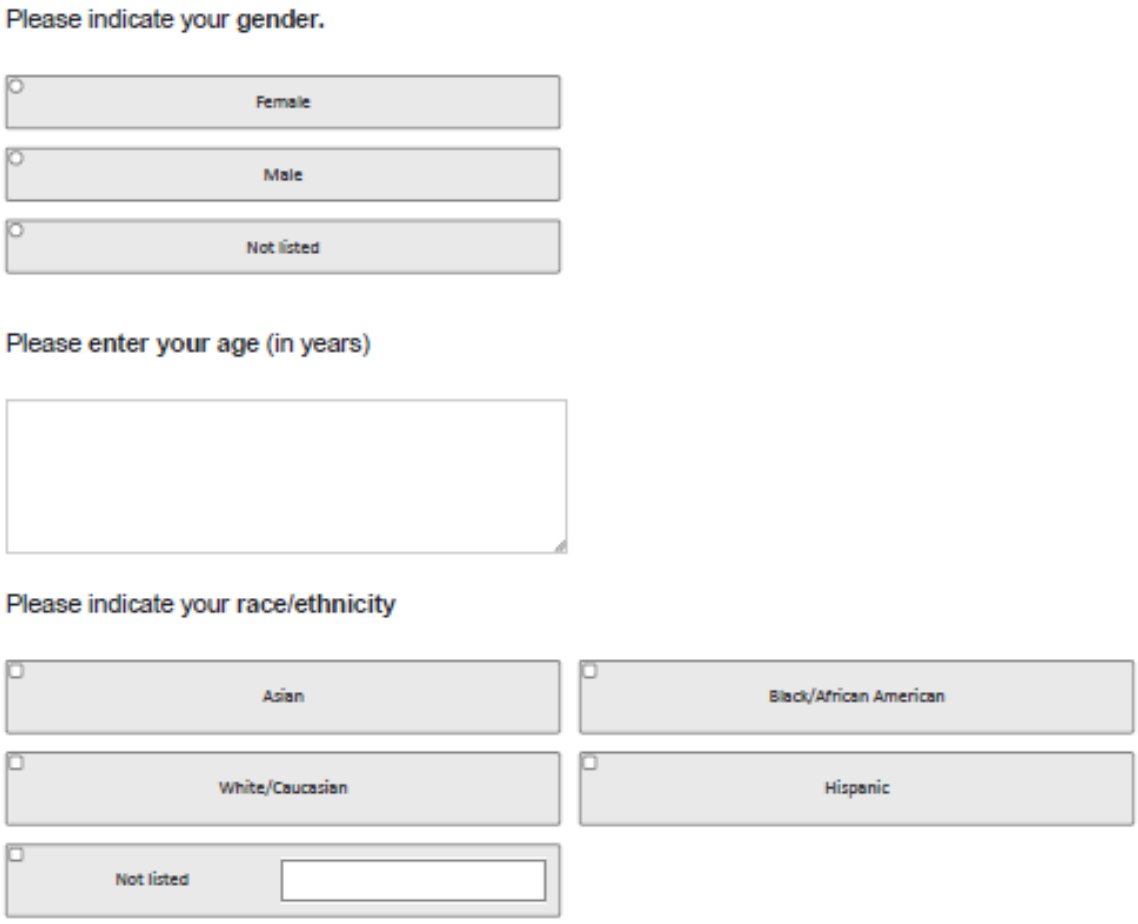

Figure 5. 13: Additional demographic information collection

5.2.2.4.5.2 Section 2 Ballot

Section 2 of the Consumer Trial was carried out exactly as in Section 1, with three differences. Firstly, in Section 2, all training and sample evaluations were carried out by participants while wearing noseclips (Figures 5.15 and 5.16). The following instructions were added regarding noseclip usage prior to any calibration or evaluation. Any bolded words were also bold in the electronic ballot:

"Before you proceed, please **put on your nose clip now**.

Put the rubber feet over your nostrils and ensure that the clip fits tightly. You should be completely unable to breathe through your nose. Please do not remove the nose clip until we instruct you to do at the end of the test!"

Proper noseclip usage can be seen in Figure 5.16. Secondly, sample evaluations occurred on only 3 days, and consumers in Section 2 therefore had a slightly different incentive structure for this reason (i.e., $5 incentives on days 1 and 2, and a $10 incentive on day 3 for completing the entire Section 2 trial). The third and final difference between Section 1 and Section 2 was that the experimental design, while based upon a Williams design, just like in Section 1, was necessarily smaller, given the smaller number of samples being assessed (i.e., 9). Summary statistics of the Section 2 experimental design are found in Table 5.4 above. Otherwise, recruitment conditions, participant training, ballot questions, serving of samples, number of samples per day, and assessment conditions such as red lighting and timing all remained the same.

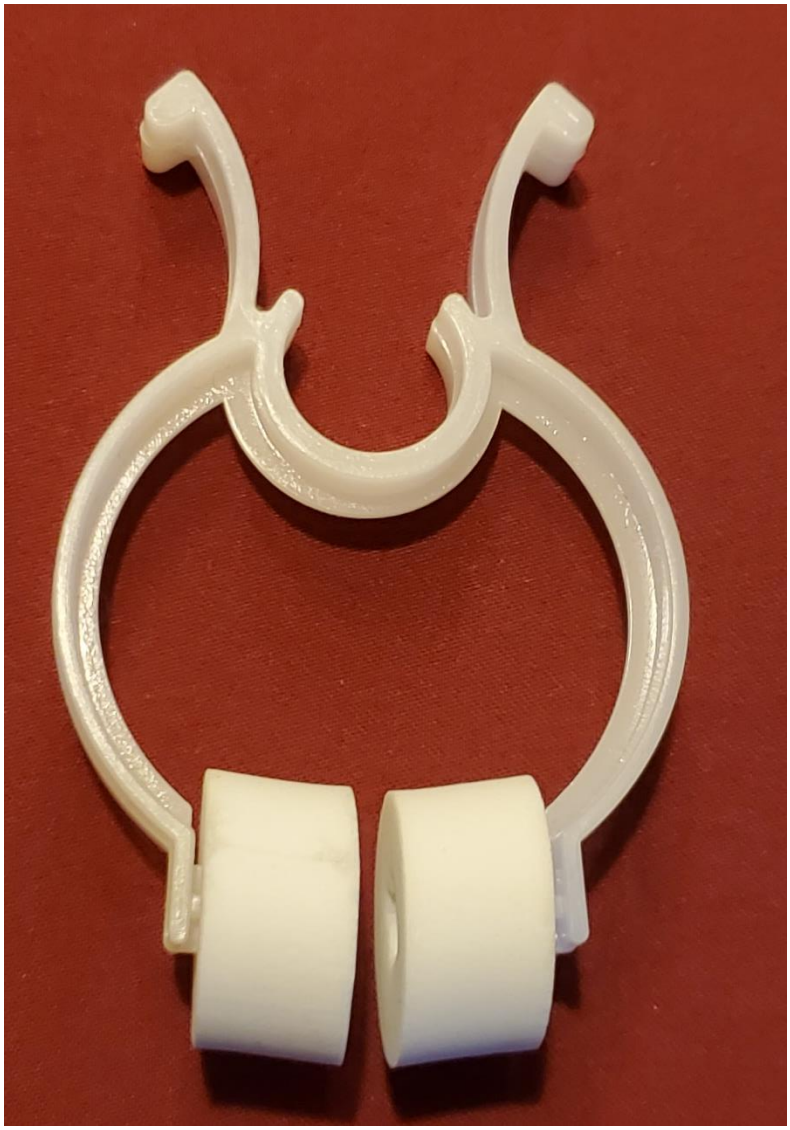

Figure 5. 14: The noseclips used in Section 2 of the consumer evaluation

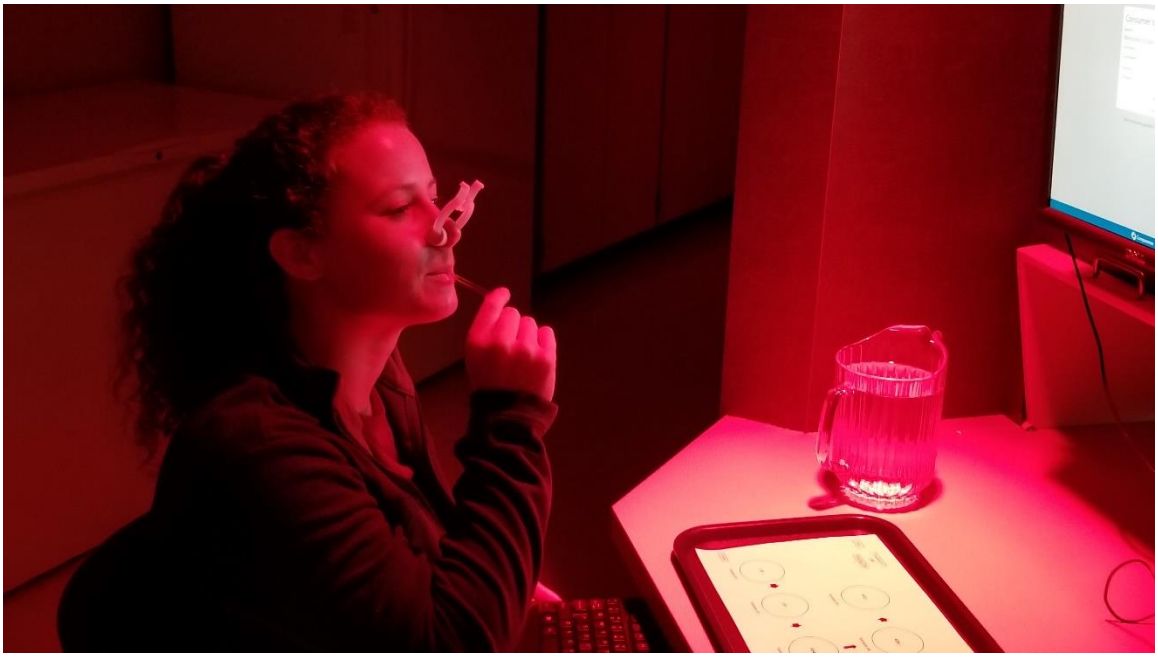

Figure 5. 15: A booth in the SEC testing room with an employee modeling a portion of the assessment with noseclips. Note the red lighting, which was used for Sections 1 and 2.

5.2.2.4.6 Statistical Analysis of Sensory Data

5.2.2.4.6.1 Method:

Principal components analysis (PCA), mixed-model linear regression, and preference mapping were carried out on consumer data. Additional details of each analysis are included in the results and discussion sections below.

5.2.2.4.6.2 Software:

Analysis of the sensory data took place in RStudio v. 1.2.1334, build 1379(f1ac3452) running R version 3.6.0 Patched (2019-06-04 r76666). Packages used include lmerTest (v. 3.1-2: Kuznetsova, Brockhoff, and Christensen, 2017) for mixed-model analysis, MuMIn (v.1.43.17: Barton, 2020) for model pseudo R^2 calculation, ggplot2(v.3.3.2:

Wickham, 2016) for PCA biplots and contour plots, and FactoMineR (v. 2.3: Le, Josse, and Husson, 2008) and SensoMineR (v.1.26: Husson, Le, and Cadoret, 2020) for preference maps and related plots.

5.3 Results & Discussion

5.3.1 Consumer Participant Summary

A summary of the number of consumer participants and other relevant details can be found in table 5.7. Additionally, 33 of the consumers who participated in Section 1, also participated in Section 2.

Table 5. 7: Summary of number of consumer participants for each section of the study

	Section 1	**Section 2**
Days of Testing	5 consecutive	3 consecutive
Noseclips Used	No	Yes
Consumers Recruited	160	108
Consumer Self-Withdraw	14	8
Consumers Dropped	1*	0
Consumers in Dataset	145	100
Males in Dataset	38	25
Incompletes in Dataset**	13	5

*The single consumer dropped from the study was a participant who was not able to accurately answer questions in a test of gLMS scale usage just after taking part in gLMS scale training.

****Incompletes refers to participants who were present for at least the first full day, but who did not remain for the entire multi-day section.

5.3.2 Consumer Ratings of Roasted Chocolate Liquors of Three Origins of Cacao Without Noseclips

5.3.2.1 Principal Components Analysis

Principal components analysis (PCA) of the data was performed, with factors consisting of all sensory-related and affective characteristics rated by consumer participants, as well as time and temperature. The PCA biplots (Figure 5.16 to 5.19) are below, and Figure 5.16, which plots data from Section 1 of the experiment with no noseclips, shows that across the ratings for all three origins, increases in Liking are most correlated with increases in Time, Temperature, and Sweetness, and to a somewhat lesser extent with increases in perceived Cocoa intensity. The relationship between Liking and Other appears orthogonal, suggesting no relationship, perhaps because participants used the Other category in a variety of unrelated ways, which is supported by the specific descriptions that people provided to describe Other, which included many items such as specific likes, texture, aroma characteristics, and much more. Finally, there is strong correlation between increases in Liking and decreases in Astringency, Sourness, and Bitterness, with Astringency appearing to be most negatively correlated with Liking. Because roasting characteristics can be said to become more pronounced as Time and Temperature in the experimental region increase, it is reasonable to surmise that consumers generally prefer some amount of roasting to unroasted cacao. Roasting is also known to result in reduced concentrations of certain bitter and astringent compounds such as the flavonoid monomer epicatechin and its oligomers (Kothe et al., 2013; Payne et al., 2010; Stanley et al., 2018), which was confirmed in chapter 4 results

for all three origins, and reduced levels of acidity (Beckett et al., 2017) and so could also explain part of the reason why Liking and perception of Bitterness, Sourness, and Astringency were negatively correlated in this experiment. Interestingly, biplots of the data from Section1 without noseclips for the individual origins (Figures 5.16 to 5.19) show very similar correlations between Liking and the other factors.

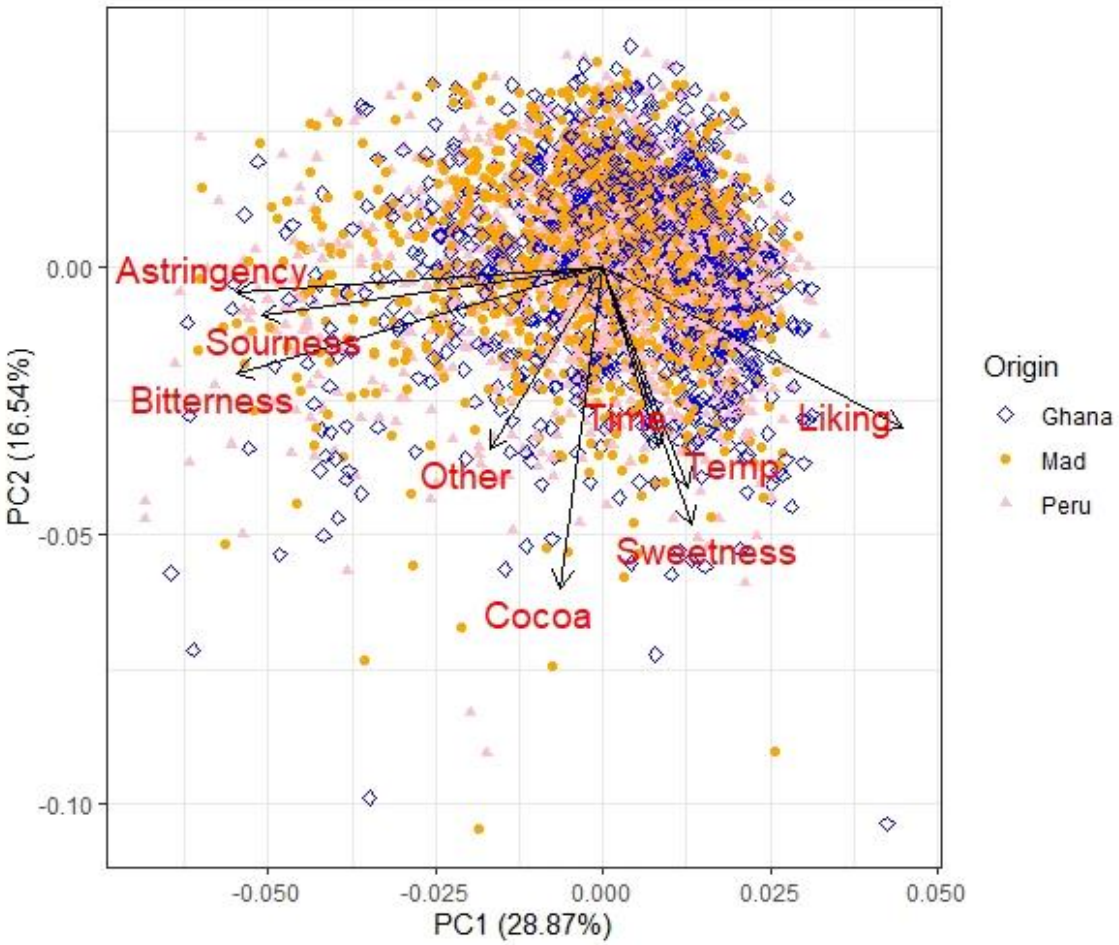

Figure 5. 16: All-Origin PCA biplot showing 45.41% of the variance in the consumer sensory data for Section 1 without noseclips, which includes factors for all rated characteristics, as well as time and temperature. Points are coded by origin.

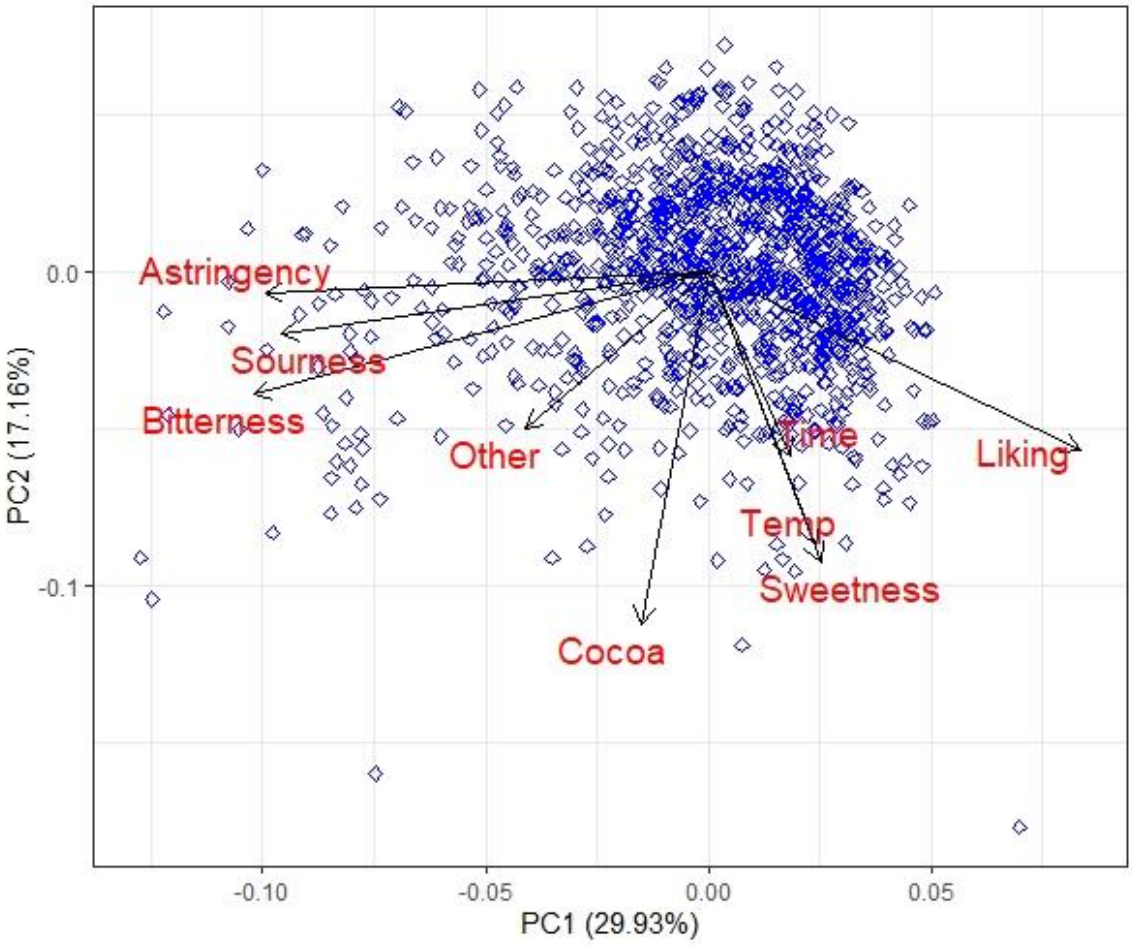

Figure 5. 17: **Ghana-Specific PCA biplot showing 47.09% of the variance in the consumer sensory data for Section 1 without noseclips, which includes factors for all rated characteristics, as well as time and temperature.**

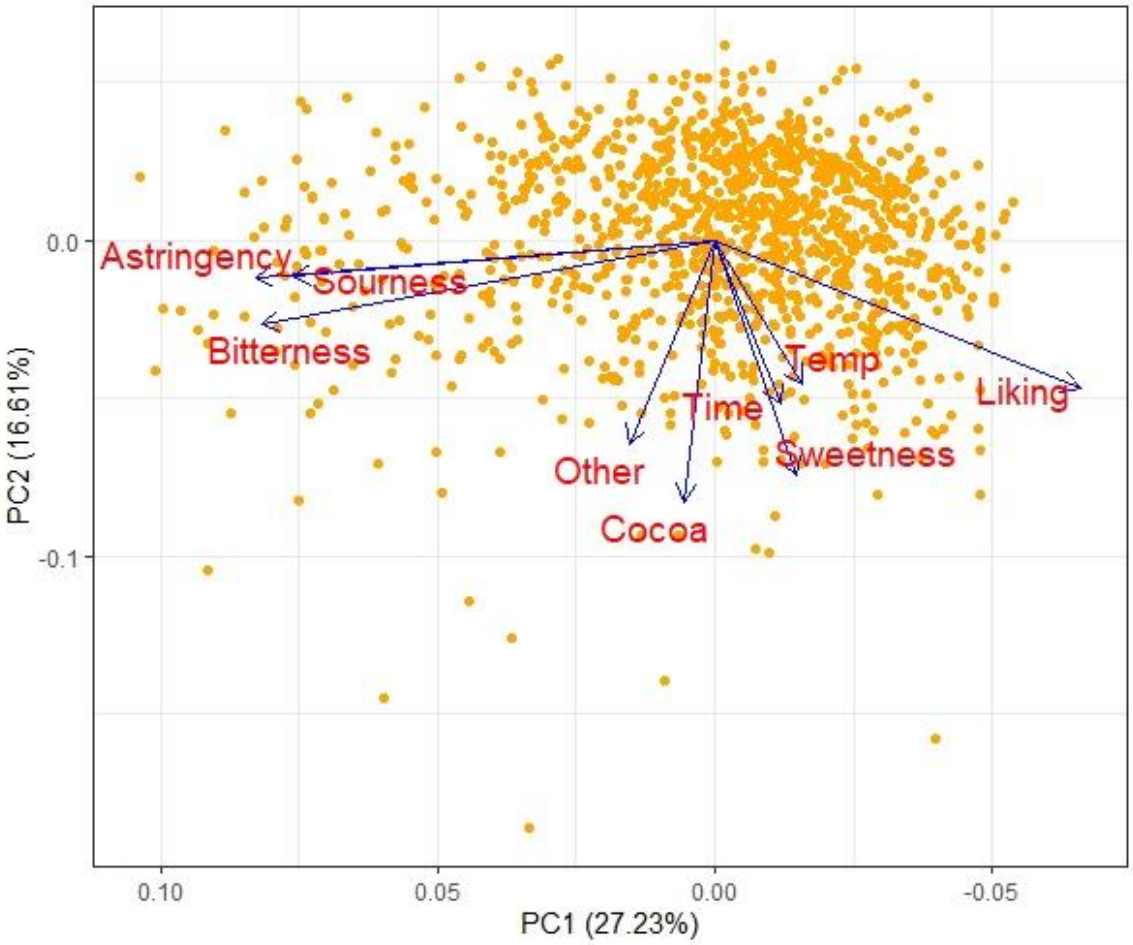

Figure 5. 18: Madagascar-Specific PCA biplot showing 43.84% of the variance in the consumer sensory data for Section 1 <u>without noseclips</u>, which includes factors for all rated characteristics, as well as time and temperature.

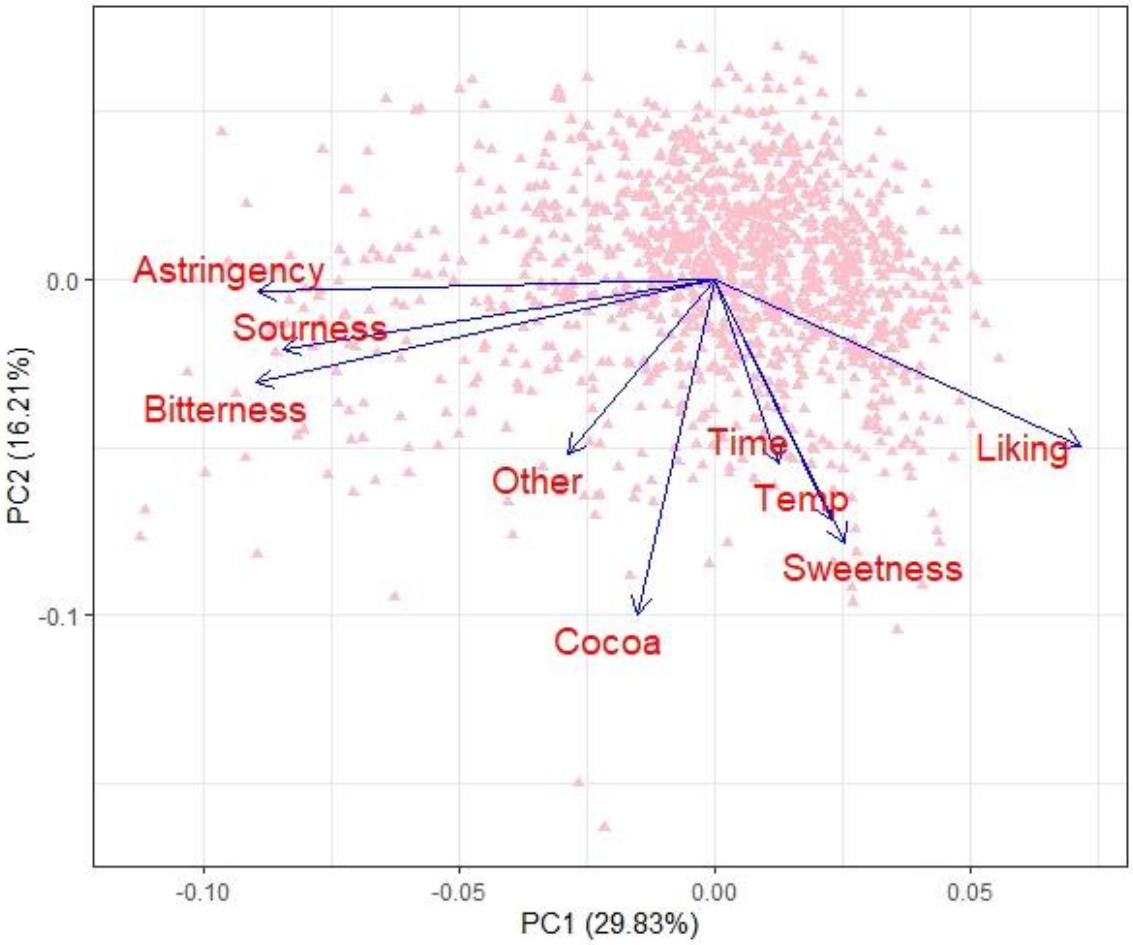

Figure 5. 19: Peru-Specific PCA biplot showing 46.04% of the variance in the consumer sensory data for Section 1 <u>without noseclips</u>, which includes factors for all rated characteristics, as well as time and temperature. **Figure 5. 20: Figure Not Available**

5.3.2.2 Mixed-Model Linear Regression

To perform mixed-model linear regression, first the predictor variables of Time and Temperature were mean-centered and scaled given that they are otherwise on quite different scales in the experimental region (i.e., minutes and degrees C respectively). Next, a mixed model was fit with each of the sensory characteristics as the response variable in turn (i.e., Bitterness, Sourness, Astringency, Sweetness, Cocoa Intensity) and

then subsequently also the rating of the affective characteristic Liking as the response variable, using the full potential model (i.e., all relevant first-order and second-order effects and their interactions, along with random effects) as follows:

$$\text{Response Variable} \sim \text{Time} + \text{Temperature} + \text{Time:Temperature} + \text{Time}^2 + \text{Temperature}^2 + \text{Temperature:Time}^2 + \text{Time:Temperature}^2 + \text{Origin} + \text{Origin:Time} + \text{Origin:Temperature} + \text{Origin:Time:Temperature} + \text{Origin:Time}^2 + \text{Origin:Temperature}^2 + \text{Origin:Temperature:Time}^2 + \text{Origin:Time:Temperature}^2 + (1|\text{OAConsumer}) + (1|\text{OAorder}) + (1|\text{Day}) + (1|\text{AgeBin}) + (1|\text{ChocPref}) + (1|\text{WDorder}) + (1|\text{OAConsumer:Day})$$

Origin is a categorical variable designating cacao from three specific geographical locations, the sampling of which is described in Materials & Methods. Terms in parentheses are random effects, and consist of the following: OAConsumer is the individual consumer coded such that each participant has a single number for the entire two-section study (e.g., OAConsumer 1 would retain that code for both Section 1 and Section 2 if s/he was present in both) , OAorder is the overall sample order across each individual section separately (all days), Day is the specific day of each Section (such as first, second, third, etc.), AgeBin is the full range of ages from 18-65 split into 5 bins and analyzed as categorical factors, ChocPref is one of six chocolate preferences chosen by each consumer as described in Materials & Methods, WDorder is within-day sample tasting order.

Model selection was carried out using backward stepwise selection with R package lmerTest (see Materials and Methods for additional information) which first selects the best random effect structure, with single-term deletions based upon p-values calculated with the likelihood ratio test. Next, the fixed-effect structure is chosen with single-term deletions relying upon Satterthwaite's method for calculating degrees of freedom and p-values, while respecting term hierarchy. Residuals of each selected model were then subjected to diagnostic plots. With this data set, residuals in selected models for each response were found to have a non-skewed distribution, but with some lack of normality in the tails, but a square-root transformation to the response substantially corrected the issue in all cases. After transformation, backward stepwise regression was yet again performed, and the residuals were once again inspected via diagnostic plots to confirm that no significant problems remained. Additionally, unless otherwise noted in a specific section below, variance inflation factor (VIF) tests showed that VIFs were less than 2.5 for all first order main effects, and generally all effects, ruling out multicollinearity as an important contributor to coefficient estimate error and term significance. Type III ANOVA for unbalanced data was then performed for each selected model to obtain estimates of p-values and coefficients for each fixed-effect term and p-values for each random-effect term in the model. Given the presence of random effects, conventional R^2 values could not be computed for the models, so pseudo R^2 values were instead computed, and all were greater than 0.5. Given the nature of the data (i.e., psychophysical data based upon sensory analysis) R^2 values over 0.25 are considered large (Cohen, 1988; Hemphill, 2003). Additionally, contour plots were

prepared for each model to help visualize the predicted values of the selected models.

All details outlined can be seen below for each response variable.

5.3.2.2.1 Bitterness

The model selected for best predicting Bitterness is:

sqrt(Bitterness) ~ Time + Temperature + Time:Temperature +

Origin + Origin:Time + Origin:Temperature + Origin:Time:Temperature +

(1 | OAConsumer) + (1 | OAorder) + (1 | Day) + (1 | OAConsumer:Day)

The summary of calculated model statistics can be seen in Tables 5.8 to 5.10.

Table 5. 8: Summary of calculated model statistics, including effect estimates

```
Fixed effects:
                  Estimate Std. Error       df t value Pr(>|t|)
(Intercept)        5.18890    0.25587    6.323  20.279 5.41e-07 ***
Time              -0.02853    0.02217 3035.405  -1.287 0.198220
Temp              -0.15290    0.02498 3070.924  -6.121 1.05e-09 ***
Origin1           -0.09477    0.03076 3103.477  -3.081 0.002079 **
Origin2           -0.01620    0.03015 3055.683  -0.537 0.591143
Time:Temp         -0.06640    0.02195 3058.582  -3.024 0.002511 **
Time:Origin1      -0.01891    0.03114 3036.081  -0.607 0.543756
Time:Origin2       0.02254    0.03228 3226.070   0.698 0.485134
Temp:Origin1      -0.12984    0.03621 3143.404  -3.586 0.000341 ***
Temp:Origin2       0.07744    0.03591 3118.558   2.156 0.031147 *
Time:Temp:Origin1 -0.11441    0.03138 3080.166  -3.645 0.000272 ***
Time:Temp:Origin2  0.05803    0.03136 3084.689   1.851 0.064336 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 9: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

                Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
Time             2.264   2.264     1 3035.4  1.6562 0.1982199
Temp            51.211  51.211     1 3070.9 37.4666 1.048e-09 ***
Origin          20.378  10.189     2 3097.9  7.4546 0.0005892 ***
Time:Temp       12.503  12.503     1 3058.6  9.1476 0.0025112 **
Time:Origin      0.781   0.391     2 3127.6  0.2857 0.7514997
Temp:Origin     17.736   8.868     2 3110.7  6.4879 0.0015424 **
Time:Temp:Origin 18.163   9.081     2 3089.1  6.6441 0.0013204 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 10: ANOVA-Like Table for Random Effects (Single-Term Deletions)

	npar	logLik	AIC	LRT	Df	Pr(>Chisq)	
<none>	17	-6060.9	12156				
(1 | OAConsumer)	16	-6352.1	12736	582.30	1	< 2.2e-16	***
(1 | OAorder)	16	-6069.1	12170	16.40	1	5.128e-05	***
(1 | Day)	16	-6078.5	12189	35.07	1	3.188e-09	***
(1 | OAConsumer:Day)	16	-6159.6	12351	197.40	1	< 2.2e-16	***

For this model, the pseudo R-squared is 0.6563748, suggesting that the combination of fixed and random effects in the model describe approximately 65.6% of the variance in the square root of Bitterness. Of the fixed-effects terms (Table 5.9), Origin and Temperature are highly significant (i.e., p-value <0.001), Time and Origin:Time are not significant (p-value >0.05), but are maintained for reasons of heredity, and all other fixed effects terms are significant (p-value <0.05). As for the random-effects terms (Table 5.10), they are all highly significant. Additionally, it can be seen in Table 5.8 that of all predictors, an increase in Temperature of one standard deviation (i.e., 43 °C degrees) has the greatest effect on a decrease in Bitterness, with an effect size on the square root of Bitterness that is approximately 5.4 times greater (-0.15290/-0.02853) than the effect of a one standard deviation increase in roasting Time (i.e., approximately 27 minutes). The Time:Temperature interaction effect is also important, with the estimated effect also being larger than that of Time alone. As for origin-based differences, the Origin effect is significant, as are the interactions of Origin:Temperature and Origin:Time:Temperature.

Response surface contour plots for each of the three origins (Figures 5.21 to 5.23) show the estimated change in the response within the experimental region due to increases in

Time and Temperature, including any interactions and quadratic effects that are included in the model. For both Peru and Madagascar, the treatment estimated to be the most bitter is raw, and Bitterness is estimated to decrease in general with increased Temperature, with slight origin-based variations. For both Madagascar and Peru, all roasted treatments have a lower estimated bitterness than the raw treatment. Ghana, however, diverges slightly from this pattern, and though increase of Temperature does result in less Bitterness in general, the interaction with Time is more important, meaning that the region near the roast treatment with the higher Temperature for the longest Time, (i.e., 135°C at 80 minutes) is estimated to result in the least Bitterness. Yet, it can still be said that the portion of the region above the horizontal line marked by approximately 100°C, which contains the five roast treatments at the highest temperatures, is predicted to result in less Bitterness than the lower portion of the plot, which contains the raw and two moderately roasted treatments.

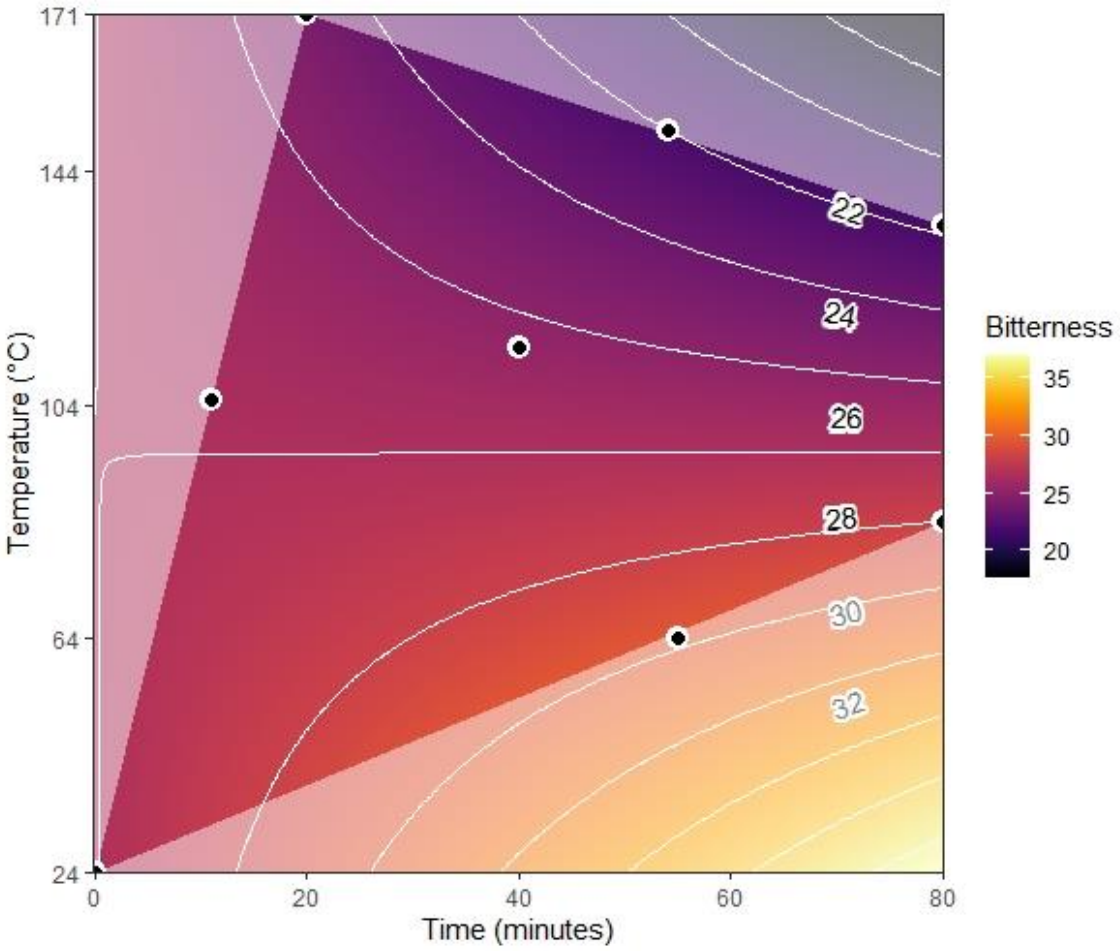

Figure 5. 21: GHANA Contour plot for predicted Bitterness for raw and roasted chocolate liquor treatments across the experimental region.

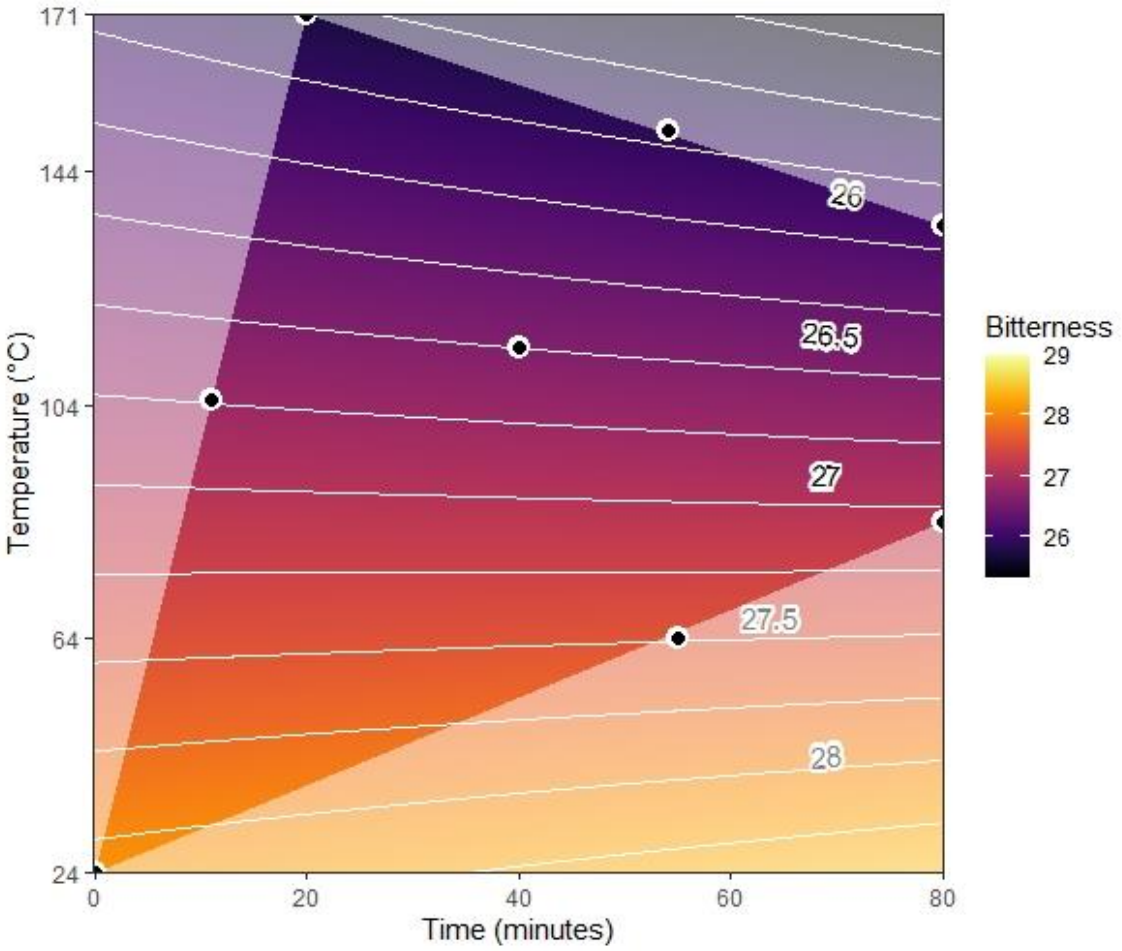

Figure 5. 22: MADAGASCAR Contour plot for predicted Bitterness for raw and roasted chocolate liquor treatments across the experimental region.

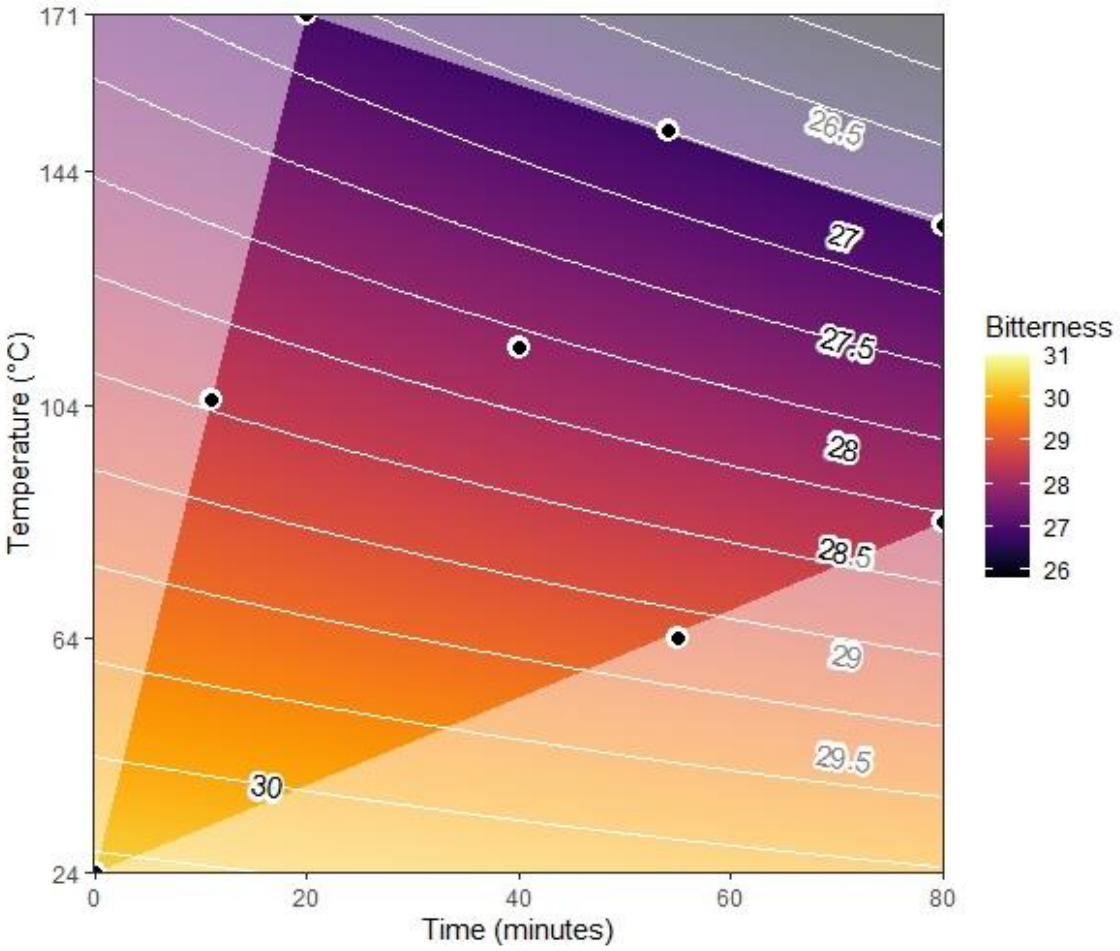

Figure 5. 23: PERU Contour plot for predicted Bitterness for raw and roasted chocolate liquor treatments across the experimental region.

5.3.2.2.2 Astringency

The model selected for best predicting Astringency is:

sqrt(Astringency)~ Time + Temperature + Time:Temperature + Temperature^2 +

Origin + Origin:Temperature + (1|OAConsumer) + (1|Day) +

(1|OAConsumer:Day) +(1|WDorder)

The summary of calculated model statistics can be seen in tables 5.11 to 5.13.

Table 5. 11: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

               Estimate Std. Error         df t value Pr(>|t|)
(Intercept)     4.77301    0.25015   11.61438  19.081 3.96e-10 ***
Time           -0.14974    0.02703 3145.99843  -5.540 3.27e-08 ***
Temp           -0.33168    0.02697 3083.60475 -12.299  < 2e-16 ***
I(Temp^2)      -0.05532    0.02507 3193.01377  -2.207 0.027399 *
Origin1        -0.23295    0.03206 3101.67809  -7.267 4.64e-13 ***
Origin2         0.07481    0.03153 3053.37938   2.373 0.017726 *
Time:Temp      -0.13509    0.02576 3089.09745  -5.244 1.67e-07 ***
Temp:Origin1   -0.04158    0.03256 3217.26510  -1.277 0.201664
Temp:Origin2    0.11428    0.03229 3125.68114   3.540 0.000407 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 12: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

             Sum Sq Mean Sq NumDF  DenDF  F value    Pr(>F)
Time         48.928  48.928     1 3146.0  30.6924 3.273e-08 ***
Temp        241.156 241.156     1 3083.6 151.2752 < 2.2e-16 ***
I(Temp^2)     7.764   7.764     1 3193.0   4.8700  0.027399 *
Origin       86.943  43.471     2 3096.3  27.2692 1.821e-12 ***
Time:Temp    43.844  43.844     1 3089.1  27.5029 1.674e-07 ***
Temp:Origin  20.562  10.281     2 3180.3   6.4492  0.001602 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 13: ANOVA-Like Table for Random Effects (Single-Term Deletions)

	npar	logLik	AIC	LRT	Df	Pr(>Chisq)	
<none>	14	-6312.3	12653				
(1 | OAConsumer)	13	-6597.1	13220	569.53	1	< 2.2e-16	***
(1 | Day)	13	-6362.2	12750	99.71	1	< 2.2e-16	***
(1 | OAConsumer:Day)	13	-6408.0	12842	191.35	1	< 2.2e-16	***
(1 | WDorder)	13	-6361.0	12748	97.32	1	< 2.2e-16	***

For this model, the pseudo R-squared is 0.6495315, suggesting that the combination of fixed and random effects in the model describe approximately 65.0% of the variance in the square root of Astringency. Of the fixed-effects terms (Table 5.12), Temperature:Origin and Temperature^2 are both significant (p-value <0.05), while all other terms are highly significant (i.e., p-value <0.001). As for the random-effects terms (Table 5.13), they are all highly significant. Additionally, it can be seen in Table 5.11 that of all predictors, an increase in Temperature of one standard deviation (i.e., 43 °C degrees) has the greatest effect on a decrease in Astringency, with an effect size on the square root of Astringency that is approximately 2.2 times greater (-0.33168/-0.14974) than the effect of a one standard deviation increase in roasting Time (i.e., approximately 27 minutes). The Time:Temperature interaction effect is also important, with the estimated effect being almost as large as that of Time itself.

Response surface contour plots for each of the three origins (Figures 5.24 to 5.26) show the predicted change in the response within the experimental region due to increases in Time and Temperature, including any interactions and quadratic effects that are

included in the model. For all three origins there is a general decrease in Astringency with Time and Temperature increase, so that the roasted treatments generally have lower Astringency than the raw treatment, and the area of predicted least Astringency is in the upper right corner of the experimental region, near the roast treatment of 151°C at 54 minutes and the roast treatment of 135°C at 80 minutes. The one exception to this general pattern is for Madagascar, where the raw treatment is estimated to be very slightly lower in Astringency than two of the moderate roasts (Figure 5.25), while still holding to the general pattern of the other five roasts being lower in Astringency than the raw treatment. It is also interesting to note that the mean Astringency of Ghana is less than that or either Peru or Madagascar, while Peru appears to be higher in Astringency than either of the other origins at the raw treatment, and yet roasting, especially as Temperature approaches the top of the experimental region, causes a drop in Astringency for Peru that leads to similar predicted values as those for Madagascar.

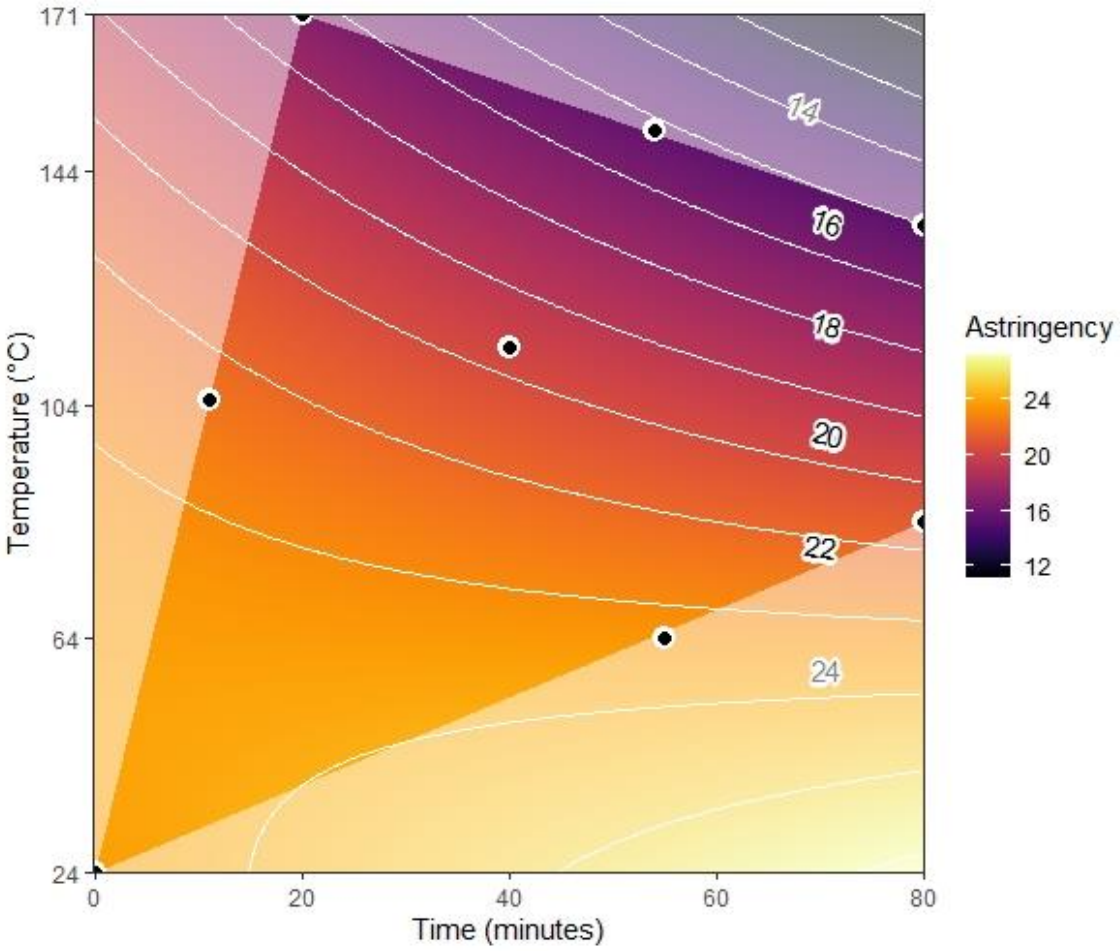

Figure 5. 24: GHANA Contour plot for predicted Astringency for raw and roasted chocolate liquor treatments across the experimental region.

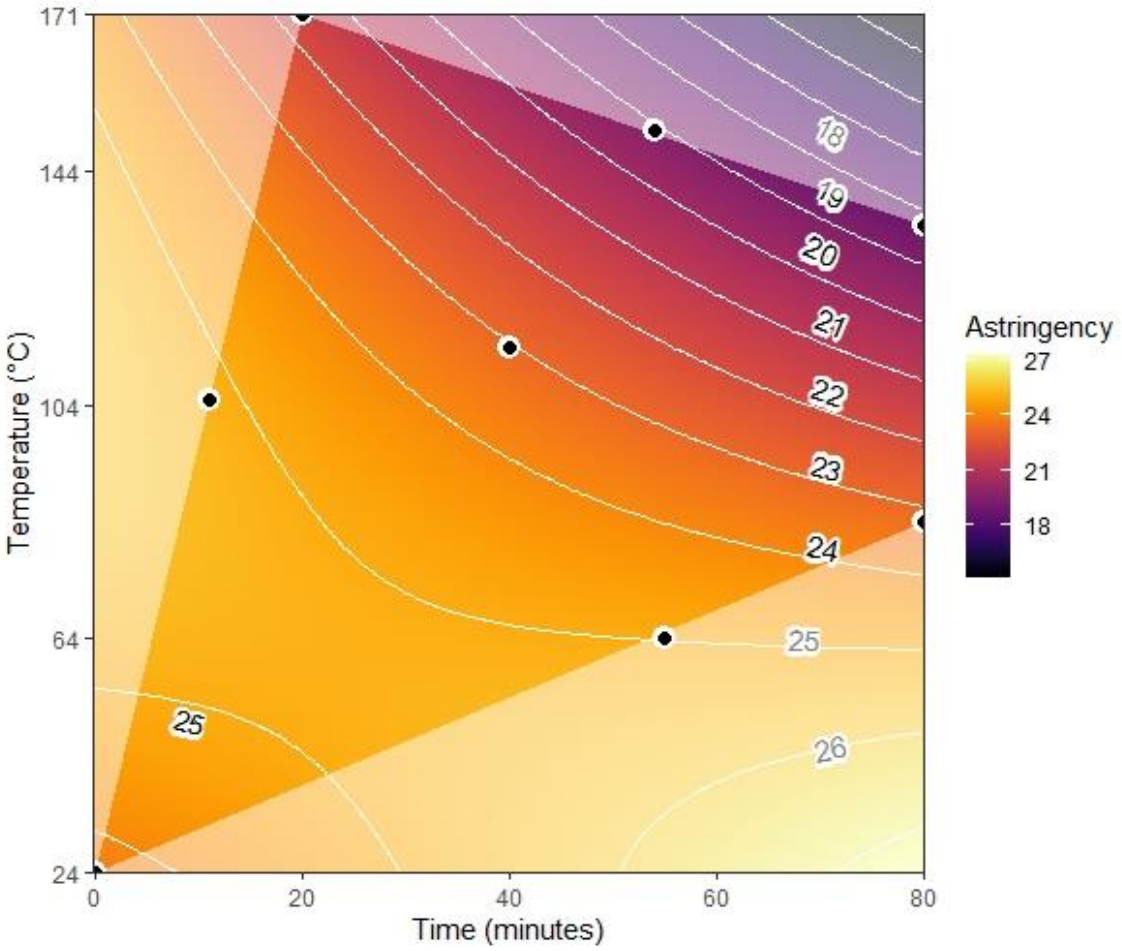

Figure 5. 25: MADAGASCAR Contour plot for predicted Astringency for raw and roasted chocolate liquor treatments across the experimental region.

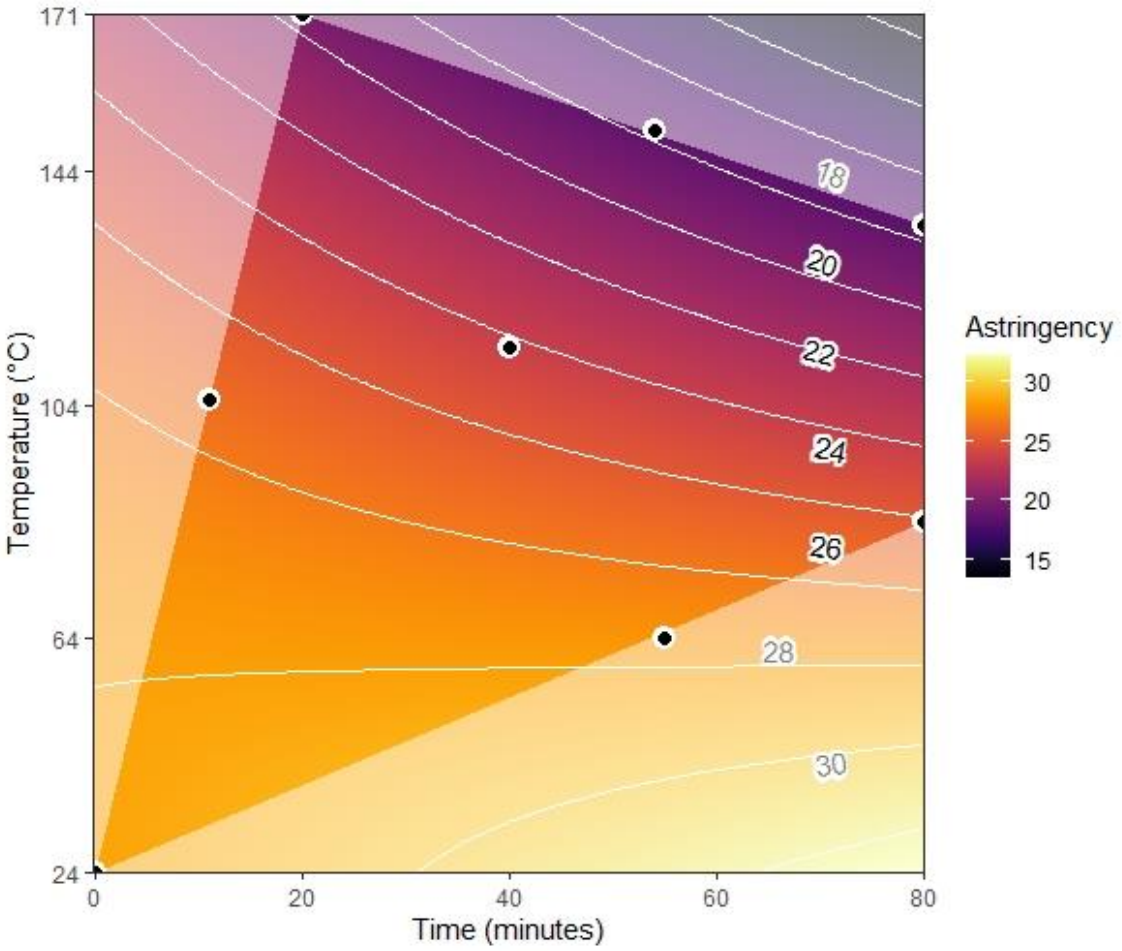

Figure 5. 26: PERU Contour plot for predicted Astringency for raw and roasted chocolate liquor treatments across the experimental region.

5.3.2.2.3 Sourness

The model selected for best predicting Sourness is:

sqrt(Sourness) ~ Time + Temperature + Time:Temperature + Temperature^2 +

Origin + (1 | OAConsumer) + (1 | Day) + (1 | OAConsumer:Day) + (1 | WDorder)

The summary of calculated model statistics can be seen in tables 5.14 to 5.16.

Table 5. 14: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

              Estimate Std. Error         df t value Pr(>|t|)
(Intercept)    4.40742    0.21610   18.78833  20.395 2.82e-14 ***
Time          -0.08194    0.02956 3203.75872  -2.772 0.005605 **
Temp          -0.25001    0.02954 3137.56182  -8.462  < 2e-16 ***
I(Temp^2)     -0.05451    0.02736 3247.00597  -1.992 0.046440 *
Origin1       -0.62361    0.03511 3157.03169 -17.763  < 2e-16 ***
Origin2        0.70831    0.03457 3097.44223  20.492  < 2e-16 ***
Time:Temp     -0.10716    0.02818 3135.24775  -3.803 0.000146 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 15: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

          Sum Sq Mean Sq NumDF  DenDF  F value    Pr(>F)
Time       14.89   14.89     1 3203.8   7.6834 0.0056054 **
Temp      138.81  138.81     1 3137.6  71.6096 < 2.2e-16 ***
I(Temp^2)   7.69    7.69     1 3247.0   3.9686 0.0464397 *
Origin    968.28  484.14     2 3148.7 249.7520 < 2.2e-16 ***
Time:Temp  28.03   28.03     1 3135.2  14.4594 0.0001459 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 16: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                      npar  logLik   AIC    LRT Df Pr(>Chisq)
<none>                  12 -6585.1 13194
(1 | OAConsumer)        11 -6898.3 13819 626.45  1  < 2.2e-16 ***
(1 | Day)               11 -6599.2 13220  28.18  1  1.106e-07 ***
(1 | OAConsumer:Day)    11 -6636.4 13295 102.57  1  < 2.2e-16 ***
(1 | WDorder)           11 -6640.8 13304 111.31  1  < 2.2e-16 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.6321368, suggesting that the combination of fixed and random effects in the model describe approximately 63.2% of the variance in the square root of Sourness. Of the fixed-effects terms (Table 5.15), Time and Temperature^2 are significant (p-value <0.05), while all other terms are highly significant (i.e., p-value <0.001). As for random-effects terms (Table 5.13), they are all highly significant. Additionally, it can be seen in Table 5.14 that of all predictors, an increase in Temperature of one standard deviation (i.e., 43 °C degrees) has the greatest effect on a decrease in Sourness, with an effect size on the square root of Sourness that is approximately 3 times greater (-0.25001/-0.08194) than the effect of a one standard deviation increase in roasting Time (i.e., approximately 27 minutes). The Time:Temperature interaction effect is also important, with the estimated effect being larger than that of Time.

As for the Origin, response surface contour plots for each of the three origins (Figures 5.27 to 5.29) show the predicted change in the response within the experimental region due to increases in Time and Temperature, including any interactions and quadratic

effects that are included in the model. Given the lack of significant interactions between Time, Temperature and Origin, the plots look substantially the same, except for differences in mean values of Sourness across the experimental region. For example, for all three origins there is a general decrease in Sourness with Time and Temperature increase, so that the roasted treatments generally have lower Sourness than the raw treatment, and the area of predicted least Sourness is in the upper right portion of the experimental region, near the roast treatment of 151°C at 54 minutes and the roast treatment of 135°C at 80 minutes. The one exception to this general pattern is that the lowest temperature roasted treatment of 64°C at 55 minutes is predicted to have very slightly more Sourness than the raw treatment for all three origins (Figure 5.25), while still holding to the general pattern of the other five roasts being lower in Sourness than the raw treatment. Though this behavior may at first appear perplexing, there is a reasonable explanation that could be confirmed with future research, even if it is perhaps of more theoretical than practical interest; as previously shown (figure 4.16), when roasting proceeds there is moisture loss due to evaporation, and this would lead to concentration of nonvolatile chemical compounds, such as certain sour organic acids known to be present in cacao (i.e., lactic acid and citric acid (Stark et al., 2006), and at certain lower roasting temperatures, even volatile acetic acid, a quite prevalent sour compound in cacao (Stark et al., 2006), could become more concentrated, made possible by a differential between the boiling points of water (100°C) and acetic acid (118°C).

It is also interesting to note that the mean Sourness of Ghana is much lower than that of the other two origins, to the extent that even the greatest Sourness in the experimental region for Ghana, is still less than the Sourness in the entire region for Madagascar, and nearly the same when compared to Peru, with only a sliver of overlap. Similarly, but in the inverse, Madagascar has an estimated Sourness that is higher throughout the whole experimental region than is the case for either Ghana or Peru.

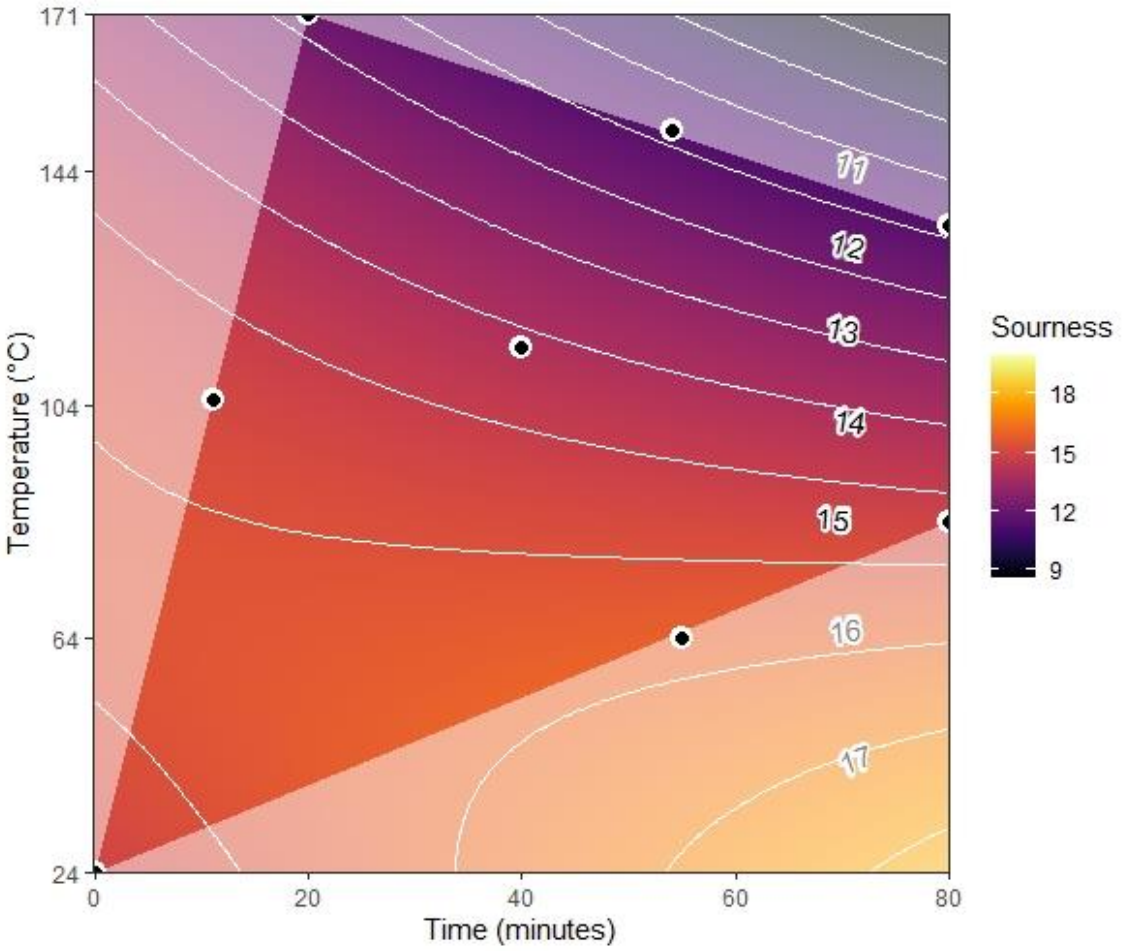

Figure 5. 27: GHANA Contour plot for predicted Sourness for roasted raw and roasted chocolate liquor treatments across the experimental region.

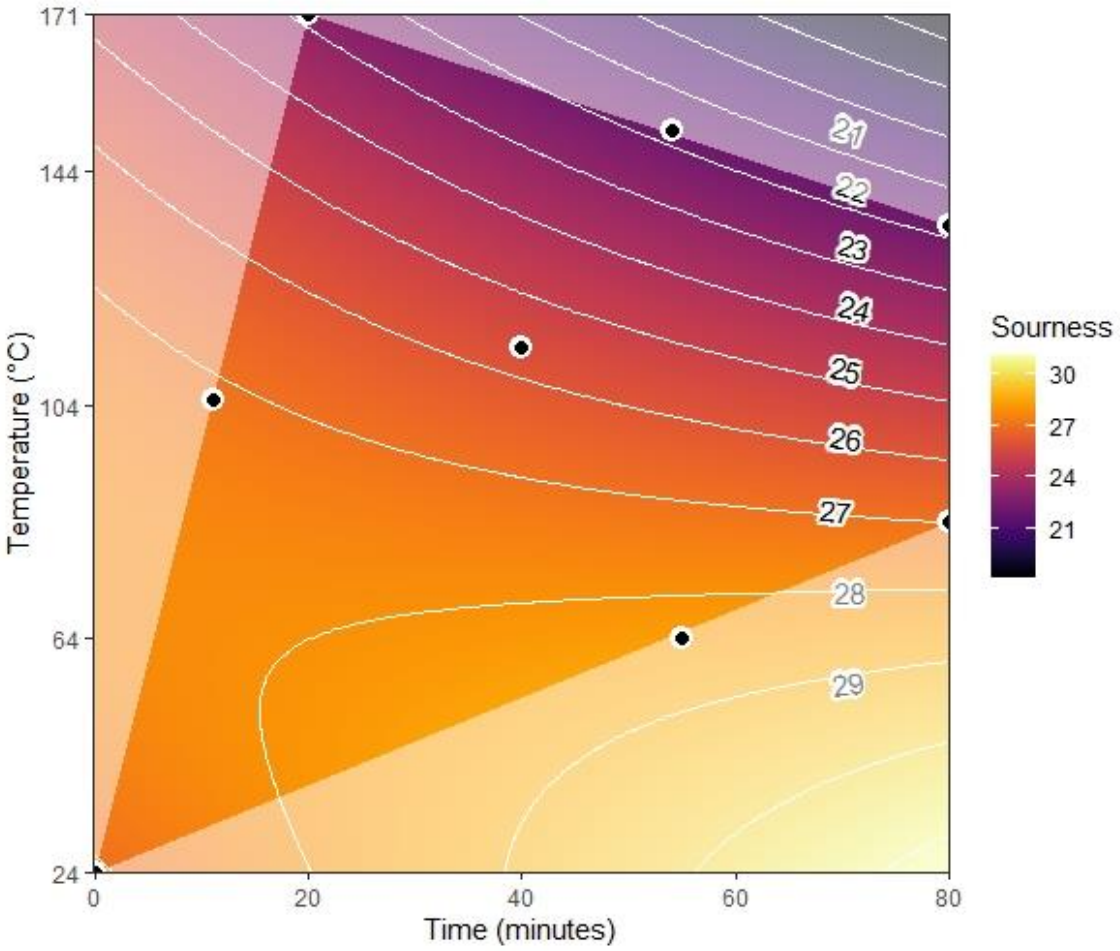

Figure 5. 28: MADAGASCAR Contour plot for predicted Sourness for raw and roasted chocolate liquor treatments across the experimental region.

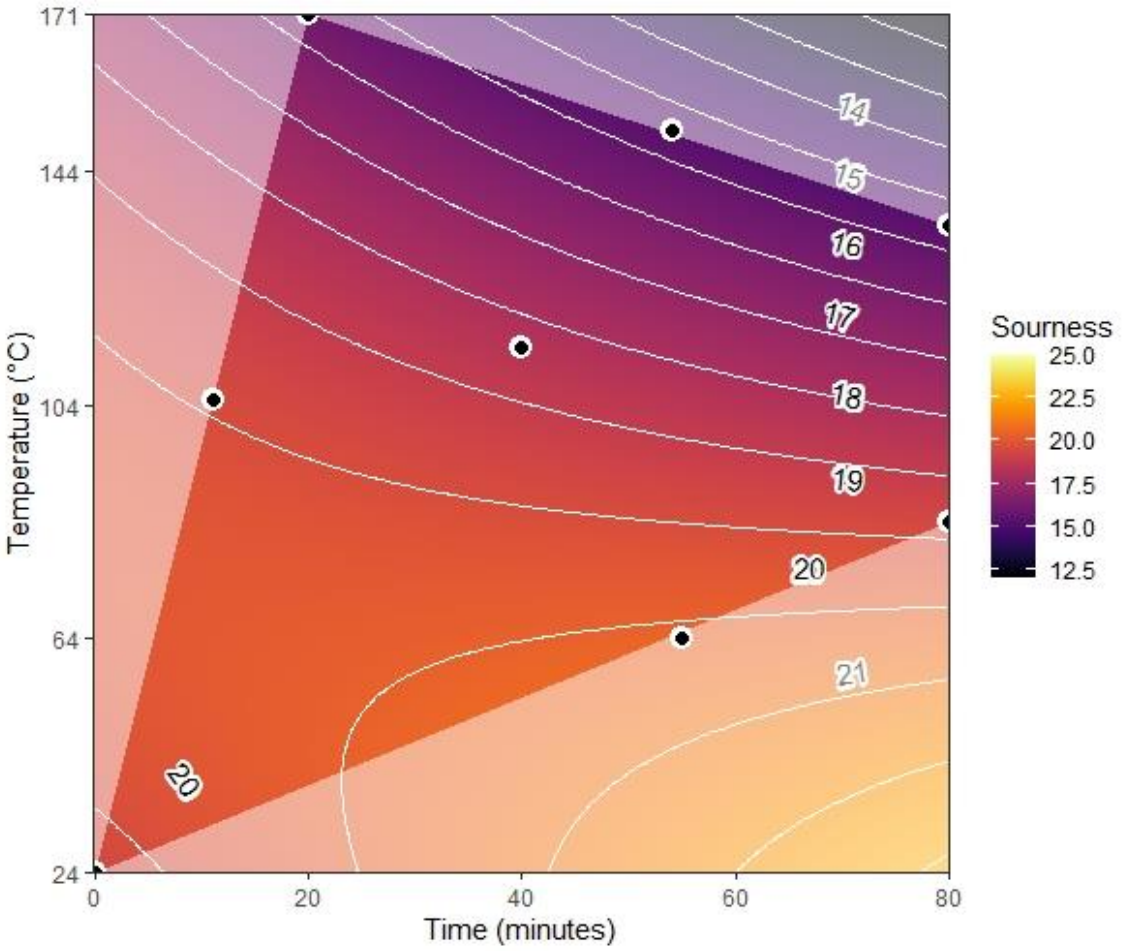

Figure 5. 29: PERU Contour plot for predicted Sourness for raw and roasted chocolate liquor treatments across the experimental region.

5.3.2.2.4 Sweetness

The model selected for best predicting Sweetness is:

sqrt(Sweetness)~ Time + Temperature + Time:Temperature + $Time^2$ +

$Temperature^2$ + Time:$Temperature^2$ + Origin + Origin:Time + Origin:Temperature

+ Origin: Time:Temperature +

(1 | OAConsumer) + (1 | ChocPref) + (1 | OAConsumer:Day) + (1 | WDorder)

The summary of calculated model statistics can be seen in tables 5.17 to 5.19.

Table 5. 17: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

                    Estimate Std. Error         df t value Pr(>|t|)
(Intercept)         1.73e+00  2.094e-01  2.180e+00   8.301 0.01092 *
Time                1.23e-01  2.731e-02  3.244e+03   4.517 6.4e-06 ***
Temp                1.79e-01  1.821e-02  3.184e+03   9.863 < 2e-16 ***
I(Time^2)          -4.68e-02  2.117e-02  3.191e+03  -2.212 0.02704 *
I(Temp^2)          -8.38e-02  3.183e-02  3.196e+03  -2.634 0.00847 **
Origin1             4.01e-02  2.051e-02  3.169e+03   1.955 0.05064 .
Origin2             1.07e-02  2.015e-02  3.123e+03   0.533 0.59378
Time:Temp           5.29e-02  1.815e-02  3.134e+03   2.919 0.00353 **
Time:I(Temp^2)     -8.95e-02  3.017e-02  3.258e+03  -2.969 0.00300 **
Time:Origin1        8.98e-03  2.079e-02  3.089e+03   0.432 0.66582
Time:Origin2       -9.60e-03  2.147e-02  3.286e+03  -0.447 0.65482
Temp:Origin1        1.04e-01  2.415e-02  3.210e+03   4.337 1.4e-05 ***
Temp:Origin2       -8.86e-02  2.400e-02  3.190e+03  -3.693 0.00022 ***
Time:Temp:Origin1   6.68e-02  2.097e-02  3.155e+03   3.189 0.00144 **
Time:Temp:Origin2  -5.35e-02  2.094e-02  3.146e+03  -2.555 0.01066 *

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 18: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

                 Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
Time             12.542  12.542     1 3243.7 20.4065 6.485e-06 ***
Temp             59.786  59.786     1 3184.3 97.2734 < 2.2e-16 ***
I(Time^2)         3.007   3.007     1 3190.9  4.8927  0.027041 *
I(Temp^2)         4.265   4.265     1 3196.4  6.9389  0.008475 **
Origin            4.123   2.062     2 3165.8  3.3545  0.035051 *
Time:Temp         5.237   5.237     1 3134.1  8.5214  0.003535 **
Time:I(Temp^2)    5.420   5.420     1 3257.6  8.8178  0.003005 **
Time:Origin       0.157   0.079     2 3182.7  0.1277  0.880095
Temp:Origin      13.303   6.652     2 3174.0 10.8224 2.069e-05 ***
Time:Temp:Origin  6.978   3.489     2 3158.1  5.6768  0.003460 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 19: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                     npar  logLik     AIC    LRT Df Pr(>Chisq)
<none>                 20 -4630.2  9300.4
(1 | OAConsumer)       19 -5006.1 10050.3 751.87  1  < 2.2e-16 ***
(1 | ChocPref)         19 -4631.7  9301.4   2.98  1    0.08445 .
(1 | OAConsumer:Day)   19 -4691.9  9421.9 123.43  1  < 2.2e-16 ***
(1 | WDorder)          19 -4645.8  9329.7  31.24  1  2.274e-08 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.6907895, suggesting that the combination of fixed and random effects in the model describe approximately 69.1% of the variance in the square root of Sweetness. Of the fixed-effects terms (Table 5.15), Time:Temperature, Time^2, Temperature^2, Time:Temperature^2 , Origin and Origin:Time:Temp are all significant (p-value <0.05), while Time, Temperature and Origin:Temperature are highly significant (i.e., p-value <0.001). Origin:Time is not significant but is maintained in the model due to heredity. As for the random-effects terms (Table 5.13), they are all highly significant, with the exception of (1|ChocPref),

which is not quite significant (i.e., >0.05 and < 0.10), but is below the default α-value used for choosing random effects values (i.e., 0.10), and the model containing it has an AIC that is approximately equivalent to the model with the lowest AIC value. In Table 5.17 it can be seen that of all predictors, an increase in Temperature of one standard deviation (i.e., 43 °C degrees) has the greatest effect on an increase in Sweetness, with an effect size on the square root of Sweetness that is approximately 1.5 times greater (0.1796/0.1234) than the effect of a one standard deviation increase in roasting Time (i.e., approximately 27 minutes).

Response surface contour plots for each of the three origins (Figures 5.30 to 5.32) show the predicted change in the response within the experimental region due to increases in Time and Temperature, including any interactions and quadratic effects that are included in the model. For all three origins there is a general increase in Sweetness as Time and Temperature increase, so that the roasted treatments generally have greater Sweetness than the raw treatment, and the area of predicted greatest Sweetness is in the upper right portion of the experimental region, near the roast treatment of 151°C at 54 minutes and the roast treatment of 135°C at 80 minutes. Additionally, the region near the raw treatment is estimated to have the least Sweetness in the entire experimental region. The one exception to this general pattern is that the region near the lowest temperature roasted treatment of 64°C at 55 minutes is predicted to have slightly less Sweetness than the raw treatment in the case of Ghana (Figure 5.25). This is not an unreasonable finding, as it was also the case that the region surrounding this

treatment for Ghana was also estimated to have the highest bitterness, and it is well known that increases in perceived bitterness can have a suppression effect on tastes such as sweetness (Lawless et al., 2010b). Still, it is important to remember that the difference between this section of the region and that near the raw treatment is quite small (<0.25 on the gLMS scale, which stretches from 0 to 100). This also means that Sweetness, in general, was not rated highly for any of the Origins, even if we were still able to successfully plot its estimated change with Time and Temperature across the experimental region. Finally, it is interesting that of all the taste modalities tested, Sweetness was predicted to be most similar between the three origins, with considerable overlap, though ultimately Ghana is estimated to have the region of greatest Sweetness of the three.

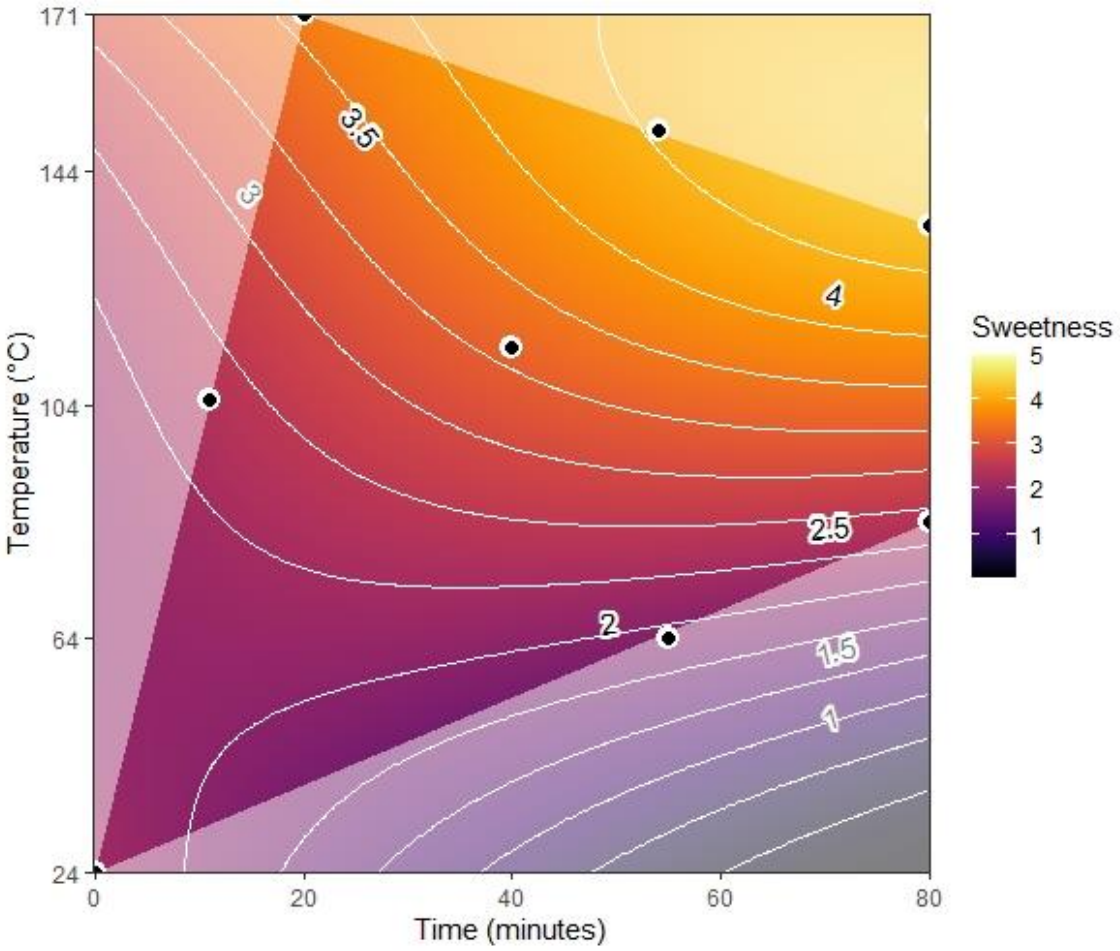

Figure 5. 30: GHANA Contour plot for predicted Sweetness for raw and roasted chocolate liquor treatments across the experimental region.

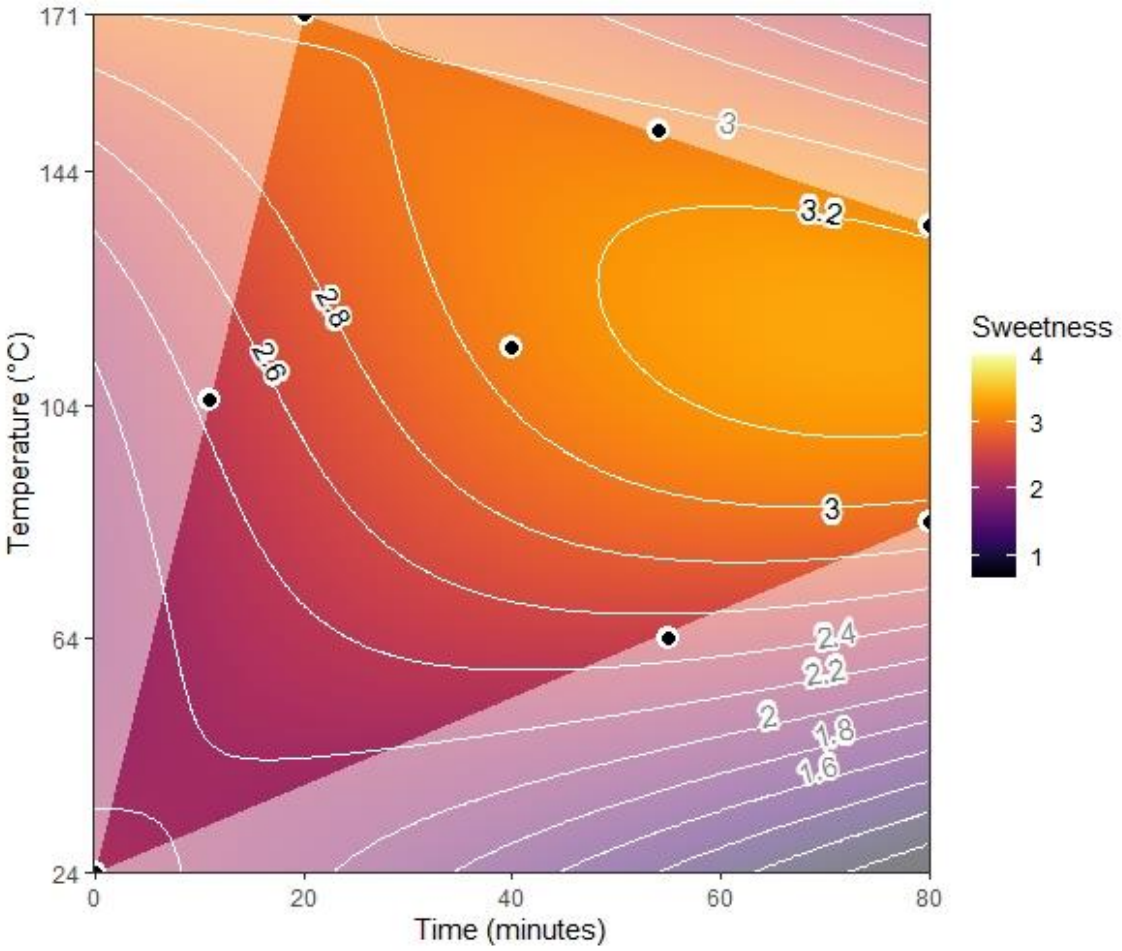

Figure 5. 31: MADAGASCAR Contour plot for predicted Sweetness for raw and roasted chocolate liquor treatments across the experimental region.

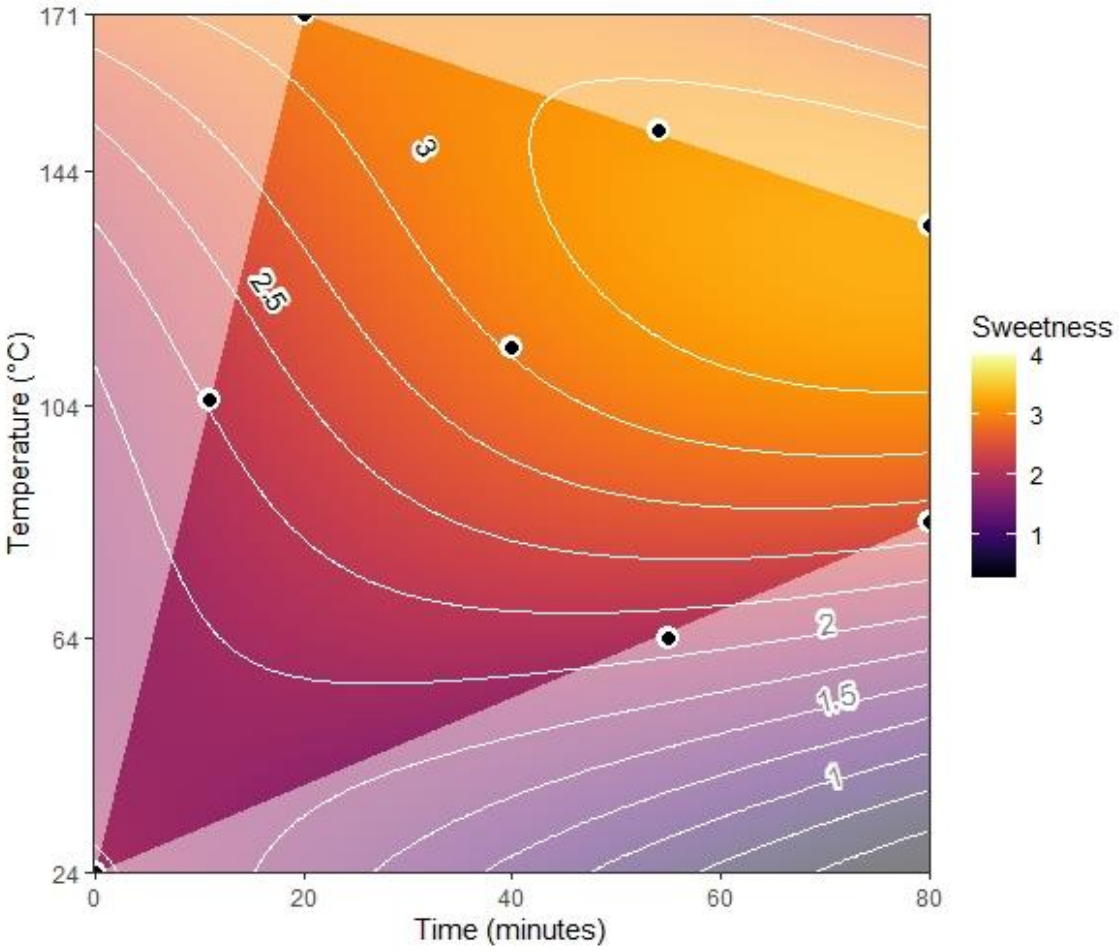

Figure 5. 32: PERU Contour plot for predicted Sweetness for raw and roasted chocolate liquor treatments across the experimental region.

5.3.2.2.5 Cocoa Flavor Intensity

Unlike the previous response surface contour plots for clearly defined taste modalities (i.e., Bitterness, Sourness, Sweetness), and perception of somatosensation (i.e., Astringency), Cocoa Flavor intensity is different in that it has no single chemical compound that can be used for a reference, making a simple calibration solution impossible. Instead, as described in Materials & Methods, participants were supplied with a written description of what Cocoa Flavor means. Consumer participants were told that Cocoa Flavor is the "richness of deep dark chocolate and cocoa flavors" that might be found in "a piece of dark chocolate, or the smell of freshly baked chocolate brownies." Still, we were able to successfully produce predictive models and response surface contour plots that should be useful for helping future researchers to understand the general perception of Cocoa by consumers as it relates to roast (Time and Temperature) and cacao Origin. As before, we used a mixed model of fixed and random effects for this purpose.

The model selected for best predicting Cocoa Flavor intensity (Cocoa) is:

$$\begin{aligned}\text{sqrt(Cocoa)}\sim\ & \text{Time} + \text{Temperature} + \text{Time:Temperature} + \text{Time}^2 + \text{Temp}^2 + \\ & \text{Temperature:Time}^2 + \text{Time:Temperature}^2 + \text{Origin} + \text{Origin:Time} + \\ & \text{Origin:Temperature} + \text{Origin:Time}^2 + \text{Origin:Temperature:Time}^2 + \\ & (1|\text{OAConsumer}) + (1|\text{Day}) + (1|\text{AgeBin}) + (1|\text{OAConsumer:Day}) + (1|\text{WDorder})\end{aligned}$$

The summary of calculated model statistics can be seen in tables 5.20 to 5.22.

Table 5. 20: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

                      Estimate  Std. Error    df t value Pr(>|t|)
(Intercept)              4.240    0.286    7.72  14.786 6.06e-07 ***
Time                     0.206    0.038 3250.79   5.293 1.28e-07 ***
Temp                     0.287    0.045 3143.62   6.324 2.90e-10 ***
I(Time^2)               -0.080    0.029 3163.42  -2.718 0.006613 **
I(Temp^2)               -0.063    0.044 3194.72  -1.406 0.159706
Origin1                  0.210    0.043 3097.95   4.844 1.33e-06 ***
Origin2                 -0.233    0.043 3250.67  -5.407 6.88e-08 ***
Time:Temp                0.114    0.025 3097.74   4.492 7.33e-06 ***
Temp:I(Time^2)           0.021    0.041 3200.05   0.526 0.599134
Time:I(Temp^2)          -0.105    0.048 3262.32  -2.163 0.030633 *
Time:Origin1             0.079    0.034 3072.66   2.286 0.022347 *
Time:Origin2            -0.081    0.035 3142.84  -2.338 0.019467 *
Temp:Origin1             0.129    0.056 3151.40   2.285 0.022357 *
Temp:Origin2            -0.154    0.055 3152.84  -2.806 0.005048 **
I(Time^2):Origin1       -0.091    0.035 3080.72  -2.557 0.010609 *
I(Time^2):Origin2        0.114    0.036 3290.83   3.114 0.001864 **
Temp:I(Time^2):Origin1  -0.093    0.039 3128.37  -2.392 0.016795 *
Temp:I(Time^2):Origin2   0.140    0.039 3195.36   3.588 0.000338 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 21: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

                      Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
Time                  32.683  32.683     1 3250.8 28.01 1.285e-07 ***
Temp                  46.665  46.665     1 3143.6 39.99 2.903e-10 ***
I(Time^2)              8.616   8.616     1 3163.4  7.38  0.006613 **
I(Temp^2)              2.308   2.308     1 3194.7  1.97  0.159706
Origin                42.064  21.032     2 3198.0 18.02 1.640e-08 ***
Time:Temp             23.537  23.537     1 3097.7 20.17 7.330e-06 ***
Temp:I(Time^2)         0.322   0.322     1 3200.1  0.27  0.599134
Time:I(Temp^2)         5.457   5.457     1 3262.3  4.67  0.030633 *
Time:Origin            8.225   4.113     2 3110.6  3.52  0.029567 *
Temp:Origin           10.182   5.091     2 3086.1  4.36  0.012813 *
I(Time^2):Origin      12.957   6.478     2 3192.1  5.55  0.003915 **
Temp:I(Time^2):Origin 15.420   7.710     2 3115.7  6.60  0.001368 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 22: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                      npar  logLik   AIC    LRT Df Pr(>Chisq)
<none>                  24 -5807.0 11662
(1 | OAConsumer)        23 -6158.0 12362 702.04  1  < 2.2e-16 ***
(1 | Day)               23 -5840.2 11726  66.28  1  3.922e-16 ***
(1 | AgeBin)            23 -5808.7 11664   3.44  1    0.06356 .
(1 | OAConsumer:Day)    23 -5899.4 11845 184.69  1  < 2.2e-16 ***
(1 | WDorder)           23 -5863.6 11773 113.08  1  < 2.2e-16 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.7180876, suggesting that the combination of fixed and random effects in the model describe approximately 71.8% of the variance in the square root of Cocoa. Variance Inflation Factors (VIFs) for first-order fixed effects are slightly higher than for other models in this work, likely due to the number of quadratic and interaction terms present, but are still well below the reasonable cutoff of 10 commonly recommended (Montgomery, Peck, & Vining, 2012; Zuur, Ieno, & Elphick, 2010) (i.e., VIFs are Time=4.2, Temperature=5.6, Origin1=3.3 and Origin2=3.4).

Of the fixed-effects terms (Table 5.21), only Temperature:Time^2 and Temperature^2 are not significant, but were both maintained in the model due to heredity. All other terms are either significant (p-value <0.05) or, in the case of Time, Temperature, Origin, and Time:Temperature, highly significant (i.e., p-value <0.001). As for the random-effects terms (Table 5.22), they are all highly significant, with the exception of (1|AgeBin), which is not quite significant (i.e., >0.05 and < 0.10), but is below the default α-value used for choosing random effects values (i.e., 0.10), and the model containing it has an AIC that is approximately equivalent to the model with the lowest AIC value. In Table

5.20 it can be seen that of all predictors, an increase in Temperature of one standard deviation (i.e., 43 °C degrees) has the greatest effect on an increase in Cocoa, with an effect size on the square root of Cocoa that is approximately 1.4 times greater (0.28726/0.20607) than the effect of a one standard deviation increase in roasting Time (i.e., approximately 27 minutes).

Response surface contour plots for each of the three origins (Figures 5.33 to 5.35) show the predicted change in the response within the experimental region due to increases in Time and Temperature, including any interactions and quadratic effects that are included in the model. For all three origins there is a general increase in Cocoa as Time and Temperature increase, so that the roasted treatments generally have greater Cocoa than the raw treatment, and the area of predicted greatest Cocoa is in the upper right portion of the experimental region, near the roast treatment of 151°C at 54 minutes and the roast treatment of 135°C at 80 minutes. Additionally, the region near the raw treatment is estimated to have the least Cocoa in the entire experimental region for all origins. Ghana, on average, is estimated to have the region of greatest Cocoa of the three origins, with the highest overall value, followed in order by Peru and then Madagascar which is estimated to have the lowest overall Cocoa across the experimental region. This is not surprising given that 70% of the cacao used worldwide is from Western Africa (Aprotosoaie et al., 2016a) meaning that consumers will be more familiar with its flavor, and so it is likely that consumer interpretation of Cocoa is based

upon significant experiences with Ghanaian or similar cocoa, as opposed to cacao from Madagascar or Peru.

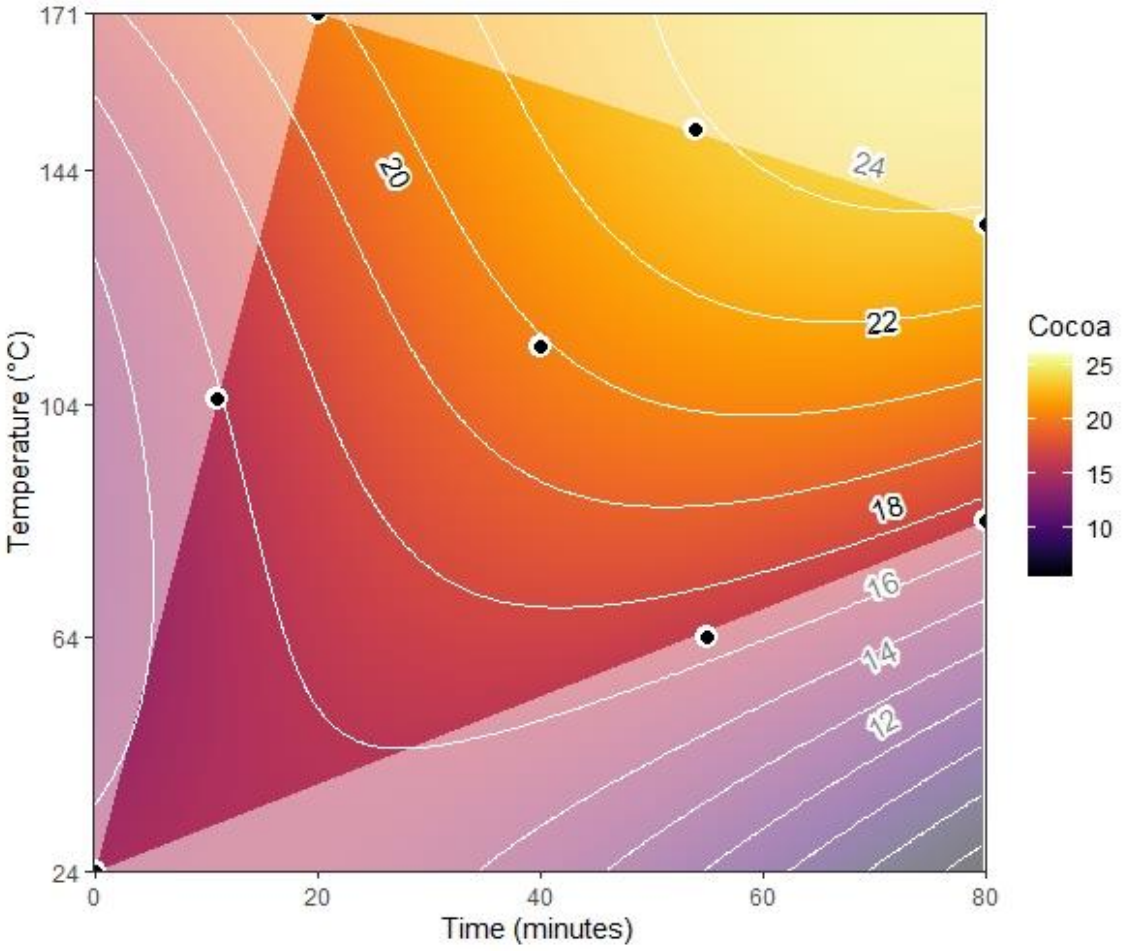

Figure 5. 33: GHANA Contour plot for predicted Cocoa Flavor intensity for raw and roasted chocolate liquor treatments across the experimental region.

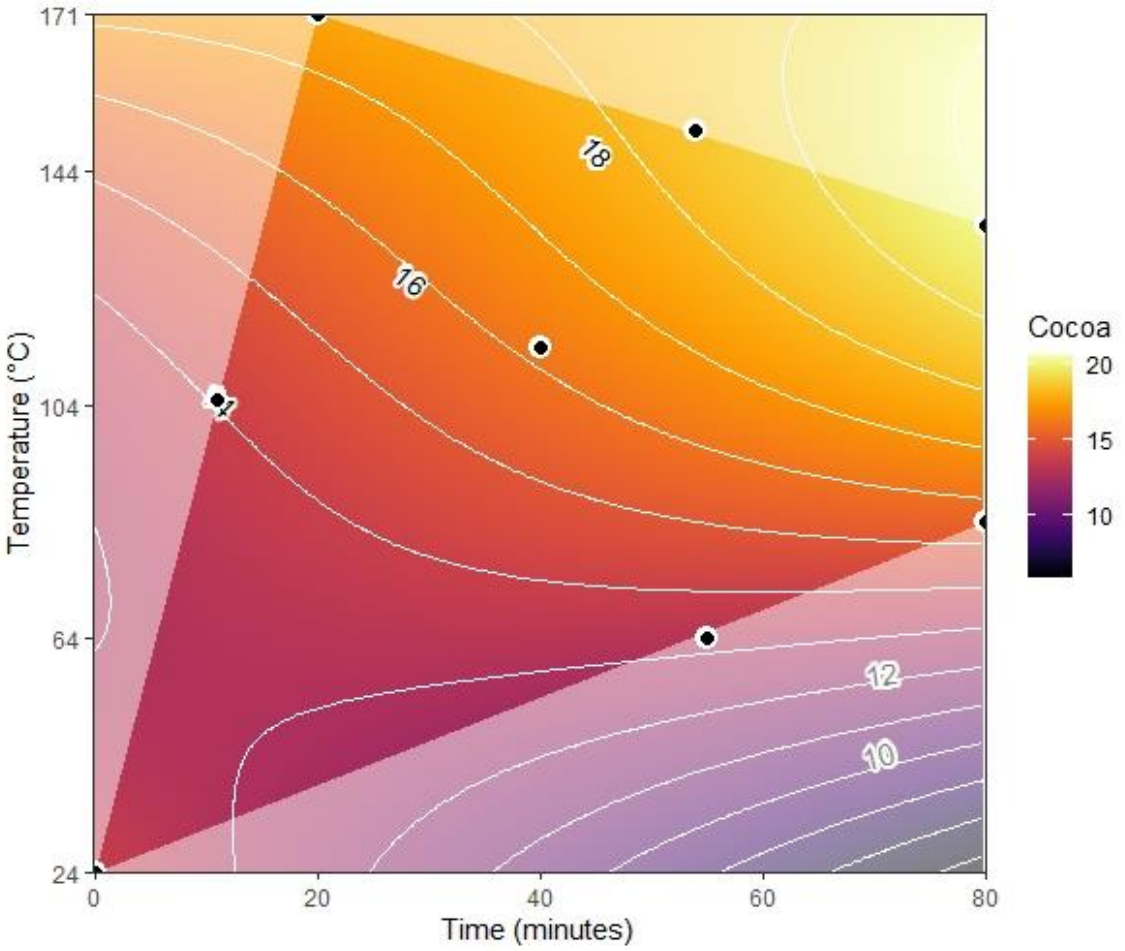

Figure 5. 34: MADAGASCAR Contour plot for predicted Cocoa Flavor intensity for raw and roasted chocolate liquor treatments across the experimental region.

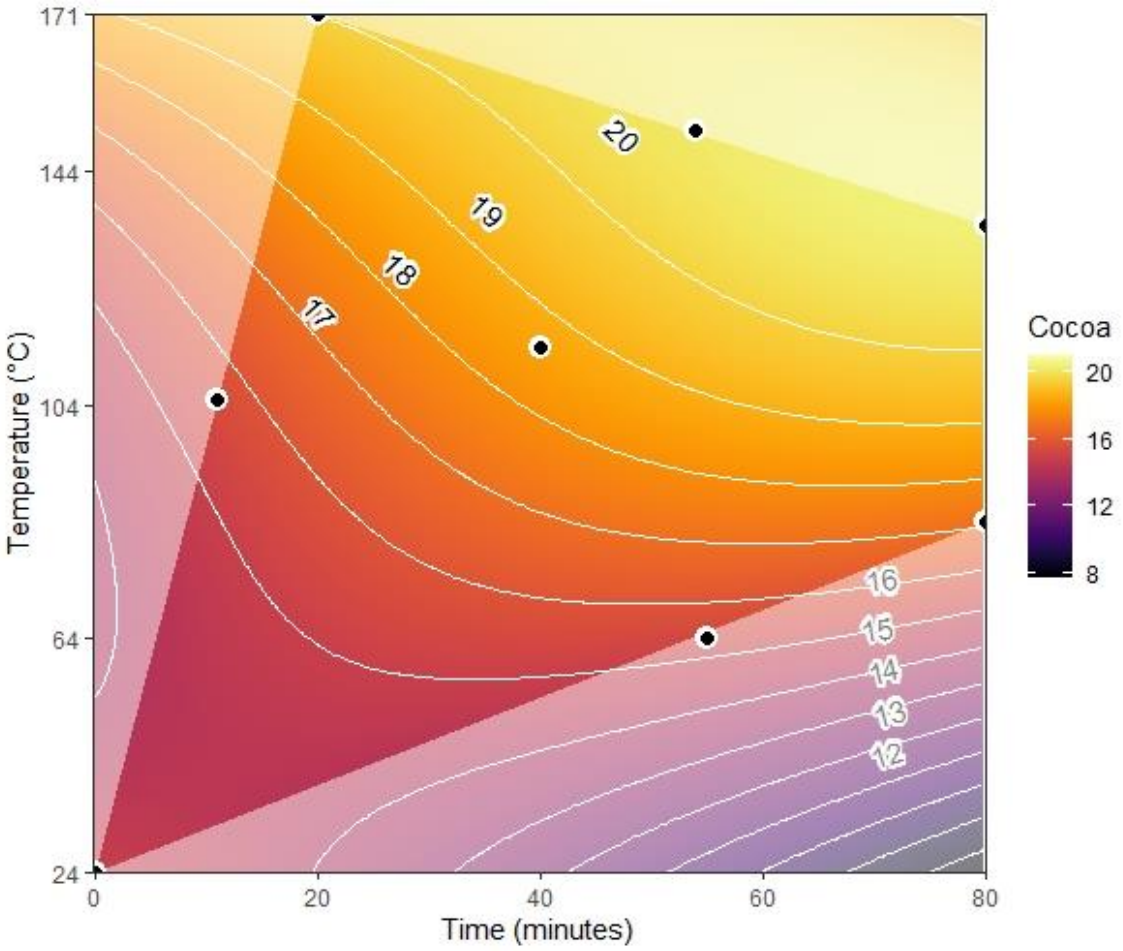

Figure 5. 35: PERU Contour plot for predicted Cocoa Flavor intensity for raw and roasted chocolate liquor treatments across the experimental region.

5.3.2.2.6 Liking

Unlike the response surface contour plots above for sensory characteristics of various types (e.g., Bitterness, Astringency, and Cocoa Intensity), Liking is different in that it is a hedonic characteristic, based not upon a single taste, aroma, or physical sensation, but upon subjective consideration of the samples, which undoubtedly is based upon previous experiences and notions about what is good and bad about chocolate generally, and perhaps dark chocolate specifically. Still, we were able to successfully produce predictive models and response surface contour plots that should be useful for helping future researchers to understand the general perception of Liking by consumers as it is related to roast (Time and Temperature) and cacao Origin. As before, we used a mixed model of fixed and random effects for this purpose. The model selected for best predicting Liking is:

sqrt(Liking)~ Time + Temperature + Time:Temperature + $Time^2$ +$Temp^2$ + Time:$Temperature^2$ + Origin + Origin:Time + Origin:Temperature + Origin:Temperature:Time + (1|OAConsumer) + (1|OAorder) + (1|Day) +(1|AgeBin) + (1|OAConsumer:Day)+(1|WDorder)

The summary of calculated model statistics can be seen in tables 5.23 to 5.25.

Table 5. 23: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

                     Estimate Std. Error       df t value Pr(>|t|)
(Intercept)         1.861e+00  7.669e-02  1.15e+01  24.27 2.84e-11 ***
Time                5.938e-02  1.141e-02  3.24e+03   5.20 2.09e-07 ***
Temp                9.046e-02  7.613e-03  3.18e+03  11.88  < 2e-16 ***
I(Time^2)          -2.459e-02  8.855e-03  3.19e+03  -2.77 0.005519 **
I(Temp^2)          -4.309e-02  1.331e-02  3.19e+03  -3.23 0.001214 **
Origin1             7.437e-02  8.578e-03  3.16e+03   8.67  < 2e-16 ***
Origin2            -6.283e-02  8.427e-03  3.12e+03  -7.45 1.15e-13 ***
Time:Temp           3.763e-02  7.588e-03  3.13e+03   4.95 7.47e-07 ***
Time:I(Temp^2)     -4.382e-02  1.260e-02  3.25e+03  -3.47 0.000514 ***
Time:Origin1        1.174e-02  8.697e-03  3.08e+03   1.35 0.177066
Time:Origin2       -4.913e-03  8.969e-03  3.28e+03  -0.54 0.583896
Temp:Origin1        4.028e-02  1.009e-02  3.21e+03   3.99 6.73e-05 ***
Temp:Origin2       -4.344e-02  1.003e-02  3.19e+03  -4.33 1.53e-05 ***
Time:Temp:Origin1   2.833e-02  8.770e-03  3.15e+03   3.23 0.001250 **
Time:Temp:Origin2  -1.787e-02  8.755e-03  3.14e+03  -2.04 0.041355 *

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 24: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

                  Sum Sq Mean Sq NumDF  DenDF  F value    Pr(>F)
Time              2.9202  2.9202     1 3246.0  27.0599 2.095e-07 ***
Temp             15.2356 15.2356     1 3185.6 141.1821 < 2.2e-16 ***
I(Time^2)         0.8322  0.8322     1 3191.5   7.7116 0.0055188 **
I(Temp^2)         1.1318  1.1318     1 3197.1  10.4877 0.0012140 **
Origin            9.5916  4.7958     2 3166.2  44.4407 < 2.2e-16 ***
Time:Temp         2.6536  2.6536     1 3133.3  24.5902 7.468e-07 ***
Time:I(Temp^2)    1.3044  1.3044     1 3259.4  12.0877 0.0005142 ***
Time:Origin       0.1998  0.0999     2 3180.6   0.9259 0.3962878
Temp:Origin       2.4727  1.2363     2 3175.1  11.4565 1.102e-05 ***
Time:Temp:Origin  1.1494  0.5747     2 3157.9   5.3256 0.0049091 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 25: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                      npar  logLik    AIC    LRT Df Pr(>Chisq)
<none>                  22 -1565.3 3174.6
(1 | OAConsumer)        21 -1857.8 3757.6 585.04  1  < 2.2e-16 ***
(1 | OAorder)           21 -1568.6 3179.2   6.62  1    0.01010 *
(1 | Day)               21 -1576.3 3194.7  22.15  1  2.526e-06 ***
(1 | AgeBin)            21 -1566.9 3175.8   3.29  1    0.06955 .
(1 | OAConsumer:Day)    21 -1617.0 3276.0 103.47  1  < 2.2e-16 ***
(1 | WDorder)           21 -1575.4 3192.7  20.20  1  6.993e-06 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.6250259, suggesting that the combination of fixed and random effects in the model describe approximately 62.5% of the variance in the square root of Liking. Variance Inflation Factors (VIFs) for first-order fixed effects are under 2.5, with the exception of Time, which has a VIF of 3.9, which is still well below the reasonable cutoff of 10 commonly recommended (Montgomery et al., 2012; Zuur et al., 2010).

Of the fixed-effects terms (Table 5.24), only Origin:Time is not significant, but this term was maintained in the model due to heredity. All other terms are either significant (p-value <0.05), or in the case of Time, Temperature, Origin, Time:Temperature, Time:Temperature^2 and Origin:Temperature, highly significant (i.e., p-value <0.001). As for the random-effects terms (Table 5.22), they are all highly significant, with the exception of (1|OAorder) which is significant (p-value <0.05), and (1|AgeBin), which is not quite significant (i.e., >0.05 and < 0.10), but is below the default α-value used for choosing random effects values (i.e., 0.10), and the model containing it has an AIC that is approximately equivalent to the model with the lowest AIC value. In Table 5.23 it can

be seen that of all predictors, an increase in Temperature of one standard deviation (i.e., 43 °C degrees) has the greatest effect on an increase in Liking, with an effect size on the square root of Liking that is approximately 1.5 times greater (0.09046/0.05938) than the effect of a one standard deviation increase in roasting Time (i.e., approximately 27 minutes).

Response surface contour plots for each of the three origins (Figures 5.36 to 5.38) show the predicted change in the response within the experimental region due to increases in Time and Temperature, including any interactions and quadratic effects that are included in the model. For all three origins there is a general increase in Liking as Time and Temperature increase, so that the roasted treatments generally have greater Liking than the raw treatment, and the area of predicted greatest Liking is in the upper right portion of the experimental region, near the roast treatment of 151°C at 54 minutes and the roast treatment of 135°C at 80 minutes. Ghana, on average, is estimated to have the region of greatest Liking of the three origins, with the highest overall value, followed in order by Peru and then Madagascar which is estimated to have the lowest overall Liking across the experimental region. This is not surprising given that 70% of the cacao used worldwide is from Western Africa (Aprotosoaie et al., 2016a) meaning that consumers will be more familiar with its flavor, and so it is likely that this familiarity with Ghanaian or similar cocoa through commercial products, as opposed to cacao from Madagascar or Peru, could influence their overall Liking.

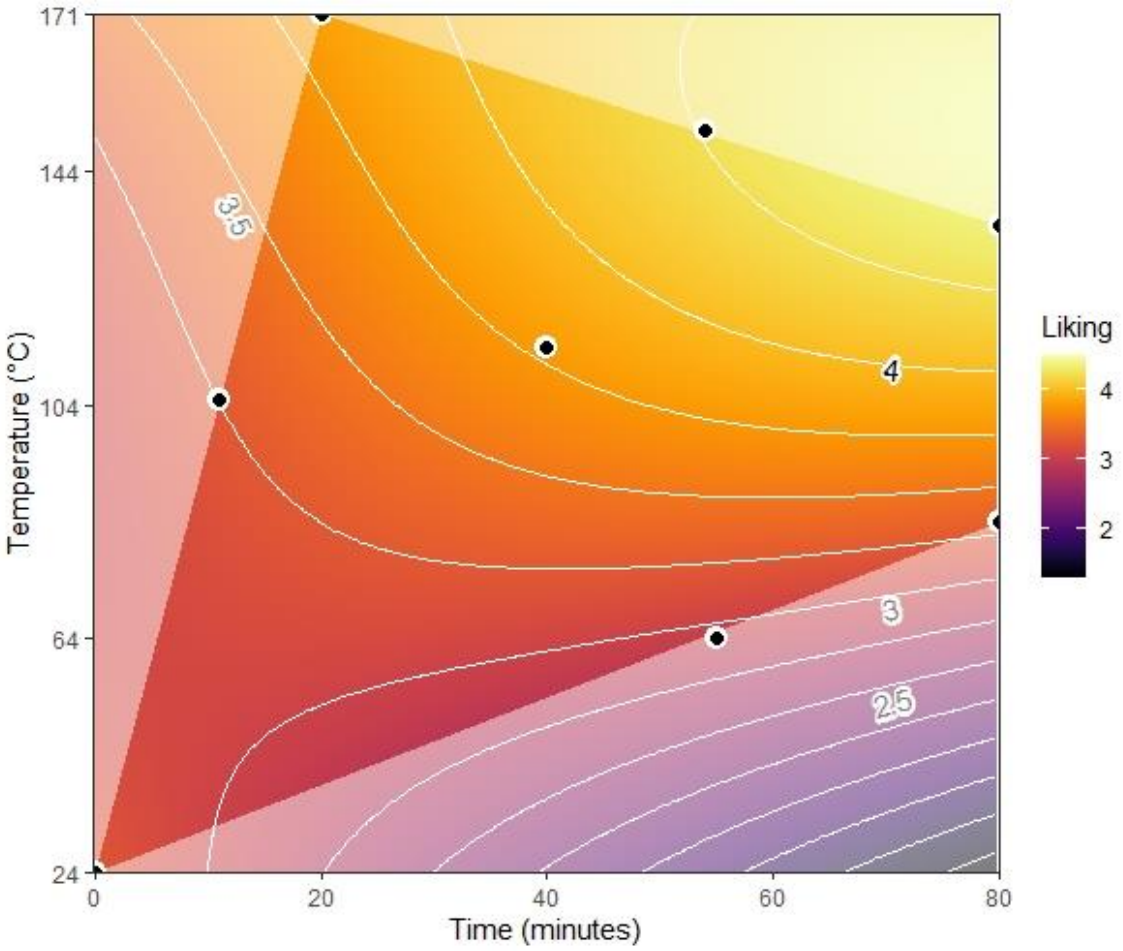

Figure 5. 36: GHANA Contour plot for predicted Liking for raw and roasted chocolate liquor treatments across the experimental region.

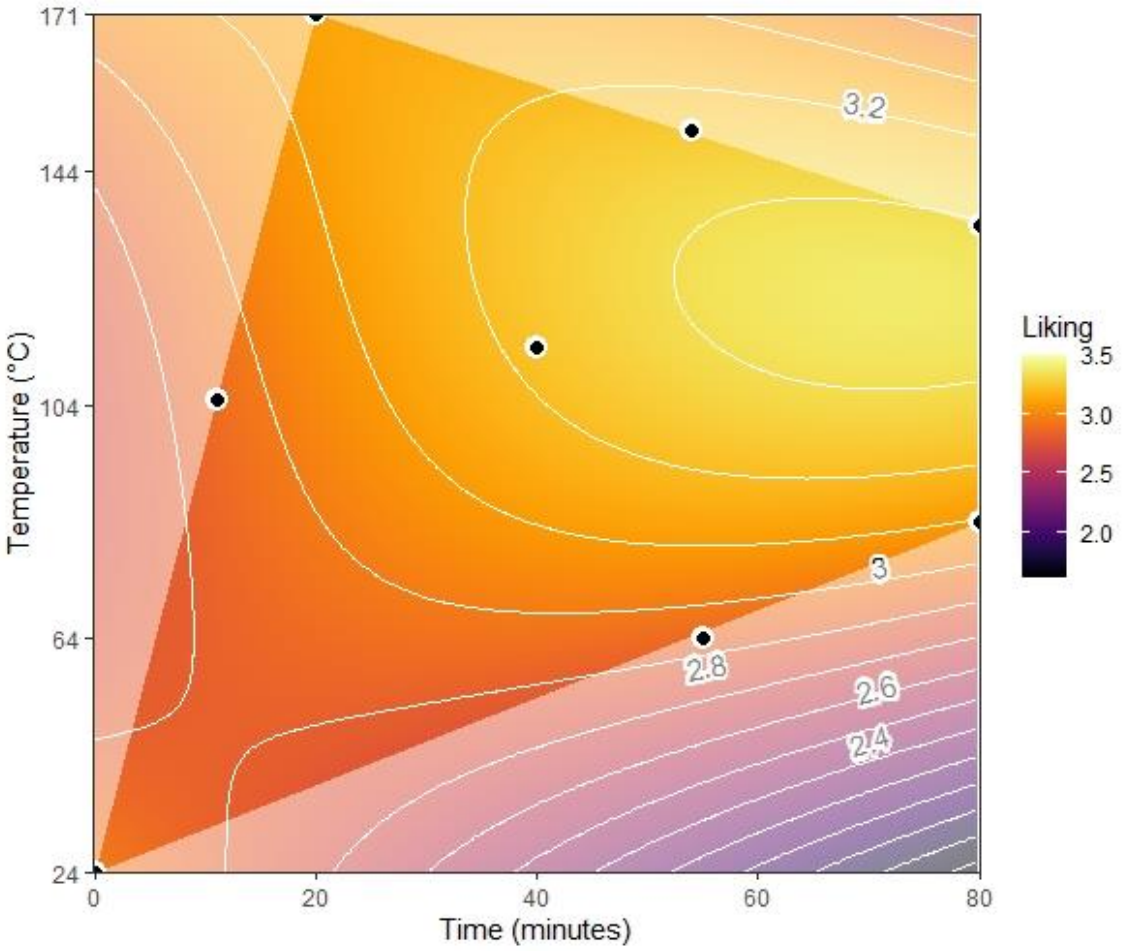

Figure 5. 37: MADAGASCAR Contour plot for predicted Liking for raw and roasted chocolate liquor treatments across the experimental region.

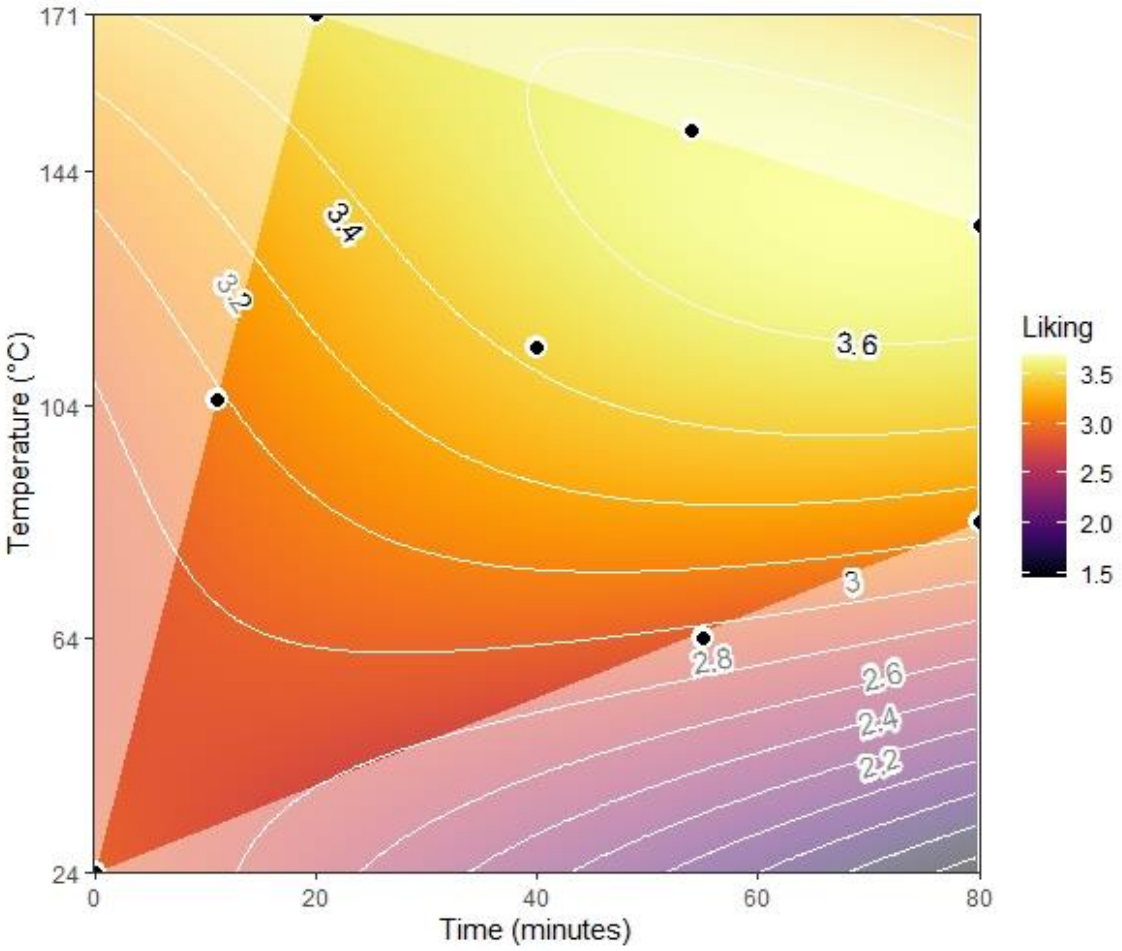

Figure 5. 38: PERU Contour plot for predicted Liking for raw and roasted chocolate liquor treatments across the experimental region.

5.3.2.3 Preference Mapping

Another way to understand Liking in relation to the various origins and treatments is through a technique called Preference Mapping, which is an attempt to determine which of the variety of sensory characteristics rated by panelists best explains differences in Liking, and how these relate to the specific rated products (Lê et al.,

2014). Such a technique, based on PCA, can give additional context to the response surface contour plots showcased above.

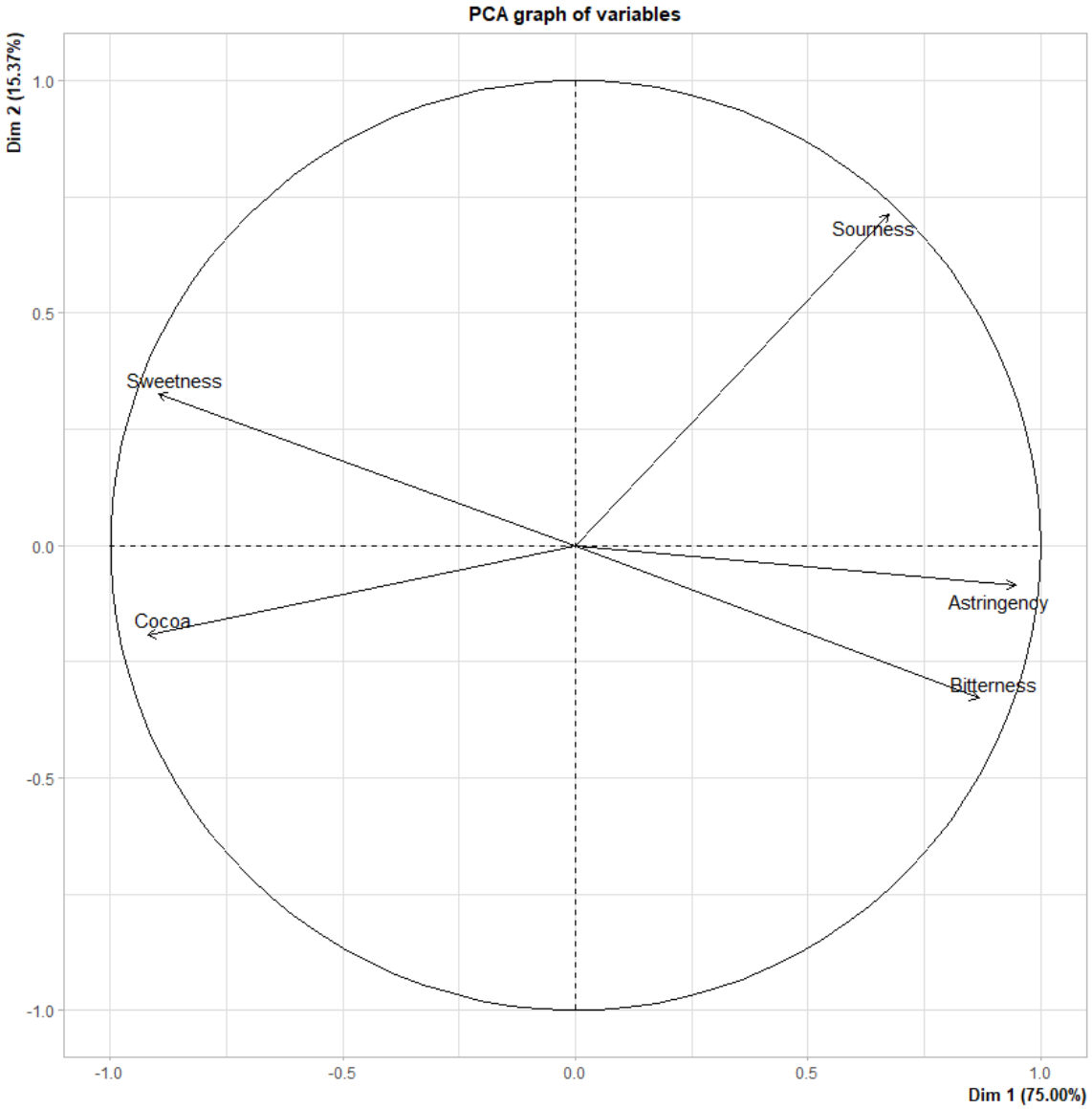

Figure 5. 39: PCA Loadings used for Preference Mapping (90.37% of variation in the treatment-wise means of these five predictors is explained with two dimensions)

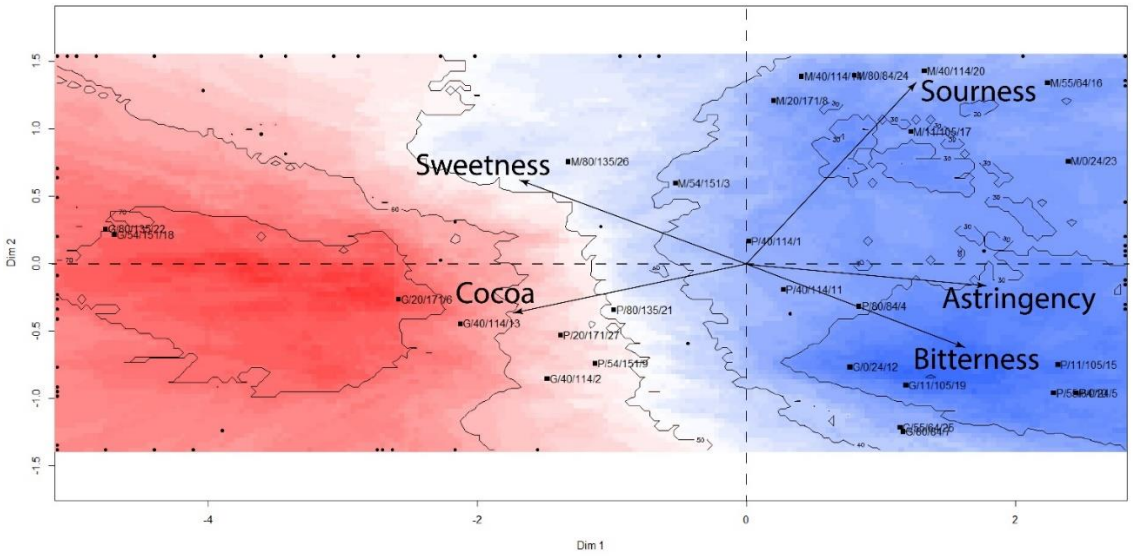

Figure 5. 40: Preference Map Showing Loadings of Sensory Characteristics Mapped to Liking of specific treatments. Dark red means most liked, and dark blue is least liked

PCA analysis was carried out on treatment-wise mean values of the sensory characteristics (i.e. Bitterness, Sourness, Astringency, Sweetness, Cocoa), and figure 5.39 shows the loadings of the two dimensions that explain the most variation in those mean values. Principal Component Regression (PCR) was carried out on treatment-wise mean values of Liking as the response, with these previously selected two dimensions as predictors, and the adjusted R^2 for this model is 0.86, meaning that 86% of the variance in the treatment-wise Liking means is explained by the first two dimensions, and the p-value is 2.167×10^{-11}, making the model highly significant (p-value≤0.001). Next, the first two dimensions of the PCA model were used in conjunction with the full consumer-wise scaled Liking ratings to produce a preference map (see SensoMineR package details in Materials & Methods) showing the relationship of the complete scaled Liking data, the

individual treatments, and the PCA loadings of the sensory variables described by dimensions 1 and 2 (figure 5.40). Figures 5.41 and 5.42 show the two halves of the map, where origins and roasting conditions (Time and Temperature) are visible for each treatment. What we can see is that Liking of particular samples is most associated with increased Cocoa Intensity and secondarily Sweetness, while Liking samples less is most associated with Astringency and Bitterness, and then Sourness. Additionally, it is clear on the left side of the map (Figure 5.41) that the most liked treatments are those that are both Ghanaian and which have undergone the greatest degree of roast in the experimental region (i.e., the combinations 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C). The next most liked samples, in approximate order, were Ghana roasted for 40 minutes at 114°C, and Peru, for the same three roasting conditions as the most liked Ghanaian samples (i.e., the 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C). Only then does a Madagascar treatment (80 minutes/135°C) appear as the next most liked, and it is more correlated with Sweetness than Cocoa.

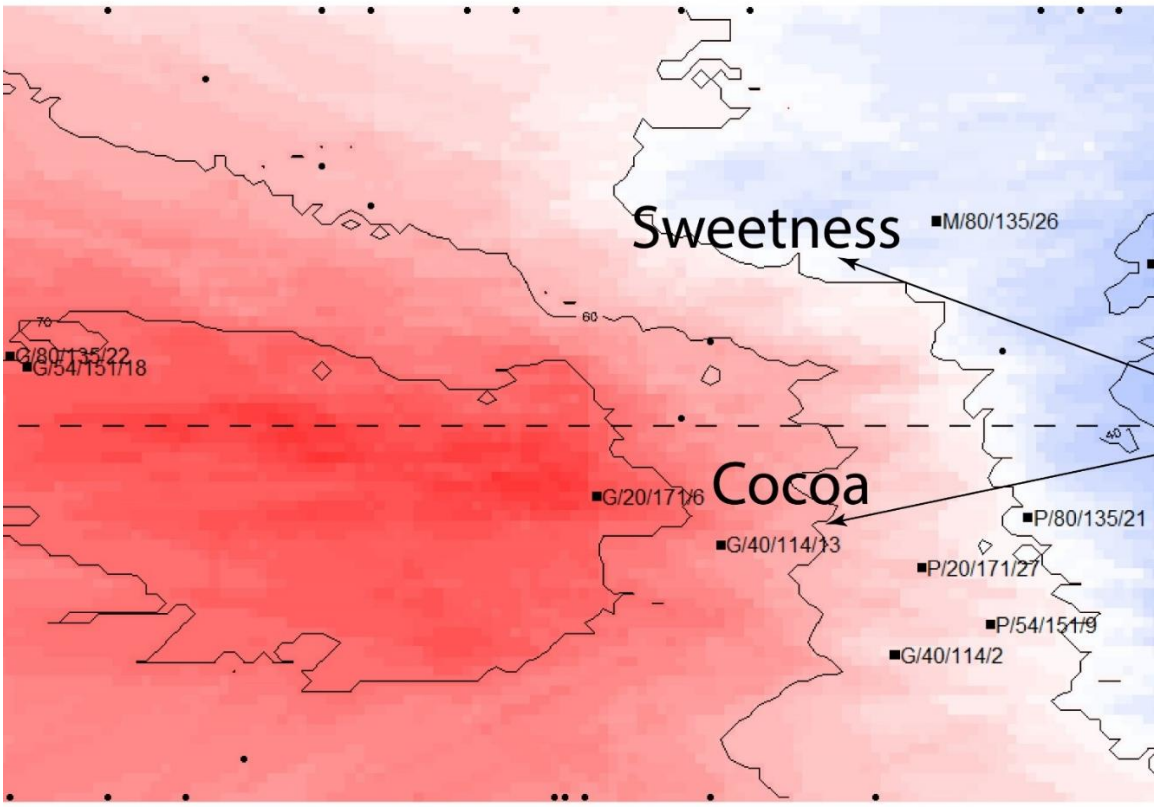

Figure 5. 41: Left side of Preference Map. Coding of Treatments is Origin/Time/Temperature/Treatment#. Therefore G/20/171/6 is Ghana roasted for 20 minutes at 171°C, and with treatment code #6. Darker Red is more liked. Blue is less liked.

On the right side of the map (Figure 5.42), we can see that most of the treatments made with Madagascar cacao are relatively highly disliked, including the raw and lightly roasted treatments (i.e., 0 minutes at 24°C, 11 minutes at 105°C, or 55 minutes at 64°C), and also highly correlated with Sourness, while the least liked treatments overall, are made of Peruvian and Ghanaian cacao that is raw or very lightly roasted (i.e., 0 minutes at 24°C, 11 minutes at 105°C, or 55 minutes at 64°C), and in these cases, the treatments are more correlated with Bitterness, and Astringency, especially Peruvian treatments in the case of the latter sensory characteristic. All of these findings support

previously detailed response surface contour plots, where it was seen that in general, as roast proceeded with Time and especially Temperature increases, Bitterness, Sourness and Astringency decreased while Cocoa Intensity, Sweetness and Liking increased. Thus, it is not surprising to see a correlation between Liking and increased Cocoa Intensity and Sweetness, nor is it surprising to find that the raw and lightly roasted treatments (i.e., 0 minutes at 24°C, 11 minutes at 105°C, or 55 minutes at 64°C) of all three origins are least liked and the most roasted treatments were generally better liked, especially in the case of Ghanaian treatments which had the estimated highest Sweetness and Cocoa Intensity levels for the most roasted treatments (i.e., the combinations 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C).

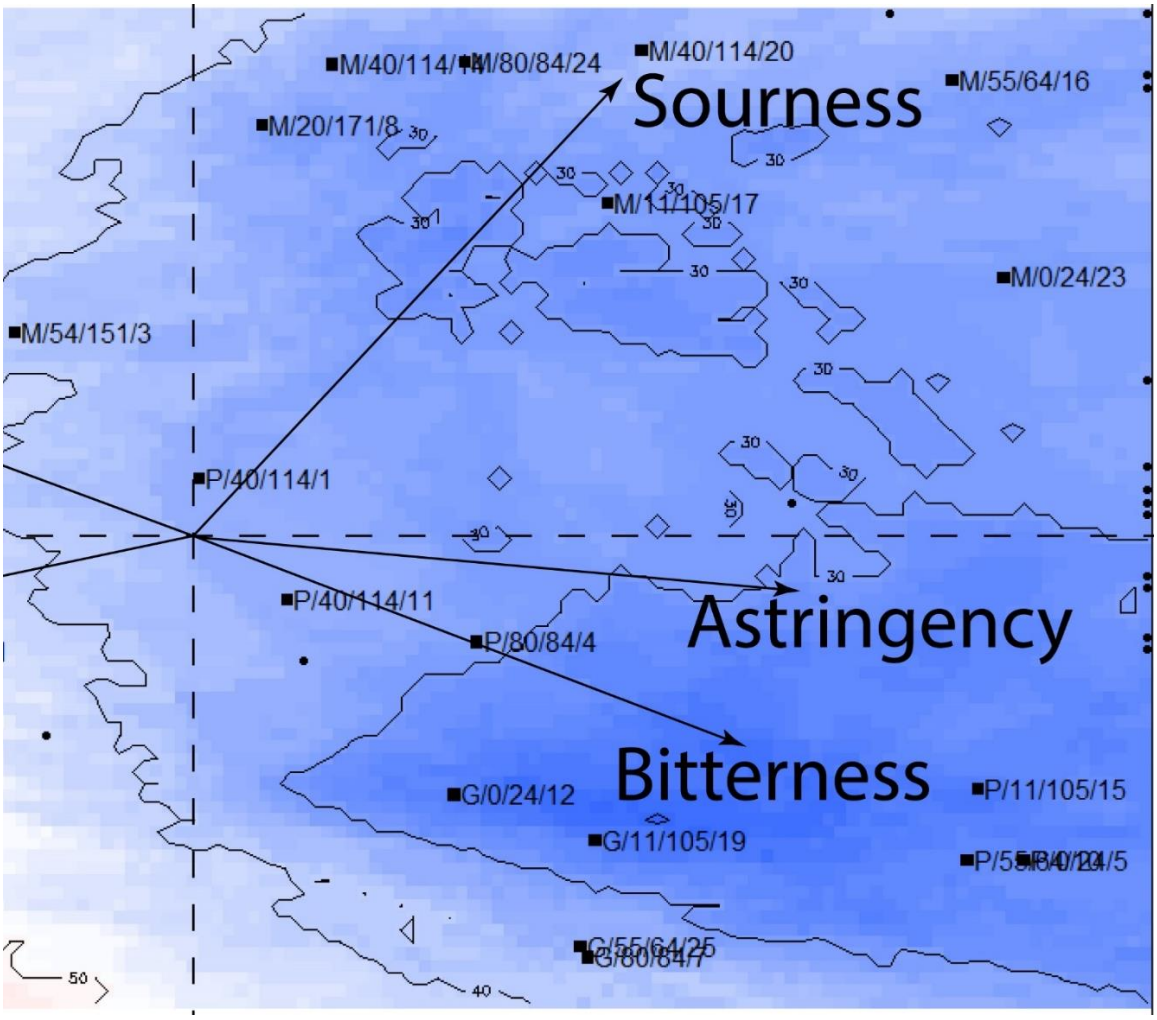

Figure 5. 42: Right side of Preference Map. Coding of Treatments is Origin/Time/Temperature/Product#. Therefore G/11/105/19 is Ghana roasted for 11 minutes at 105°C, and with product code #19. Darker Blue is less liked.

5.3.3 The Effect of Noseclips on Consumer Ratings of Sensory and Hedonic Characteristics of Roasted Ghanaian Chocolate Liquor

Section 2 of the Consumer study was conducted to better understand the impact of aroma on the perception of sensory modalities such as Bitterness, and hedonic

characteristics such as Liking when it comes to chocolate. It is well known that there may be an interplay of the various sensory modalities that can lead to suppression or enhancement of a specific perception (i.e., suppression of bitterness in the presence of sweetness (Lawless et al., 2010b). Such a study was achieved by asking consumer participants to taste all of the previously rated Ghanaian treatments from Section 1 of the study, and to rate the same characteristics, but this time to do so while wearing noseclips, which restrict the ability to smell. Exploratory analysis was then carried out with PCA, and mixed-model linear regression was used on the dataset consisting of all of the Ghanaian treatments tasted with and without noseclips.

5.3.3.1 Principal Components Analysis

The biplot of data from Section 2 of our study (Figure 5.43), where noseclips were worn, and which consists of samples of roasted treatments from Ghanaian cacao, shows slightly different correlations between factors than previously seen with the non-noseclip rating of the same roasted Ghana treatments (5.44). For example, Liking is more correlated with Other and less correlated with Cocoa, a seeming switch from Section 1, making Cocoa almost orthogonal to Liking. Additionally, Liking is most negatively correlated with Bitterness, instead of Astringency, as was the case with the biplot of Section 1 data (Figure 5.44). Further, plots of loadings for Time and Temperature are dramatically smaller. An important takeaway is that without the ability to smell the samples, most indicators of extent of roast, which are predominantly Maillard reaction-related volatile aroma compounds (Aprotosoaie et al., 2016a) are not

able to be sensed, and so Time and Temperature, of which increases in both are associated with roasting, generally explain much less of the variance in the overall data for Section 2 than do Liking or the non-aroma-based sensory characteristics. The lack of perceived roasted Cocoa Intensity (i.e., cocoa flavor/aroma) could also help to explain the orthogonal position of Cocoa to Liking, and while the Cocoa loading is still large, that may simply imply that participants are still using the Cocoa category, when unable to actually sense Cocoa aroma, to rate one or more other tastes or mouth-oriented sensations, such as astringency, which they might perceive to be related to Cocoa as a food, and of course this may also not be consistent from participant to participant.

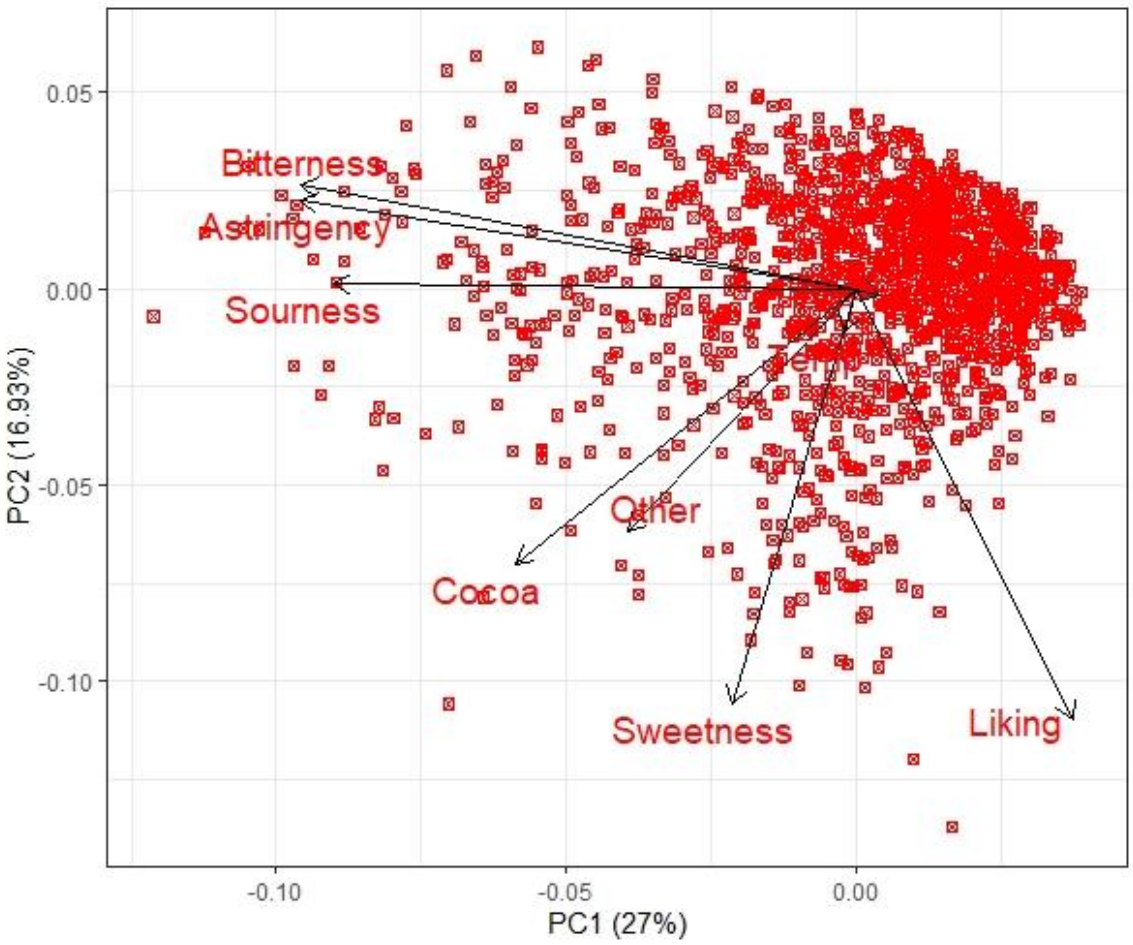

Figure 5. 43: WITH Noseclips Ghana PCA biplot showing 43.93% of the variance in the consumer sensory data for Section 2, which includes factors for all rated characteristics, as well as time and temperature.

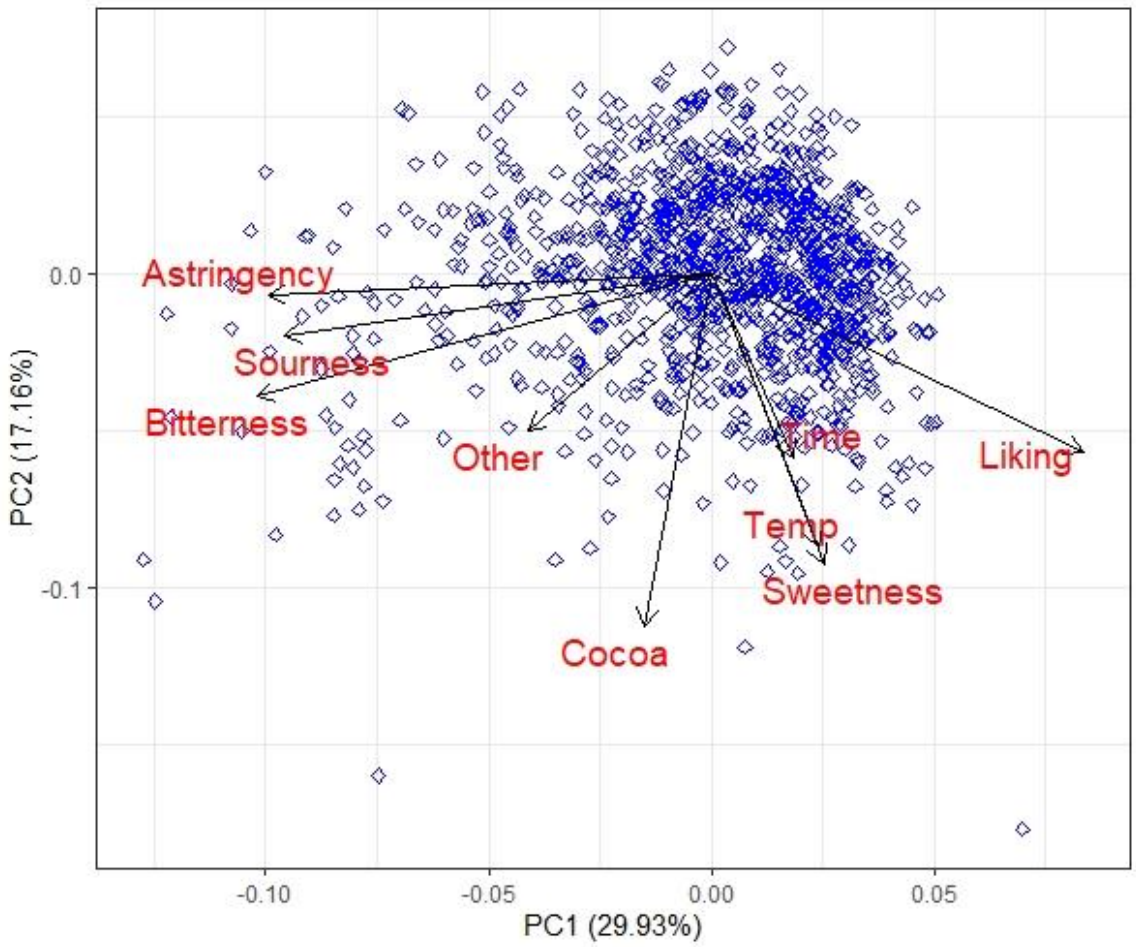

Figure 5. 44: WITHOUT Noseclips Ghana PCA biplot showing 47.09% of the variance in the consumer sensory data for Section 1, which includes factors for all rated characteristics, as well as time and temperature

5.3.3.2 Mixed Model Linear Regression

To carry out mixed-model linear regression, first the predictor variables of Time and Temperature were mean-centered and scaled given that they are otherwise on quite different scales in the experimental region. Next, a mixed model was fit with each of the sensory characteristics as the response variable in turn (i.e., Bitterness, Sourness, Astringency, Sweetness, Cocoa Intensity) and then subsequently also Liking as the

response variable, using the full potential model (i.e., all relevant first order and second-order effects and their interactions, as well as random effects) as follows:

$$\text{Response Variable} \sim \text{Time} + \text{Temperature} + \text{Time:Temperature} + \text{Time}^2 + \text{Temperature}^2 + \text{Temperature:Time}^2 + \text{Time:Temperature}^2 + \text{Nose} + \text{Nose:Time} + \text{Nose:Temperature} + \text{Nose:Time:Temperature} + \text{Nose:Time}^2 + \text{Nose:Temperature}^2 + \text{Nose:Temperature:Time}^2 + \text{Nose:Time:Temperature}^2 + (1|\text{OAConsumer}) + (1|\text{OAorder}) + (1|\text{Day}) + (1|\text{AgeBin}) + (1|\text{ChocPref}) + (1|\text{WDorder}) + (1|\text{OAConsumer:Day})$$

Nose is a categorical variable encoding use or lack of noseclips during consumer assessments. Terms in parentheses are random effects, and consist of the following: OAConsumer is the individual consumer coded such that each participant has a single number for the entire two-section study (e.g., OAConsumer 1 would retain that code for both Section 1 and Section 2 if s/he is present in both) , OAorder is the overall sample order across each individual section separately (all days), Day is the specific day of each Section (such as first, second, third, etc.), AgeBin is the full range of ages from 18-65 split into 5 bins and analyzed as categorical factors, ChocPref is one of six chocolate preferences chosen by each consumer as described in Materials & Methods, WDorder is within-day sample tasting order.

Model selection was carried out using backward stepwise selection with R package lmerTest (see Materials and Methods for full information on this package), which first selects the best random effect structure, with single-term deletions based upon p-

values calculated with the REML-likelihood ratio test. Next, the fixed-effect structure is chosen with single-term deletions relying upon Satterthwaite's method for calculating degrees of freedom and p-values, while respecting hierarchy when interactions occur. Residuals of each selected model were then subjected to diagnostic plots. With this data set, residuals in selected models for each response were found to have a non-skewed distribution, but with some lack of normality in the tails, but a square-root transformation to the response substantially corrected the issue in all cases. After transformation, backward stepwise regression was yet again performed, resulting in equivalent models in most cases, and the residuals were once again inspected via diagnostic plots to confirm that no significant problems remained. Additionally, unless otherwise noted below, variance inflation factor (VIF) tests showed that VIF were less than 2.5 for all first order main effects, ruling out multicollinearity as an important contributor to coefficient estimate error and term significance. Type III ANOVA was then performed for each selected model to obtain estimates of p-values and coefficients for each fixed-effect term and p-values for each random-effect term in the model. Given the presence of random effects, conventional R^2 values could not be computed for the models, so pseudo R^2 values were instead computed, and all were greater than 0.5. Given the nature of the data (i.e., psychophysical data based upon sensory analysis) R^2 values over 0.25 are considered large (Cohen, 1988; Hemphill, 2003). Additionally, contour plots were prepared for each model to help visualize the predicted values of the selected models. All details outlined can be seen below for each response variable.

5.3.3.2.1 Bitterness

The model selected for best predicting Bitterness is:

sqrt(Bitterness) ~ Time + Temperature + Time:Temperature +

Nose + Nose:Time + Nose:Temperature + Nose:Time:Temperature +

(1 | OAConsumer) + (1 | OAorder) + (1 | Day) + (1 | OAConsumer:Day)

The summary of calculated model statistics can be seen in tables 5.26 to 5.28.

Table 5. 26: Summary of calculated model statistics, including effect estimates

```
Fixed effects:
                    Estimate Std. Error         df t value Pr(>|t|)
(Intercept)       4.635e+00  2.202e-01  5.876e+00  21.049 9.31e-07 ***
Time             -5.922e-02  2.713e-02  2.204e+03  -2.182   0.0292 *
Temp             -1.539e-01  3.099e-02  2.256e+03  -4.966 7.36e-07 ***
Nose1            -4.842e-01  4.854e-02  2.322e+03  -9.976  < 2e-16 ***
Time:Temp        -1.404e-01  2.732e-02  2.284e+03  -5.139 3.00e-07 ***
Time:Nose1        2.223e-04  2.714e-02  2.207e+03   0.008   0.9935
Temp:Nose1        1.301e-01  3.105e-02  2.263e+03   4.191 2.88e-05 ***
Time:Temp:Nose1   6.487e-02  2.731e-02  2.280e+03   2.375   0.0176 *

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 27: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

                Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
Time             7.053   7.053     1 2204.2  4.7631   0.02918 *
Temp            36.508  36.508     1 2256.2 24.6564 7.365e-07 ***
Nose           147.353 147.353     1 2322.2 99.5185 < 2.2e-16 ***
Time:Temp       39.103  39.103     1 2284.1 26.4095 2.996e-07 ***
Time:Nose        0.000   0.000     1 2206.8  0.0001   0.99346
Temp:Nose       26.009  26.009     1 2262.8 17.5660 2.882e-05 ***
Time:Temp:Nose   8.354   8.354     1 2280.2  5.6418   0.01762 *

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 28: ANOVA-Like Table for Random Effects (Single-Term Deletions)

	npar	logLik	AIC	LRT	Df	Pr(>Chisq)	
<none>	13	-4693.9	9413.7				
(1 | OAConsumer)	12	-4886.8	9797.7	385.97	1	< 2.2e-16	***
(1 | OAorder)	12	-4697.3	9418.5	6.84	1	0.008939	**
(1 | Day)	12	-4705.7	9435.5	23.79	1	1.075e-06	***
(1 | OAConsumer:Day)	12	-4742.2	9508.4	96.67	1	< 2.2e-16	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

For this model, the pseudo R-squared is 0.6359074, suggesting that the combination of fixed and random effects in the model describe approximately 63.6% of the variance in the square root of Bitterness. Of the fixed-effects terms (Table 5.27), Nose, Temp, Time:Temp, and Temp:Nose are highly significant (i.e., p-value <0.001), Time and Time:Temp:Nose are significant (p-value <0.05), and Time:Nose is not significant (p-value >0.05), but remains in the model due to heredity. As for the random-effects terms (Table 5.28), they are all highly significant (i.e., p-value <0.001), with the exception of (1|OAorder) which is significant (p-value≤0.05).

In Figure 5.45, we see the previously outlined general trend of Bitterness change with roast, with the highest Bitterness in the region near the raw treatment and light roasts in the bottom half of the figure, and lowest Bitterness in the uppermost portion of the figure near the most extreme roasts. However, it is interesting that once aroma is no longer perceived, the pattern of change in Bitterness across the region is substantially different (Figure 5.46). First, noseclips clearly have the effect of Bitterness depression across the experimental region, or put otherwise, perceiving aroma with the nose

clearly has an effect of Bitterness enhancement across the experimental region. Additionally, it is interesting to see that with noseclips, while the area of least Bitterness is estimated to be in the upper-right quadrant of the experimental region, near the 80 minutes/135°C roast treatment, it is also true that the area near the raw treatment (0 minutes, 24°C) is estimated to be very close to that same level, while the region near the 20 minutes/171°C is estimated to have the highest Bitterness. As there are changes in bitter tasting chemical compounds across the experimental region during roasting (see Chapter 4), with some increasing while others decrease, such a finding, in the absence of aroma, is not necessarily surprising, and the direct relationship between Bitterness and various chemical compounds will be inspected in detail in Chapter 6, in order to ascertain potential chemical causes for this behavior.

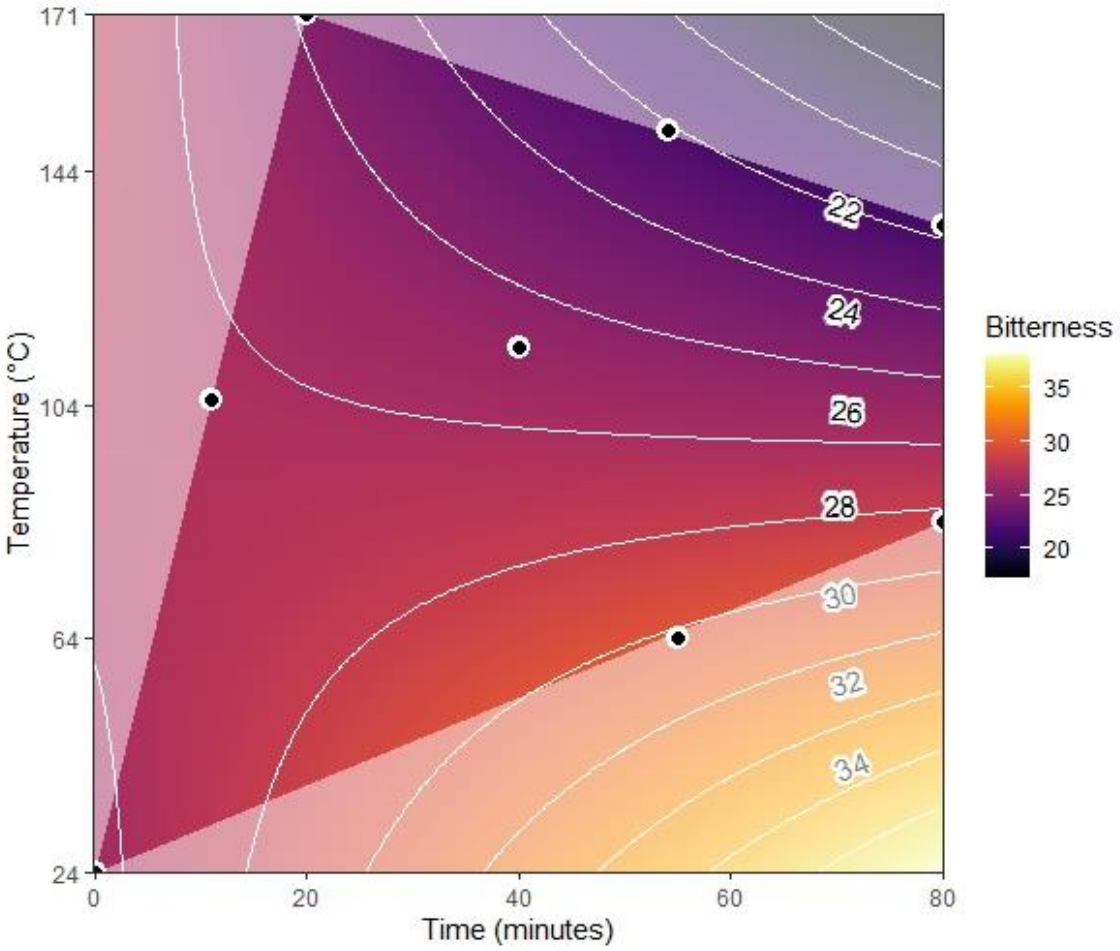

Figure 5. 45: WITHOUT Noseclips GHANA Contour plot for predicted Bitterness for raw and roasted chocolate liquor treatments across the experimental region.

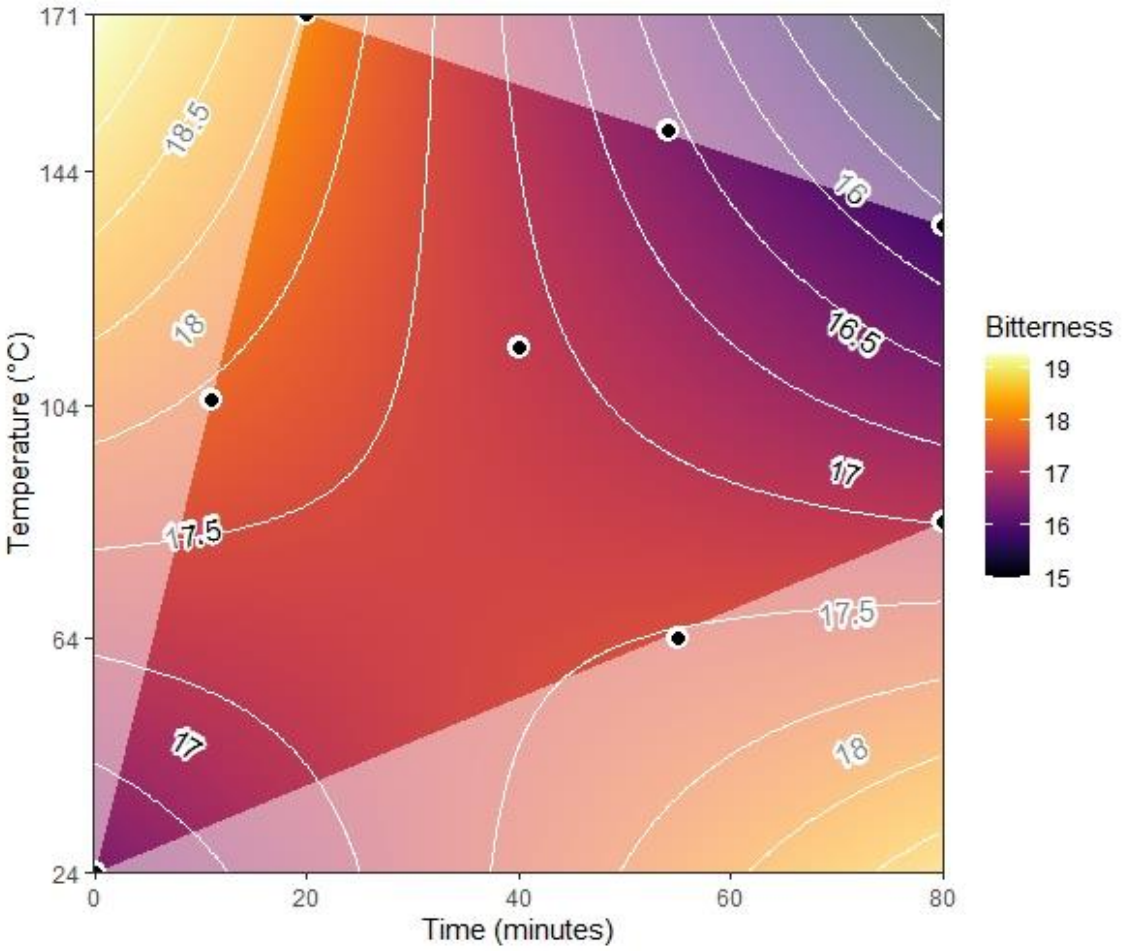

Figure 5. 46: WITH Noseclips Ghana Contour plot for predicted Bitterness for raw and roasted Ghanaian chocolate liquor treatments across the experimental region.

5.3.3.2.2 Astringency

The model selected for best predicting Astringency is:

sqrt(Astringency) ~ Time + Temperature + Time:Temperature +

Nose + Nose:Time + Nose:Temperature + Nose:Time:Temperature +

(1 | OAConsumer) + (1 | Day) + (1 | OAConsumer:Day) + (1|WDorder)

The summary of calculated model statistics can be seen in tables 5.29 to 5.31.

Table 5. 29: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

                  Estimate Std. Error         df t value Pr(>|t|)
(Intercept)        4.10006    0.20921    9.67560  19.597 4.13e-09 ***
Time              -0.08274    0.02837 2195.94696  -2.916  0.00358 **
Temp              -0.22878    0.03241 2251.52093  -7.058 2.23e-12 ***
Nose1             -0.37870    0.05113 2358.59893  -7.407 1.79e-13 ***
Time:Temp         -0.13115    0.02858 2283.77870  -4.588 4.72e-06 ***
Time:Nose1         0.05718    0.02837 2198.96481   2.015  0.04401 *
Temp:Nose1         0.18384    0.03248 2259.12311   5.661 1.70e-08 ***
Time:Temp:Nose1    0.08032    0.02857 2280.03006   2.811  0.00498 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 30: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

               Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
Time           13.715  13.715     1 2195.9  8.5037  0.003580 **
Temp           80.355  80.355     1 2251.5 49.8209 2.234e-12 ***
Nose           88.486  88.486     1 2358.6 54.8625 1.790e-13 ***
Time:Temp      33.950  33.950     1 2283.8 21.0492 4.719e-06 ***
Time:Nose       6.550   6.550     1 2199.0  4.0610  0.044006 *
Temp:Nose      51.686  51.686     1 2259.1 32.0456 1.697e-08 ***
Time:Temp:Nose 12.746  12.746     1 2280.0  7.9029  0.004978 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 31: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                      npar  logLik     AIC    LRT Df Pr(>Chisq)
 <none>                 13 -4821.8  9669.6
 (1 | OAConsumer)       12 -5027.8 10079.5 411.90  1  < 2.2e-16 ***
 (1 | Day)              12 -4851.9  9727.7  60.09  1  9.046e-15 ***
 (1 | OAConsumer:Day)   12 -4870.5  9765.0  97.36  1  < 2.2e-16 ***
 (1 | WDorder)          12 -4835.2  9694.4  26.79  1  2.262e-07 ***
 ---
 Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.6466555, suggesting that the combination of fixed and random effects in the model describe approximately 64.7% of the variance in

the square root of Astringency. Of the fixed-effects terms (Table 5.30), Nose, Temp, Time:Temp, and Temp:Nose are highly significant (i.e., p-value <0.001), and Time, Time:Nose, and Time:Temp:Nose are significant (p-value <0.05). All random-effects terms (Table 5.28), are highly significant (i.e., p-value <0.001). In terms of the estimated effect on Astringency, or specifically the square root of Astringency, of the main effects, Nose is by far most important.

This can be seen by comparing Figures 5.47 and 5.48, where we note in Figure 5.47 the previously outlined general trend of Astringency change with roast, with the highest Astringency in the region near the raw treatment or light roasts in the bottom half of the figure and lowest Astringency in the uppermost portion of the region near the most extreme roasts. However, once aroma is no longer perceived, the pattern of change in Astringency is substantially different (Figure 5.48). First, noseclips clearly have the effect of mean Astringency depression across the experimental region, or put otherwise, perceiving aroma with the nose clearly has an effect of mean Astringency enhancement across the experimental region. Additionally, it is interesting to see that while wearing noseclips, the area of least Astringency is estimated to be in the upper-right quadrant of the experimental region, near the 80 minutes/135°C roast treatment, but it is also true that the area near the raw treatment (0 minutes, 24°C) is estimated to be very similar in Astringency, while the region near the 20 minutes/171°C is estimated to have even slightly higher Astringency than for raw, though not quite so high as the area near the 55 minutes at 64°C treatment, which is estimated to have the highest Astringency. As

there are changes in astringent chemical compounds across the experimental region during roasting (see Chapter 4), with some increasing while others decrease, such a finding, in the absence of aroma, is not necessarily surprising, and the direct relationship between Astringency and various chemical compounds should be inspected in further research in order to ascertain potential chemical causes for this behavior.

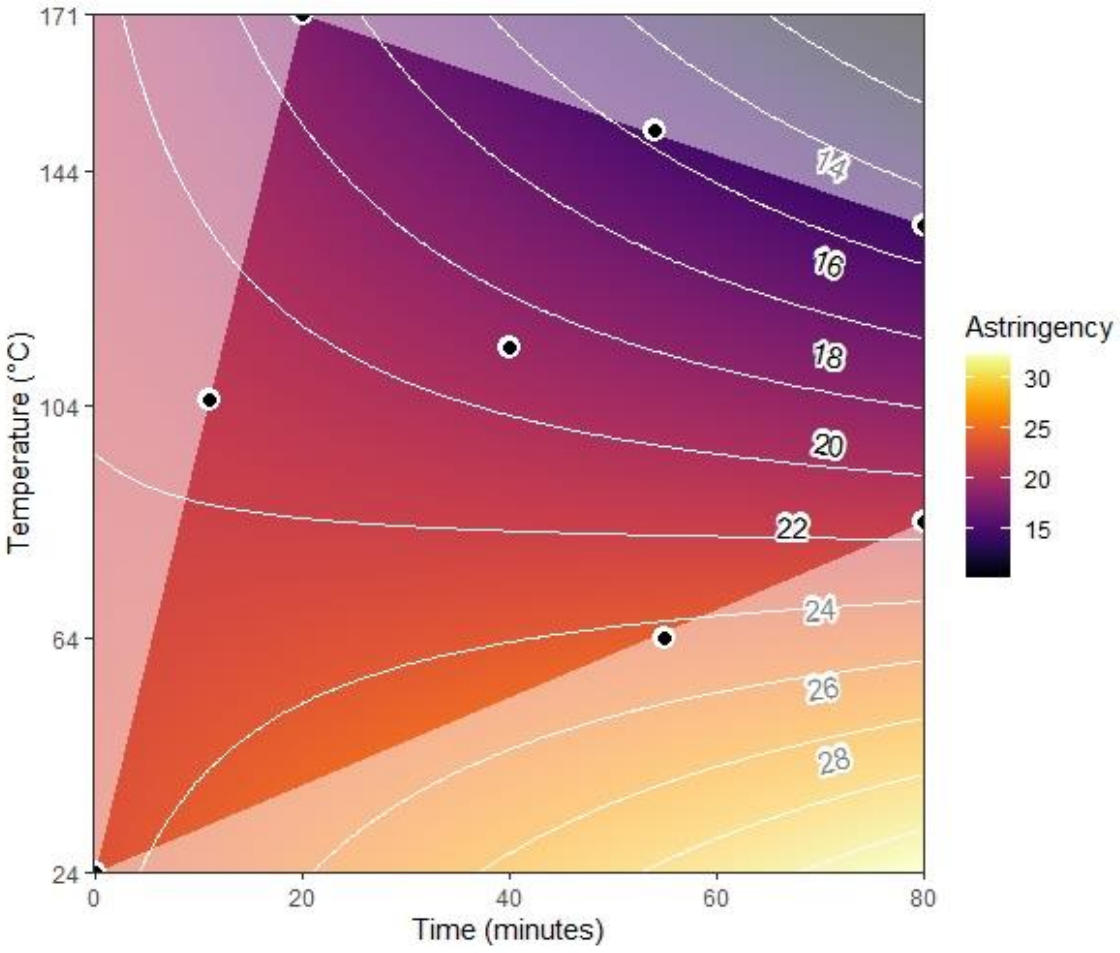

Figure 5. 47: WITHOUT Noseclips GHANA Contour plot for predicted Astringency for raw and roasted chocolate liquor treatments across the experimental region.

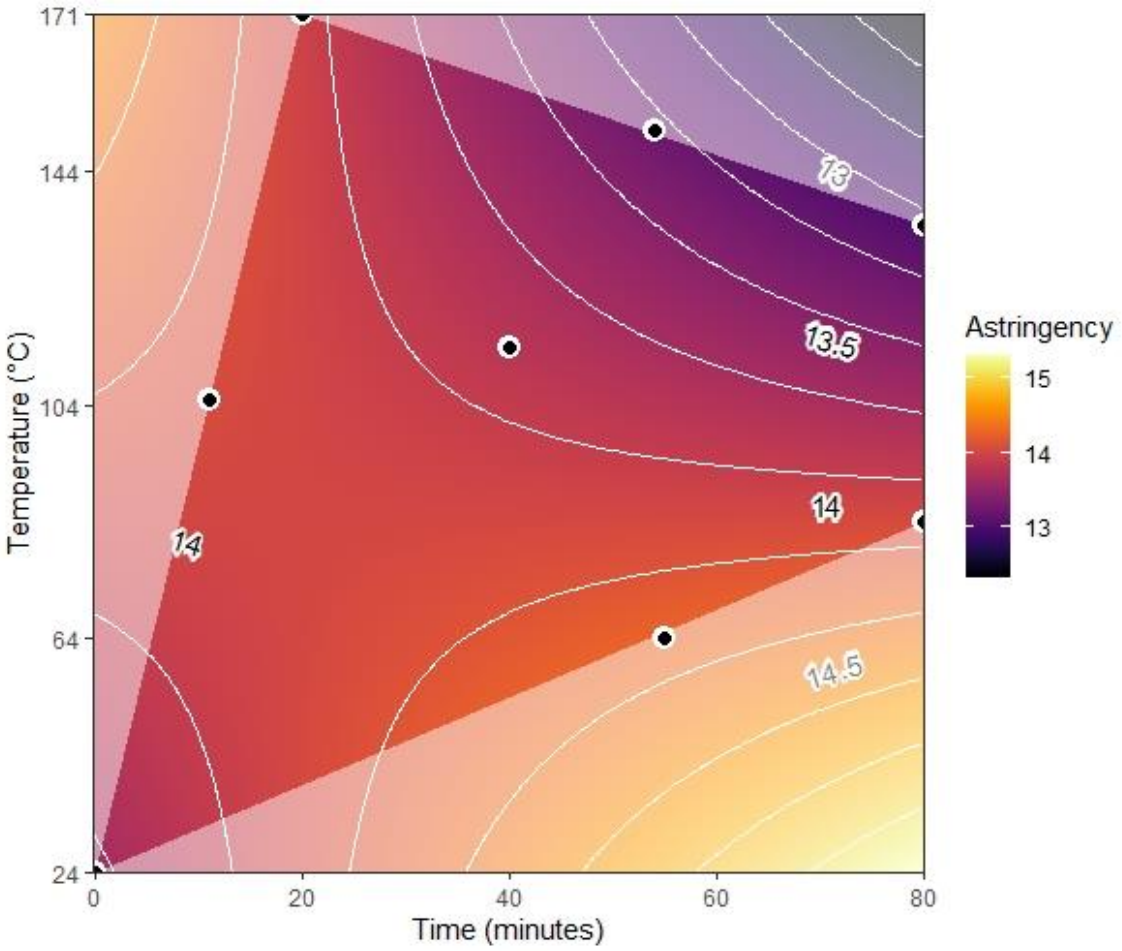

Figure 5. 48: WITH Noseclips Ghana Contour plot for predicted Astringency for raw and roasted Ghanaian chocolate liquor treatments across the experimental region.

5.3.3.2.1 Sourness

The model selected for best predicting Sourness is:

sqrt(Sourness) ~ Time + Temperature + Time:Temperature +

Nose + Nose:Temperature +

(1 | OAConsumer) + (1 | Day) + (1 | OAConsumer:Day) + (1|WDorder)

The summary of calculated model statistics can be seen in tables 5.32 to 5.34.

Table 5. 32: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

             Estimate Std. Error         df t value Pr(>|t|)
(Intercept)   3.42393    0.15661   25.63436  21.863  < 2e-16 ***
Time         -0.03802    0.02730 2164.24910  -1.393 0.163912
Temp         -0.13876    0.03133 2254.17256  -4.429 9.93e-06 ***
Nose1        -0.21595    0.04883 1862.83312  -4.423 1.03e-05 ***
Time:Temp    -0.09652    0.02728 2213.79617  -3.538 0.000412 ***
Temp:Nose1    0.14522    0.02612 2337.35724   5.561 2.99e-08 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 33: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

          Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
Time       3.062   3.062     1 2164.2  1.9391 0.1639123
Temp      30.969  30.969     1 2254.2 19.6147 9.926e-06 ***
Nose      30.881  30.881     1 1862.8 19.5590 1.032e-05 ***
Time:Temp 19.761  19.761     1 2213.8 12.5159 0.0004118 ***
Temp:Nose 48.820  48.820     1 2337.4 30.9209 2.994e-08 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 34: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                      npar  logLik     AIC    LRT Df Pr(>Chisq)
<none>                  11 -4730.8  9483.6
(1 | OAConsumer)        10 -5015.8 10051.5 569.90  1  < 2.2e-16 ***
(1 | Day)               10 -4736.2  9492.4  10.74  1   0.001046 **
(1 | OAConsumer:Day)    10 -4747.6  9515.2  33.53  1  7.019e-09 ***
(1 | WDorder)           10 -4749.2  9518.4  36.77  1  1.330e-09 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.6435359, suggesting that the combination of fixed and random effects in the model describe approximately 64.4% of the variance in

the square root of Sourness. Of the fixed-effects terms (Table 5.33), all are highly significant (i.e., p-value <0.001), with the exception of Time, which is not significant, (p-value >0.05), but remains in the model due to heredity. All random-effects terms (Table 5.34), are highly significant (i.e., p-value <0.001), with the exception of (1|Day), which is significant (p-value≤0.05). In terms of the estimated effect on Sourness, or to be specific the square root of Sourness, of the main effects, Nose is by far most important. This can be seen when comparing Figures 5.49 and 5.50.

We note in figure 5.49 the previously outlined general trend of Sourness change with roast, with the highest Sourness in the region near the raw treatment (i.e., 0 minutes at 24°C), and lowest Sourness in the uppermost portion of the region near the most extreme roasts (i.e., the combinations 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C). However, it is interesting that once aroma is no longer perceived, the pattern of change in Sourness is significant (Figure 5.48). First, noseclips clearly have a mean effect of Sourness depression across the experimental region, or put otherwise, perceiving aroma with the nose clearly has a mean effect of Sourness enhancement across the experimental region. Additionally, it is interesting to see that with noseclips, there is an area of reduced Sourness in the upper-right quadrant of the experimental region, near the 80 minutes/135°C roast treatment, and it is also true that the area near the raw treatment (0 minutes, 24°C) is estimated to be even slightly lower in Sourness, while the region near the 20 minutes/171°C is estimated to have the highest Sourness in the entire experimental region. This complex behavior is not necessarily surprising,

given that there are changes in concentration of sour tasting chemical compounds during roasting (Beckett et al., 2017), with volatile acids such as acetic acid decreasing on average (Afoakwa et al., 2008), while nonvolatile acids such as citric acid would become more concentrated with moisture loss during roasting. In order to better understand the relevance of each of the chemicals present in relation to this change, these relationships should be inspected in future research.

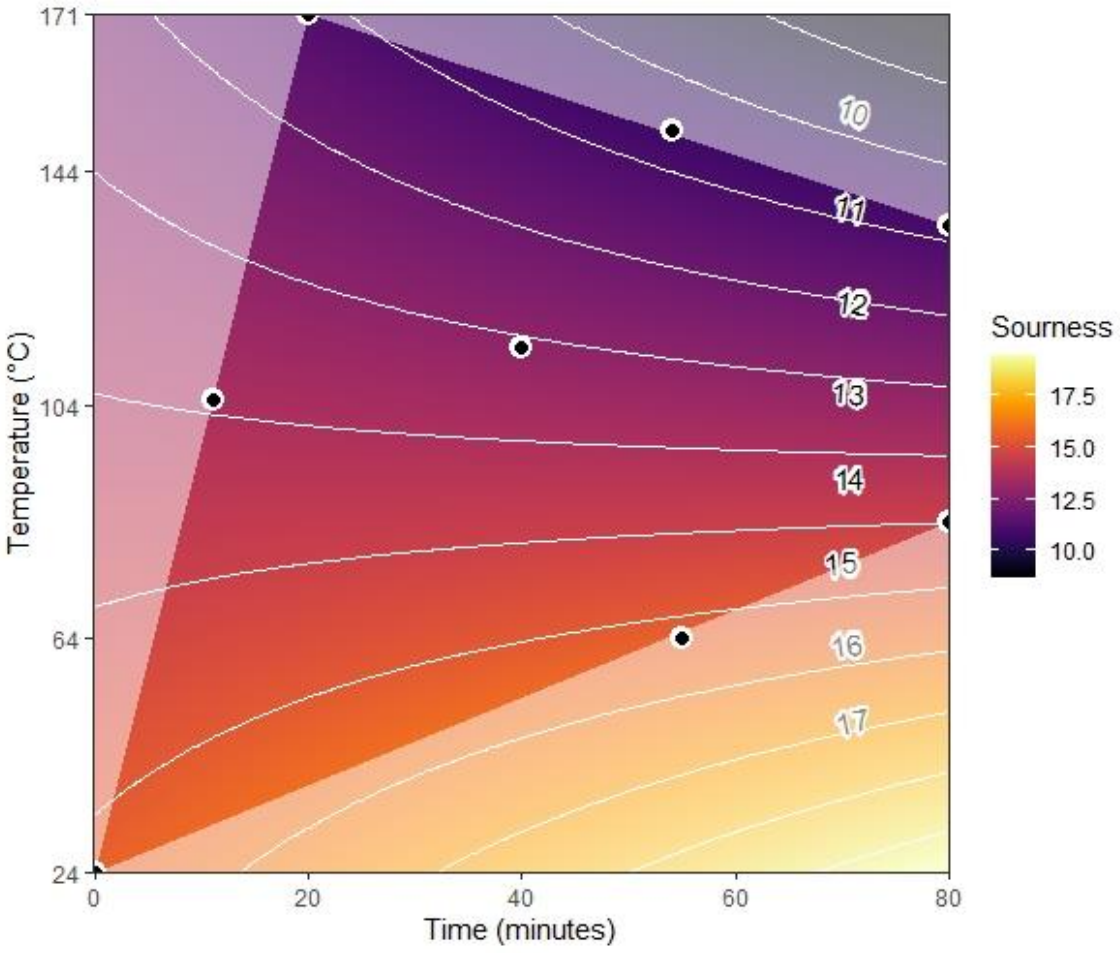

Figure 5. 49: WITHOUT Noseclips GHANA Contour plot for predicted Sourness for raw and roasted chocolate liquor treatments across the experimental region.

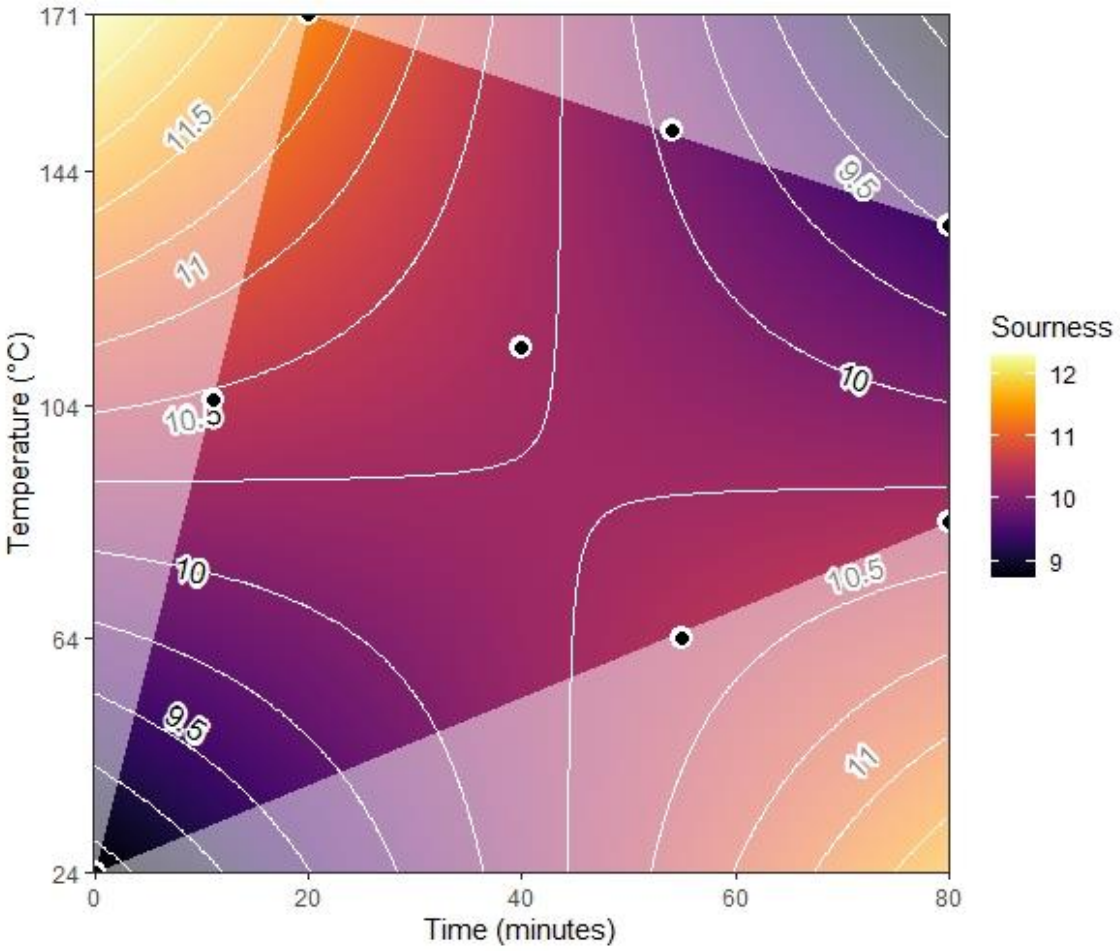

Figure 5. 50: WITH Noseclips Ghana Contour plot for predicted Sourness for raw and roasted Ghanaian chocolate liquor treatments across the experimental region.

5.3.3.3 Sweetness

The model selected for best predicting Sweetness is:

sqrt(Sweetness) ~ Time + Temperature + Time:Temperature +

Nose + Nose:Temperature +

(1 | OAConsumer) + (1 | Day) + (1 | OAConsumer:Day) + (1|WDorder)

The summary of calculated model statistics can be seen in tables 5.35 to 5.37.

Table 5. 35: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

                  Estimate Std. Error         df t value Pr(>|t|)
(Intercept)        1.46803    0.08336  111.63936  17.610  < 2e-16 ***
Time               0.05420    0.01894 2399.21486   2.861 0.004256 **
Temp               0.16249    0.01932 2399.75196   8.412  < 2e-16 ***
I(Temp^2)          0.03932    0.01732 2397.82033   2.270 0.023267 *
Nose1             -0.10486    0.02914 2553.74158  -3.599 0.000326 ***
Time:Temp          0.05923    0.01842 2403.29019   3.215 0.001322 **
Temp:Nose1        -0.11778    0.01930 2400.04458  -6.102 1.22e-09 ***
Time:Nose1        -0.02914    0.01698 2398.95635  -1.716 0.086269 .
Time:Temp:Nose1   -0.05926    0.01699 2403.42577  -3.488 0.000495 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 36: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

               Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
Time            5.091   5.091     1 2399.2  8.1871 0.0042556 **
Temp           43.996  43.996     1 2399.8 70.7540 < 2.2e-16 ***
I(Temp^2)       3.206   3.206     1 2397.8  5.1551 0.0232669 *
Nose            8.054   8.054     1 2553.7 12.9529 0.0003255 ***
Time:Temp       6.427   6.427     1 2403.3 10.3365 0.0013215 **
Temp:Nose      23.150  23.150     1 2400.0 37.2300  1.22e-09 ***
Time:Nose       1.831   1.831     1 2399.0  2.9451 0.0862693 .
Time:Temp:Nose  7.567   7.567     1 2403.4 12.1690 0.0004946 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 37: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                  npar  logLik    AIC     LRT Df Pr(>Chisq)
  <none>            12 -3433.4 6890.9
  (1 | OAConsumer)  11 -4443.9 8909.7 2020.82  1  < 2.2e-16 ***
  (1 | WDorder)     11 -3437.0 6896.0    7.15  1   0.007513 **

  ---
  Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.663528, suggesting that the combination of fixed and random effects in the model describe approximately 66.4% of the variance in

the square root of Sweetness. Of the fixed-effects terms (Table 5.36), Temperature, Nose, Nose:Temperature, and Nose:Time:Temperature are highly significant (i.e., p-value <0.001), and with the exception of Time:Nose, which is not significant, (p-value >0.05), but remains in the model due to heredity, all other fixed-effect terms are significant (p-value≤0.05). As for random-effects terms (Table 5.37), OAConsumer is highly significant (i.e., p-value <0.001), and WDorder is significant (p-value≤0.05). In terms of the estimated effect on Sweetness, or to be more specific the square root of Sweetness, Temperature is the most important of the main effects, followed by Nose. This can be seen when comparing Figures 5.51 and 5.52.

We note in figure 5.51 the previously outlined general trend of Sweetness change with roast, with the lowest Sweetness in the region near the raw treatment (i.e., 0 minutes at 24°C), and greatest predicted Sweetness in the uppermost portion of the region near the most extreme roasts (i.e., the combinations 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C). However, the Nose effect and interactions suggest that once aroma is no longer perceived, the pattern of change in Sweetness is significant (Figure 5.52). We see that with noseclips, differences in Sweetness across the experimental region are moderated, making them very similar in areas of roasted and raw treatments. Interestingly, with noseclips, the most extreme roasts in the experimental region (i.e., the combinations 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C) are still estimated to be slightly sweeter than raw or lightly roasted treatments, even without the presence of aroma. Though the cause for this is unclear, this region of highest

Sweetness matches, to a large extent, the inverse of the pattern seen in Astringency perception with noseclips. Given the known effect of sensory-related suppression/enhancement effects (Lawless et al., 2010b), this potential interaction should be considered in future research to determine if it is valid.

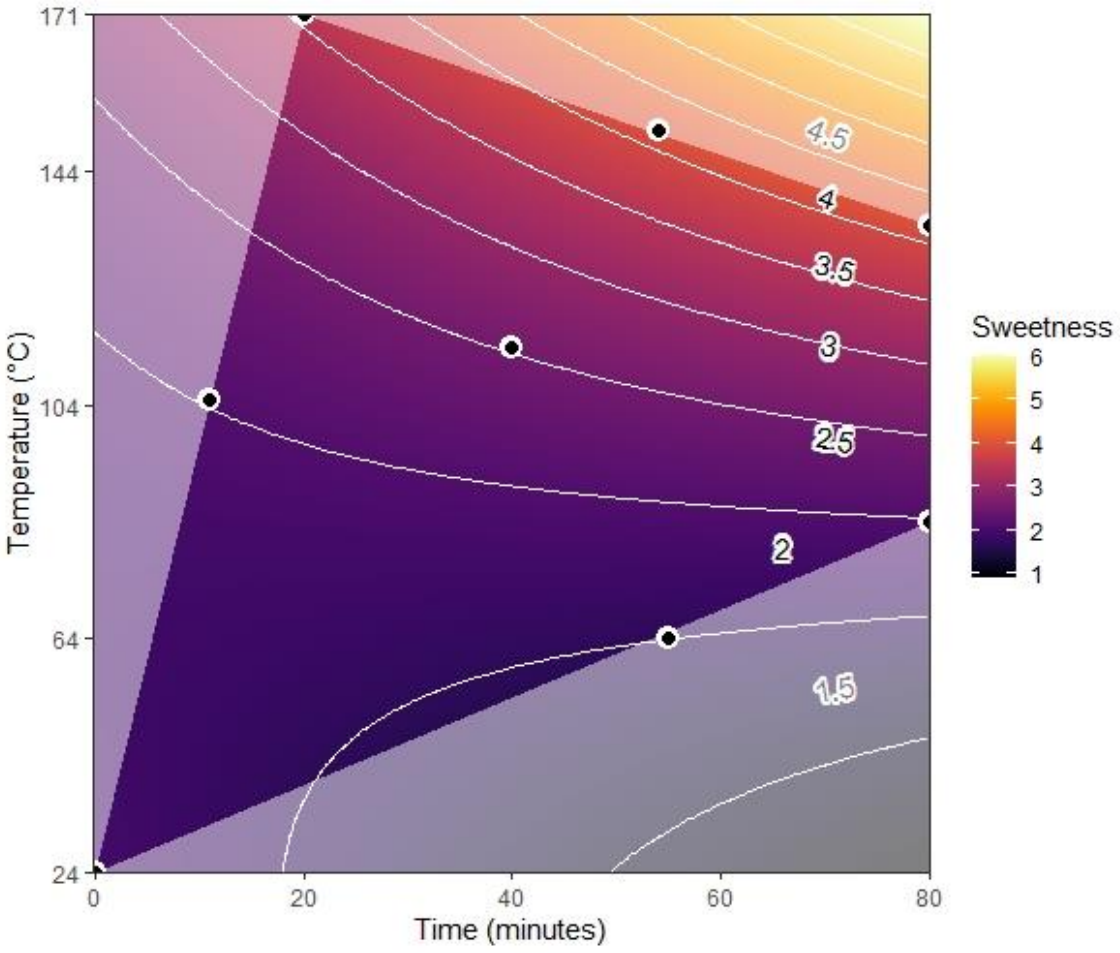

Figure 5. 51: WITHOUT Noseclips GHANA Contour plot for predicted Sweetness for raw and roasted chocolate liquor treatments across the experimental region.

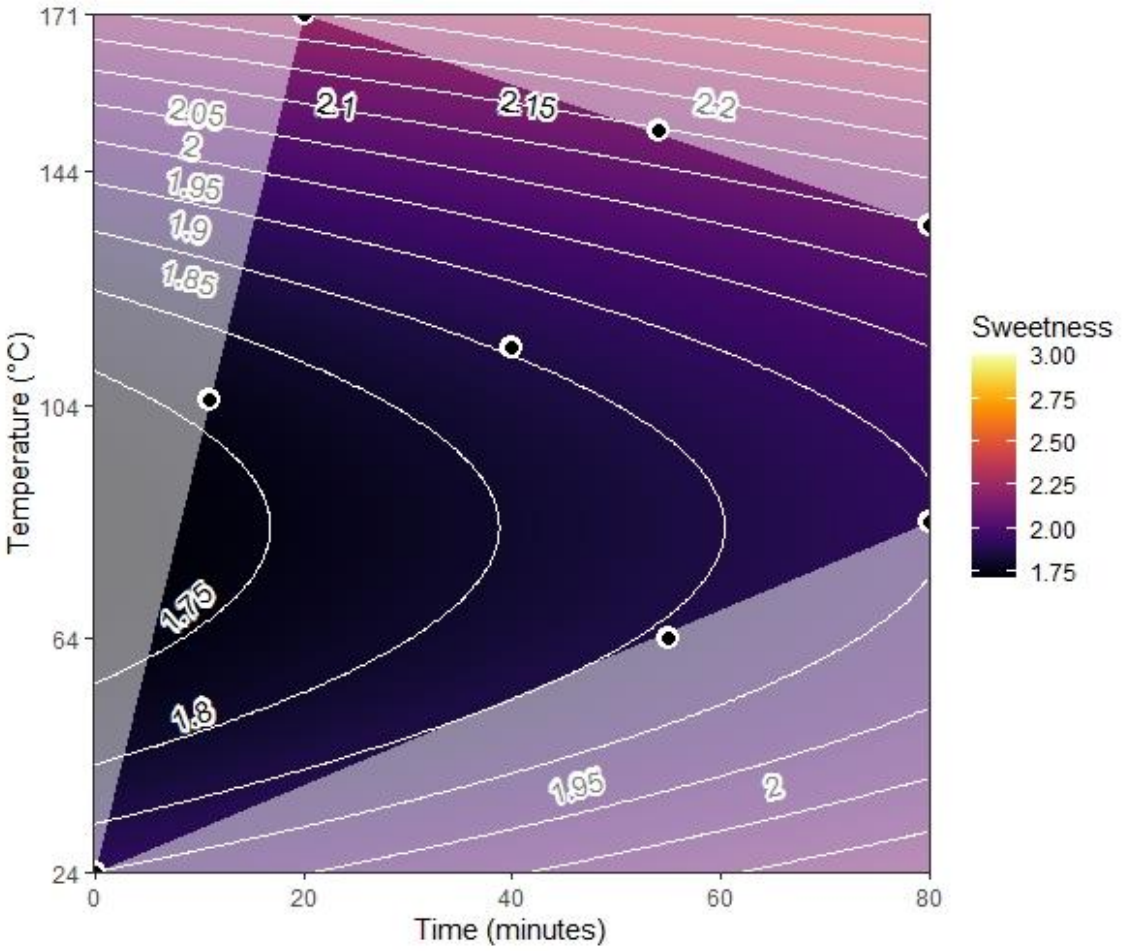

Figure 5. 52: WITH Noseclips Ghana Contour plot for predicted Sweetness for raw and roasted Ghanaian chocolate liquor treatments across the experimental region.

5.3.3.4 Cocoa Intensity

Unlike the previous sections where clearly defined taste modalities (i.e., Bitterness, Sourness, Sweetness), and perception of somatosensation (i.e., Astringency) were modeled, Cocoa Flavor intensity is different in that it has no single chemical compound that can be used for a reference, making a simple calibration solution impossible. Instead, as described in Materials & Methods, participants were supplied with a written description of what Cocoa Flavor means. Consumer participants were told that Cocoa Flavor is the "richness of deep dark chocolate and cocoa flavors" that might be found in "a piece of dark chocolate, or the smell of freshly baked chocolate brownies." Given that Cocoa was described as being related to flavors and especially smell of very specific cocoa-containing foods, one might already anticipate that noseclips would dramatically reduce the perception of Cocoa intensity. What is less clear is how this perception of Cocoa intensity will change across the experimental region, if at all. The model selected for best predicting Cocoa Flavor intensity (Cocoa) is:

$$\text{sqrt(Cocoa)} \sim \text{Time} + \text{Temperature} + \text{Time:Temperature} + \text{Time}^2 + \text{Temperature}^2$$
$$\text{Nose} + \text{Nose:Time} + \text{Nose:Temperature} + \text{Nose:Time:Temperature} + \text{Nose:Time}^2$$
$$(1 \mid \text{OAConsumer}) + (1 \mid \text{Day}) + (1 \mid \text{AgeBin}) + (1 \mid \text{OAConsumer:Day}) +$$
$$(1 \mid \text{WDorder})$$

The summary of calculated model statistics can be seen in tables 5.38 to 5.40.

Table 5. 38: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

                  Estimate Std. Error         df t value Pr(>|t|)
(Intercept)        3.54936    0.27172    5.92811  13.063 1.36e-05 ***
Time               0.08970    0.02689 2129.08826   3.336 0.000864 ***
Temp               0.22618    0.02726 2293.84603   8.299  < 2e-16 ***
I(Time^2)         -0.05289    0.03070 2344.27183  -1.723 0.085066 .
I(Temp^2)          0.07768    0.02415 2087.02571   3.216 0.001319 **
Nose1             -0.84444    0.05096 2509.17262 -16.569  < 2e-16 ***
Time:Temp          0.09959    0.02697 2348.82006   3.693 0.000227 ***
Time:Nose1        -0.11116    0.02353 2140.43550  -4.723 2.47e-06 ***
Temp:Nose1        -0.08444    0.02729 2297.00227  -3.094 0.002000 **
I(Time^2):Nose1    0.12558    0.02979 2363.53546   4.215 2.59e-05 ***
Time:Temp:Nose1   -0.07061    0.02571 2304.79520  -2.747 0.006069 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 39: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

                Sum Sq Mean Sq NumDF  DenDF  F value    Pr(>F)
Time            11.849  11.849     1 2129.1  11.1295 0.0008642 ***
Temp            73.320  73.320     1 2293.8  68.8681 < 2.2e-16 ***
I(Time^2)        3.160   3.160     1 2344.3   2.9679 0.0850663 .
I(Temp^2)       11.013  11.013     1 2087.0  10.3439 0.0013189 **
Nose           292.286 292.286     1 2509.2 274.5376 < 2.2e-16 ***
Time:Temp       14.521  14.521     1 2348.8  13.6389 0.0002266 ***
Time:Nose       23.753  23.753     1 2140.4  22.3110 2.469e-06 ***
Temp:Nose       10.191  10.191     1 2297.0   9.5719 0.0019996 **
I(Time^2):Nose  18.916  18.916     1 2363.5  17.7677 2.590e-05 ***
Time:Temp:Nose   8.031   8.031     1 2304.8   7.5436 0.0060690 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 40: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                      npar  logLik    AIC    LRT Df Pr(>Chisq)
  <none>                17 -4337.0 8708.0
  (1 | OAConsumer)      16 -4653.3 9338.6 632.56  1  < 2.2e-16 ***
  (1 | Day)             16 -4354.0 8739.9  33.90  1  5.804e-09 ***
  (1 | AgeBin)          16 -4340.0 8712.1   6.07  1    0.01376 *
  (1 | OAConsumer:Day)  16 -4375.6 8783.2  77.14  1  < 2.2e-16 ***
  (1 | WDorder)         16 -4344.8 8721.5  15.50  1  8.256e-05 ***

  ---
  Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.7761084, suggesting that the combination of fixed and random effects in the model describe approximately 77.6% of the variance in the square root of Cocoa. Of the fixed-effects terms (Table 5.39), all are highly significant (i.e., p-value <0.001), with the exception of Temperature^2, Temperature:Nose, and Time:Temperature:Nose, which are *only* significant, (p-value <0.05), and Time^2, which is not significant (p>0.05), but remains in the model due to heredity. As for random-effects terms (Table 5.40), all are highly significant (i.e., p-value <0.001), with the exception of (1|AgeBin), which is significant (p-value≤0.05). In terms of the estimated effect on Cocoa, or more specifically the square root of Cocoa, Nose is the largest of all main effects, followed by Temperature. This can be seen when comparing Figures 5.53 and 5.54.

We note in figure 5.53 the previously outlined general trend of Cocoa change with roast, with the lowest estimated Cocoa perception in the region near the raw treatment (i.e., 0 minutes at 24°C), and greatest predicted Cocoa in the uppermost portion of the region near the most extreme roasts (i.e., the combinations 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C), and especially the 54 minutes/151°C treatment. However, as suspected, once aroma can no longer be perceived, the mean value of Cocoa and the difference in its general pattern of change over the experimental region are significant (Figure 5.54). First, noseclips clearly have a large effect of mean Cocoa depression across the experimental region, or put otherwise, the lack of noseclips clearly allows Cocoa to be more readily perceived across the experimental region,

increasing its mean value. Additionally, it is interesting to see that with noseclips, while there is an area of increased Cocoa in the upper-left quadrant of the experimental region, near the 20 minutes/171°C treatment, it is also perceived to be relatively high near the raw treatment (i.e., 0 minutes at 24°C), while the area near 55 minutes/64°C has the lowest estimated Cocoa intensity. While it is not possible to be certain what people were rating as Cocoa while wearing noseclips, one might suspect that it is some other sensation or combination of sensations in the mouth that are present due to dark chocolate. Interestingly, the pattern of increase seen (Figure 5.54) in estimated Cocoa Intensity with noseclips does not match exactly any of the response surface plots for the sensory modalities previously modeled. Perhaps the closest match, however, is to Sweetness (Figure 5.55 below), where some of the general trends of increase and decrease in intensity are similar, and which also makes sense given the relatively, though not exactly, correlated nature of Cocoa and Sweetness in the PCA biplot (Figure 5.43 above) for treatments rated with noseclips. However, the region of least Sweetness and least Cocoa Intensity do not exactly overlap, so even if it is possible that Sweetness may have become somewhat of a stand-in for Cocoa, there likely to be at least one or more other factors also playing a role in Cocoa ratings. Future studies could clarify this matter. Regardless, none of this impacts our results, which quite clearly show a dramatic mean reduction (>50%) in Cocoa when aroma is not present, as was anticipated.

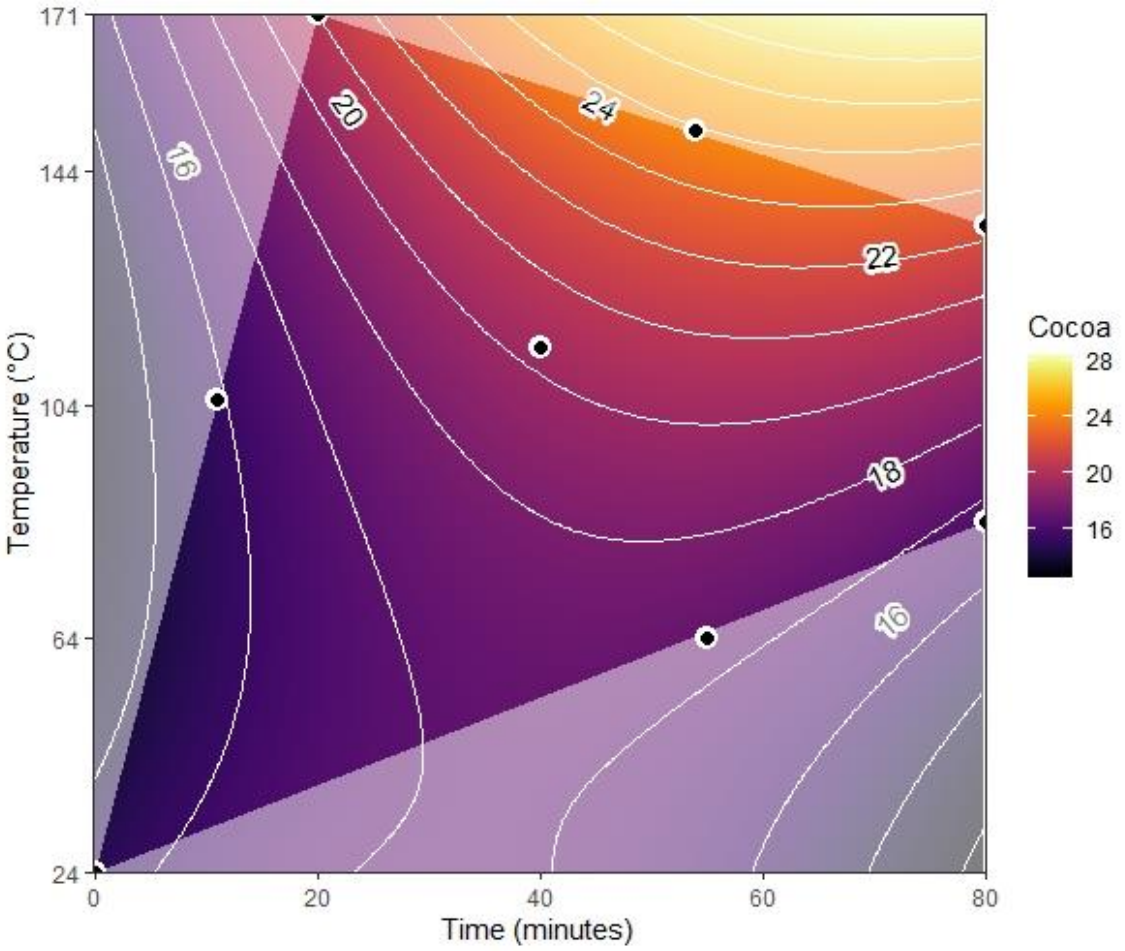

Figure 5. 53: WITHOUT Noseclips GHANA Contour plot for predicted Cocoa for raw and roasted chocolate liquor treatments across the experimental region.

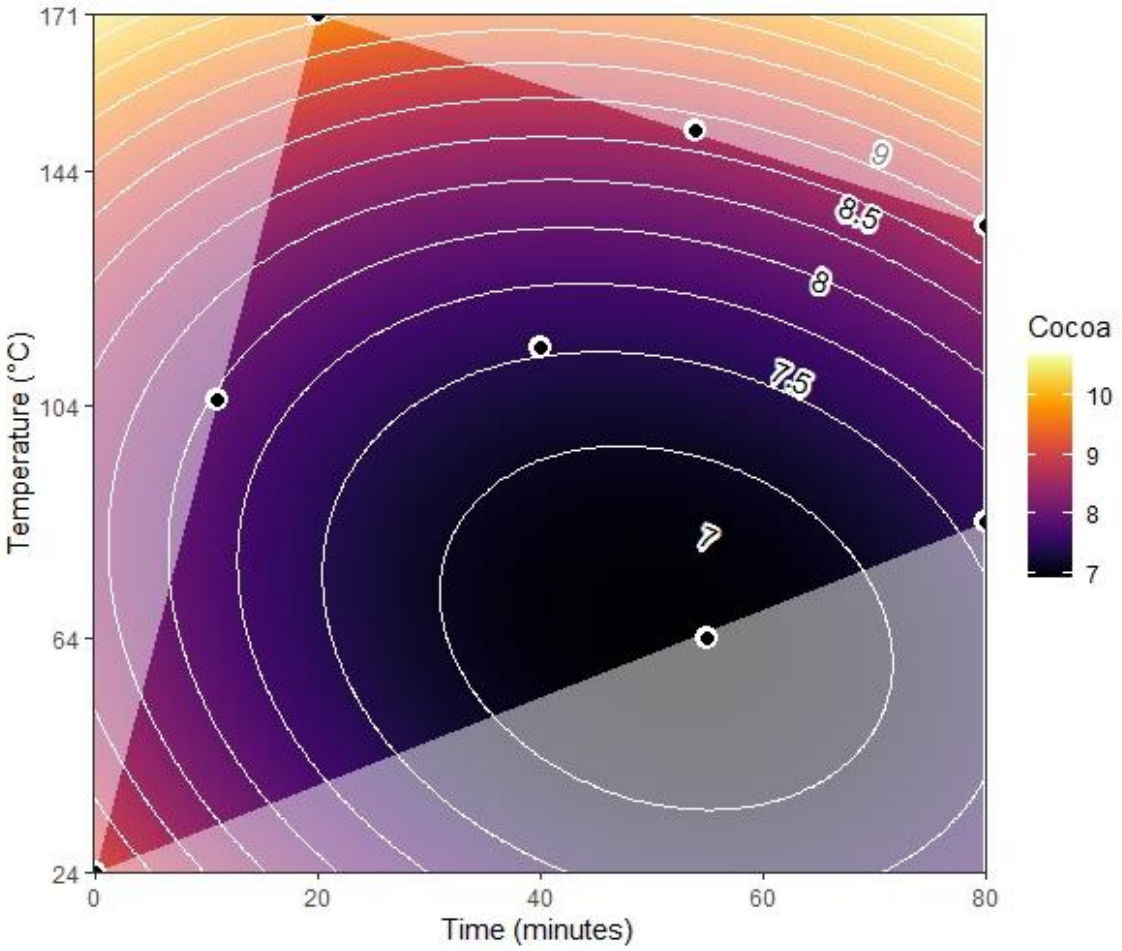

Figure 5. 54: WITH Noseclips Ghana Contour plot for predicted Cocoa for raw and roasted Ghanaian chocolate liquor treatments across the experimental region.

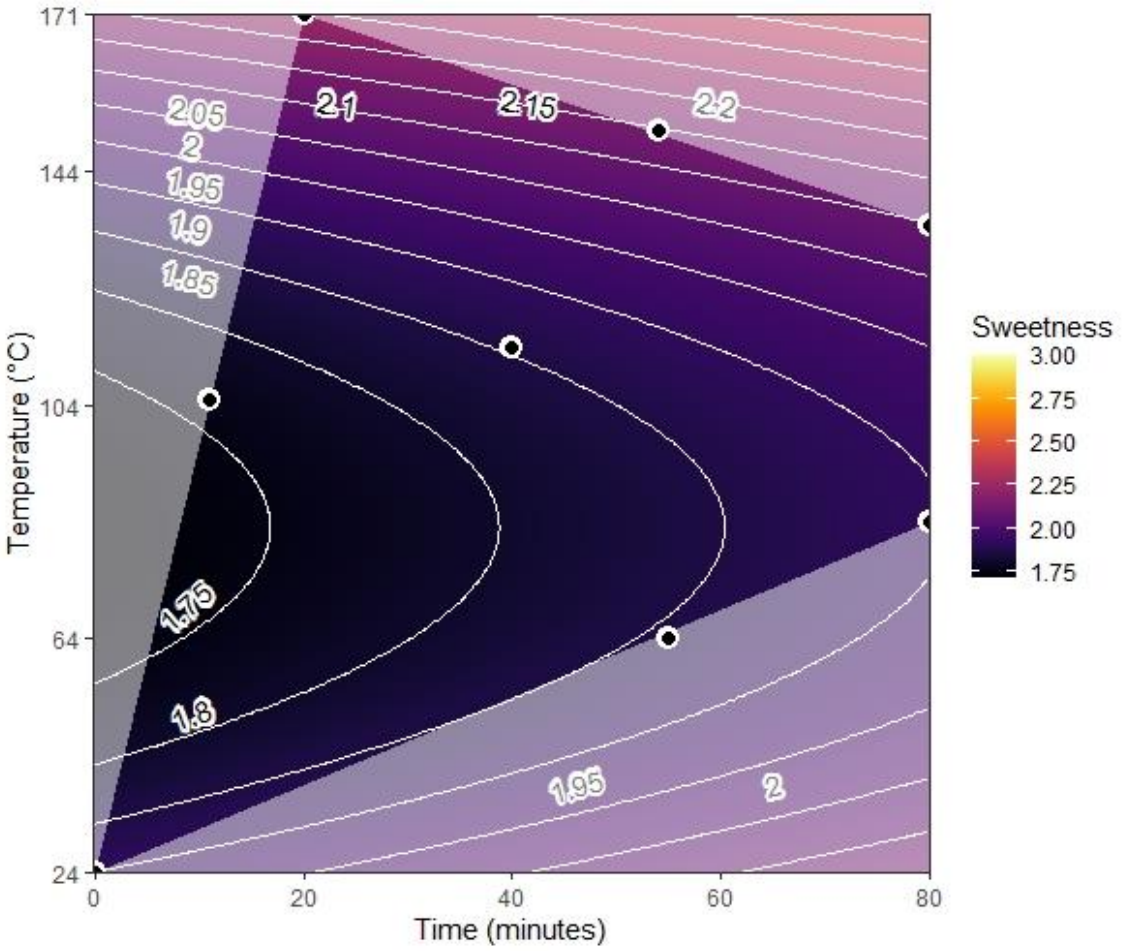

Figure 5. 55: Perceived Sweetness WITH noseclips in Ghanaian chocolate liquor treatments across the experimental region.

5.3.3.5 Liking

Unlike the response surface contour plots above for sensory characteristics of various types (e.g., Bitterness, Astringency, and Cocoa Intensity), Liking is different in that it is a hedonic characteristic, based not upon a single taste, aroma, or physical sensation, but upon subjective consideration of the samples, which undoubtedly is based upon previous experiences and notions about what is good and bad about chocolate generally, and perhaps dark chocolate specifically. Still, we were able to successfully

produce predictive models and response surface contour plots using mixed models that should be useful for helping future researchers to understand the general perception of Liking by consumers as it is related to roast (Time and Temperature) of cacao and the impact of aroma (Nose) on Liking. As before, we used a mixed model of fixed and random effects for this purpose. The model selected for best predicting Liking is:

sqrt(Liking) ~ Time + Temperature + Time:Temperature + $Time^2$ + $Temperature^2$ Nose + Nose:Time + Nose:Temperature + Nose:Time:Temperature + $Nose:Temp^2$ + (1 | OAConsumer) + (1 | OAorder) + (1 | Day) + (1 | AgeBin) + (1|OAConsumer:Day) + (1|WDorder)

The summary of calculated model statistics can be seen in tables 5.41 to 5.43.

Table 5. 41: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

                  Estimate Std. Error         df t value Pr(>|t|)
(Intercept)      1.960e+00  6.253e-02  9.317e+00  31.345 9.21e-11 ***
Time             5.413e-02  1.282e-02  2.268e+03   4.222 2.51e-05 ***
Temp             6.865e-02  8.696e-03  2.264e+03   7.894 4.50e-15 ***
I(Time^2)       -2.429e-02  1.015e-02  2.337e+03  -2.393 0.016771 *
I(Temp^2)       -1.449e-02  1.514e-02  2.225e+03  -0.957 0.338485
Nose1            3.736e-02  1.405e-02  2.355e+03   2.659 0.007892 **
Time:Temp        3.250e-02  8.844e-03  2.370e+03   3.675 0.000243 ***
Time:I(Temp^2)  -3.221e-02  1.429e-02  2.296e+03  -2.254 0.024264 *
Time:Nose1      -3.745e-03  7.876e-03  2.241e+03  -0.475 0.634517
Temp:Nose1      -5.221e-02  8.043e-03  2.268e+03  -6.492 1.04e-10 ***
I(Temp^2):Nose1  1.502e-02  7.295e-03  2.258e+03   2.058 0.039672 *
Time:Temp:Nose1 -2.928e-02  7.739e-03  2.351e+03  -3.784 0.000158 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 42: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

               Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
Time           1.8173  1.8173     1 2267.9 17.8271 2.515e-05 ***
Temp           6.3530  6.3530     1 2264.3 62.3218 4.505e-15 ***
I(Time^2)      0.5840  0.5840     1 2336.6  5.7284 0.0167707 *
I(Temp^2)      0.0934  0.0934     1 2224.6  0.9166 0.3384848
Nose           0.7207  0.7207     1 2354.8  7.0700 0.0078916 **
Time:Temp      1.3767  1.3767     1 2370.0 13.5048 0.0002432 ***
Time:I(Temp^2) 0.5181  0.5181     1 2296.3  5.0824 0.0242637 *
Time:Nose      0.0230  0.0230     1 2240.5  0.2260 0.6345167
Temp:Nose      4.2960  4.2960     1 2267.9 42.1430 1.039e-10 ***
I(Temp^2):Nose 0.4319  0.4319     1 2258.2  4.2368 0.0396722 *
Time:Temp:Nose 1.4594  1.4594     1 2350.8 14.3161 0.0001584 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 5. 43: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                        npar  logLik    AIC    LRT Df Pr(>Chisq)
  <none>                  19 -1163.9 2365.7
  (1 | OAConsumer)        18 -1376.1 2788.3 424.54  1  < 2.2e-16 ***
  (1 | OAorder)           18 -1165.5 2367.1   3.35  1   0.067025 .
  (1 | Day)               18 -1168.1 2372.2   8.51  1   0.003535 **
  (1 | AgeBin)            18 -1166.8 2369.7   5.92  1   0.014958 *
  (1 | OAConsumer:Day)    18 -1181.4 2398.8  35.09  1  3.148e-09 ***
  (1 | WDorder)           18 -1168.2 2372.4   8.63  1   0.003306 **

  ---
  Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

For this model, the pseudo R-squared is 0.5962359, suggesting that the combination of fixed and random effects in the model describe approximately 59.6% of the variance in the square root of the response. Variance Inflation Factors (VIFs) for first-order fixed effects are under 2.5, with the exception of Time, which has a VIF of 3.9, which is still well below the reasonable cutoff of 10, commonly recommended (Montgomery et al., 2012; Zuur et al., 2010).

Of the fixed-effects terms (Table 5.42), Time, Temp, Time:Temp, Nose:Temp, and Nose:Time:Temp are all highly significant (i.e., p-value <0.001), while Nose, $Time^2$, $Time:Temperature^2$, $Nose:Temperature^2$, are significant, (p-value <0.05), and $Temp^2$ and Nose:Time are not significant (p>0.05), but both remain in the model due to heredity. As for random-effects terms (Table 5.43), OAorder is not quite significant (i.e., >0.05 and < 0.10), but is below the default α-value used for choosing random effects values (i.e., 0.10), and the model containing it has an AIC that is approximately equivalent to the model with the lowest AIC value. All other terms are highly significant (i.e., p-value <0.001), or significant (p-value≤0.05).

When the response surface is plotted (Figures 5.56 and 5.57), we note the previously outlined general trend of Liking increasing with roast (Figure 5.56), with the lowest Liking in the region near the raw treatment (i.e., 0 minutes at 24°C), and greatest predicted Liking in the uppermost portion of the region near the most extreme roasts (i.e., the combinations 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C), and especially the 54 minutes/151°C treatment. However, when aroma is removed from the equation through the usage of noseclips, the estimated mean Liking for the entire experimental region increases slightly, though differences in Liking across the experimental region are moderated, making them very similar in areas of roasted and raw treatments. Interestingly, the area in the upper right quadrant is still estimated to be slightly higher in Liking (Figure 5.57). Since that area happens to correspond with the region of least Bitterness, Astringency, and Sourness in the presence of noseclips, given

the negative correlations of Liking with these sensory modalities, as seen in the previous PCA biplot (Figure 5.43), it is not particularly surprising.

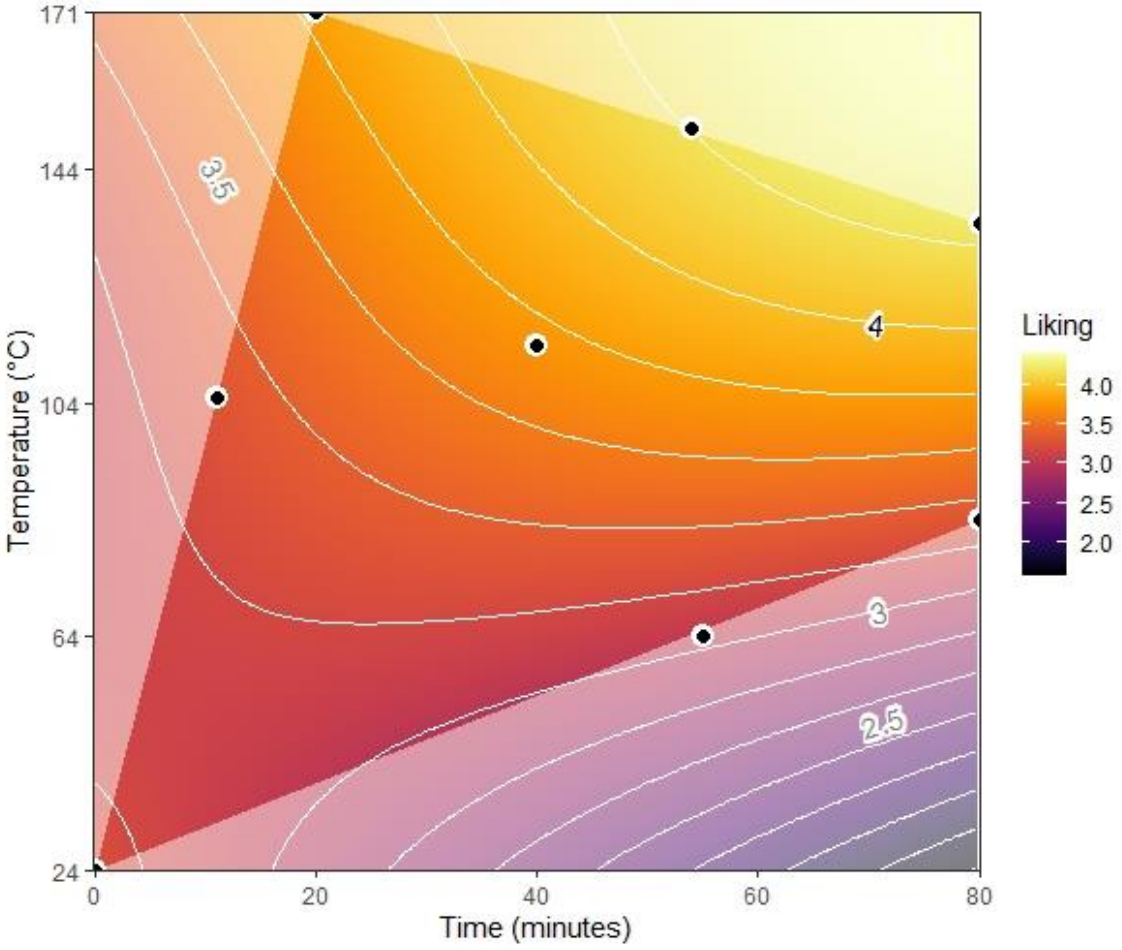

Figure 5. 56: WITHOUT Noseclips GHANA Contour plot for predicted Liking for raw and roasted chocolate liquor treatments across the experimental region.

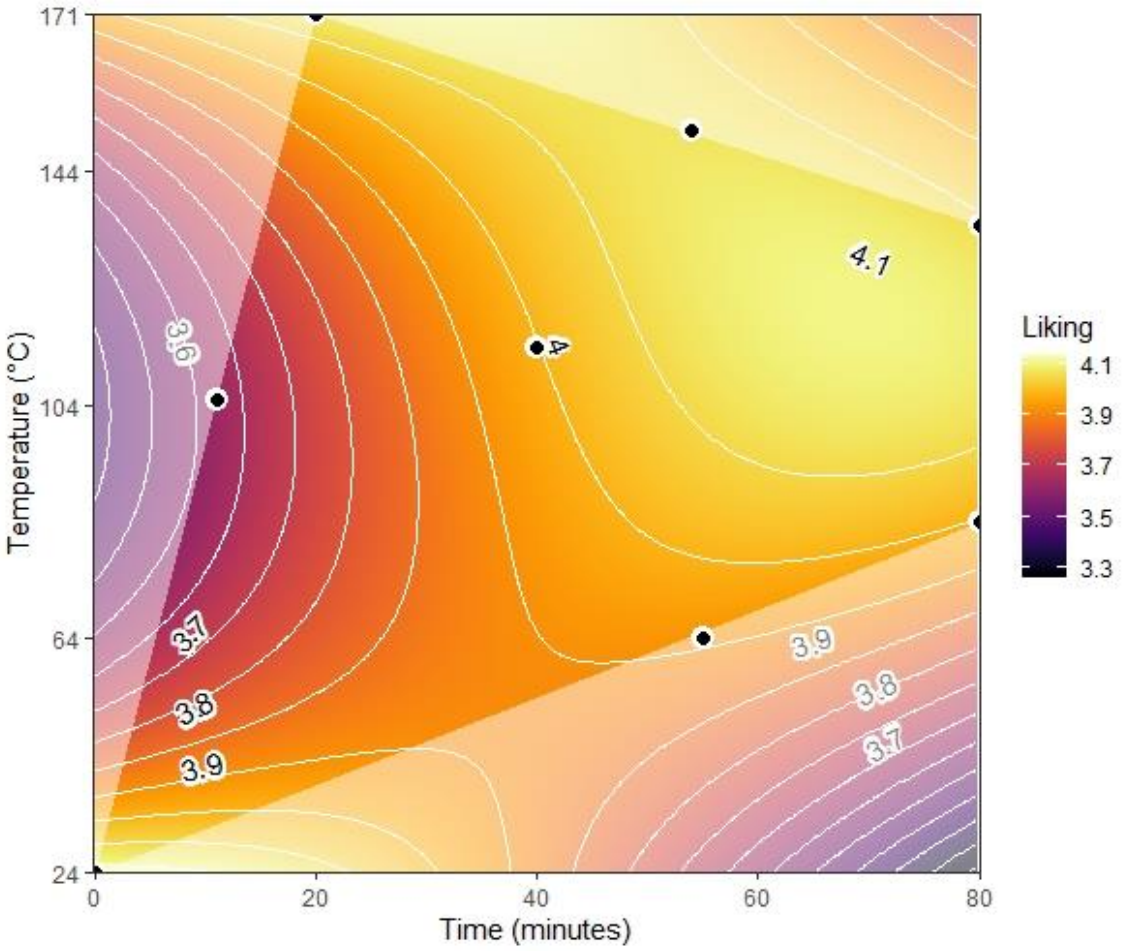

Figure 5. 57: WITH Noseclips Ghana Contour plot for predicted Liking for raw and roasted Ghanaian chocolate liquor treatments across the experimental region.

5.3.4 Results Summary

We have seen that without noseclips the taste modalities of Bitterness, Sourness, and Astringency are all similarly affected as roast progresses with Time and Temperature increases, leading to reductions of different degrees, and showing that in general, the area of the experimental region close to the raw or lightly roasted treatments is expected to be higher in these three characteristics. On the other hand, Sweetness is estimated to increase with roast, as is Cocoa Flavor intensity, and Liking. Given the PCA and preference mapping analysis, none of this is surprising given how positively correlated these three characteristics are, and also given the negative correlation between Liking and the three taste modalities of Bitterness, Sourness, and Astringency. As for Origin differences, though we cannot say that our samples are entirely representative of all cacao that comes from Ghana, Peru, or Madagascar, the characteristics so far discovered help to create a foundation for further study. Our Ghana sample has estimated mean values of Bitterness, Sourness, and Astringency that are lower than the samples of the other origins, and mean values of Cocoa, Sweetness, and Liking that are higher. Additionally, the Marañon cacao from Peru was shown to have the highest estimated Astringency near the raw treatment, but it was reduced substantially during roasting so that it matched more closely the Astringency of the Madagascar cacao. As for Madagascar, it had, by far, the highest estimated values for Sourness, and on average, was the least liked cacao origin.

As expected, noseclips had a significant impact on all characteristics rated, whether sensory (e.g., Bitterness) or hedonic (e.g., Liking). From largest to smallest mean noseclip effect on the reduction of the square root of the response were: Cocoa Intensity, Bitterness, Astringency, Sourness, and Sweetness, with an interesting and significant, but relatively small, mean increase in the hedonic characteristic of Liking. What could account for this increase in Liking? Is it truly the case that consumers would prefer chocolate with no aroma? This is probably not a reasonable supposition, given that humans suffering from the anosmia, a medical condition that results in an inability to smell food or other aromas, have higher rates of depression (Ahn et al., 2016). Rather, it could be that given the strong negative correlations between Liking and the taste modalities of Bitterness, Sourness, and Astringency, and with the dramatic reduction in those three sensory characteristics due to noseclips, many of the more bitter, sour, and astringent treatments are simply slightly more well-liked without aroma present to enhance those characteristics. In fact, it is still the case that the estimated most liked region, even if only marginally so, requires perception of aroma, and is found in the surrounding the treatment levels 80 minutes/135°C, and 54 minutes/151°C.

5.4 Conclusions

Reduction of Bitterness, Sourness, and Astringency is correlated with increased Liking. Choosing raw material naturally low in these characteristics (e.g., Ghana in our study), and optimizing roast can therefore minimize these characteristics and increase Liking. Also, consumers appear to have a preference for increased Cocoa flavor. So, choosing raw materials accordingly (e.g., Ghana in our study), and roasting to maximize this characteristic may be advised. Roast profiles that minimize and maximize each of these characteristics can vary at least slightly by origin, and there is not a complete understanding of this variation within or between origins, but temperature combinations such as 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C, or times and temperatures that would be analogous to these with other roasting equipment, should be considered. Likewise, if maximizing Liking is a consideration, there is not currently support for using raw and lightly roasted cacao (i.e., 0 minutes at 24°C, 11 minutes at 105°C, or 55 minutes at 64°C or similar, based upon specific roasting equipment).

Caveats do also exist, as our study involved analysis of consumer ratings of unsweetened chocolate. Additions of sugar, salt, and other ingredients could introduce significant main and interaction effects that would be relevant, especially given the correlation with Sweetness and Liking in chocolate noted in our models. Additionally, there are other aroma profiles that may be relevant for cacao aside from Cocoa flavor, such as Floral, Fruity or Nutty (Aprotosoaie et al., 2016a), and these may also be

impacted by ingredient addition in a variety of ways (i.e., mixture suppression or enhancement (Lawless et al., 2010b) relevant to Liking. Finally, it is now understood that aroma plays a significant and large role in the perception of not just Cocoa flavor/aroma, but also taste modalities (i.e., Bitterness, Sourness, Sweetness), and the somatosensation Astringency, as well as Liking in chocolate, and this should be an important consideration for any future sensory studies of cacao and chocolate.

CHAPTER 6

6 RELATIONSHIP BETWEEN BITTER COMPOUNDS AND CONSUMER PERCEPTION OF CHOCOLATES MADE FROM THREE ORIGINS OF CACAO USING EIGHT ROAST PROFILES

6.1 Introduction

6.1.1 Cacao and Chocolate

6.1.1.1 Cacao

Also known as cocoa, cacao consists of the fermented and dried seeds of the fruit of the tropical *Theobroma cacao* tree in the *Malvaceae* family (Aprotosoaie et al., 2016a). Cacao is a significant food commodity, with annual global consumption reaching approximately 4.6 million metric tons as of 2018 with an increase in demand of 3.9% over 2017 (Barchart, 2019), and plays an important role as the primary ingredient in the manufacture of chocolate (Aprotosoaie et al., 2016a).

6.1.1.2 Chocolate

Chocolate is a usually sweetened, uniquely flavored, solid paste that melts smoothly at human body temperature due to the presence and unique fatty acid composition of cacao fat, called cocoa butter (Aprotosoaie et al., 2016a). The aforementioned unique chocolate aroma is due mostly to the presence of cacao, required in American-made chocolate by FDA Title 21 (2019). Prior to transformation into chocolate, cacao is roasted to obtain a more complex flavor and character that is generally preferred by consumers over that of raw cacao (Aprotosoaie et al., 2016a).

6.1.2 Bitterness and its Sensation and Perception

6.1.2.1 Bitterness

Bitterness is one of the five taste modalities (i.e., salty, sweet, sour, bitter, umami) sensed by the tongue (Gaudette et al., 2013; Keast et al., 2003b). It has been noted that toxic compounds are often bitter (Keast et al., 2003b), and the ability of humans to taste bitter substances is likely to have evolved as a form of toxin detection (Keast et al., 2003a). This may explain why bitterness is generally disliked by humans (Drewnowski et al., 2000; Fischer et al., 2005) and even rejected in most food (Gaudette et al., 2013), despite famous exceptions such as coffee, beer, red wine, and dark chocolate (Gaudette et al., 2013; Keast et al., 2003b; Roy, 1997) that highlight the sometimes complex nature of human food choices (Gaudette et al., 2013).

6.1.2.2 Bitterness Sensation

Bitter sensation begins as a ligand/taste molecule contacts a bitter receptor cell, and bitter taste transduction results, with nerve impulses being sent to the brain (Drewnowski, 2001; Lawless et al., 2010b). Bitter taste receptor cells are clustered as part of taste buds, located on the papillae across the tongue, with more located on the palate, and in the throat (Drewnowski, 2001; Lawless et al., 2010b). Of the basic tastes (i.e., bitter, sweet, sour, salty, umami), bitter is the most complex (Drewnowski, 2001), and it is now known that in humans there are approximately 25 subtypes of G-protein-

coupled receptors called TAS2Rs which are responsible for the transduction of bitter taste from many thousands of compounds (Dagan-Wiener et al., 2018; Maehashi et al., 2009). Each bitter taste sensation is complex and unique, with four properties to define it: quality, intensity, temporal patterns and spatial patterns (Keast et al., 2003b).

6.1.2.3 Bitterness Perception

Bitterness perception starts with the sensation of bitter compounds, but also includes processing by the brain of incoming signals from other sensory modalities (i.e., other tastes, aromas, and somatosensory, aural, and visual inputs) (Lawless et al., 2010a). In short, bitterness perception is a combination of sensation and central cognitive effects resulting from other concurrently processed sensory information (Keast et al., 2003b). For example, aural stimulation (i.e. music) (Carvalho et al., 2017) can even affect bitterness perception.

6.1.2.4 Variation in Bitterness Sensation/Perception

There is a great deal of variation across human individuals regarding sensation of bitterness (Drewnowski, 2001; Mennella et al., 2005; Negri et al., 2012), and the ability to sense certain bitter compounds at all, which can be inherited (e.g., phenylthiocarbamide (PTC) and 6-n-propylthiouracil (PROP) (Drewnowski, 2001)). Genotype therefore leads to a large variance between individuals regarding sensitivity to bitter compounds, where approximately 1/3 of people are unable to sense certain bitter compounds, while others, so-called supertasters, show hypersensitivity to bitter compounds, due in part to larger numbers of taste buds, with all other individuals falling

upon a spectrum between these extremes (Lawless et al., 2010b). Additionally, children avoid bitter foods to a larger extent than do adults (Negri et al., 2012), and sex-based differences have also been described (Bartoshuk et al., 1994).

Variation in bitterness perception depends on more than genotype, age, and sex, however. It is an interesting case that perception of overall bitterness intensity of a mixture of bitter compounds at known concentrations is less than that of the sum of intensities of the individual compounds at the same concentrations (Keast et al., 2003b). It is also the case that bitterness may be suppressed by certain compounds with sweet, salty, and umami tastes, and enhanced by those with sour ones (Calviño et al., 1993; Drewnowski, 2001; Fischer et al., 1994). Short-term variance in perception also exists, when exposure to bitter compounds can lead to adaptation (i.e., decreased responsiveness) to bitterness (Lawless et al., 2010b), and bitterness can even be altered by the concentration of calcium ions in an individual's saliva (Neyraud et al., 2004).

6.1.3 Known Important Bitter Compounds in Cacao

6.1.3.1 Methylxanthines

It is well known that the bitter compounds theobromine and caffeine are present in cacao at levels above their sensory detection thresholds (Bonvehí et al., 2000; Stark et al., 2006). However, the ratio of theobromine to caffeine can change substantially depending upon cacao variety (Bonvehí et al., 2000; Trognitz et al., 2013; Ziegleder, 2009) covering a ratio range from approximately 1 to 6 or more (Trognitz et al., 2013). Still, theobromine is present in cacao at a higher concentration than caffeine (Stark et

al., 2006), which one might assume makes it a more important contributor of bitterness in cacao, though this assumption is complicated by the much lower solubility of theobromine than caffeine in near-neutral aqueous solutions (Spiller, 1997; Stark et al., 2006). Furthermore, any potential differences in solubility kinetics between the two compounds are likely to be important for overall bitterness, as chocolate passes through the mouth only briefly, and such a comparison has not been published as far as this author is aware. However, theobromine is known to dissolve quite slowly (Wadsworth, 1921), an observation also noted by this author, suggesting the possibility that a given amount of theobromine in chocolate could result in a lower perceived bitterness than that of an equivalent amount of caffeine. This does not even consider the potential difference in bitterness in equimolar solutions of the two compounds due to possible variation in bitterness receptor binding.

6.1.3.2 Flavan-3-ols

Flavan-3-ols are the main polyphenolic compounds in cacao, even when origin and varietal are considered (Oracz et al., 2015). These healthful cacao flavonoids (Aprotosoaie et al., 2016b), which are known to be bitter (Stark et al., 2006), are present in cocoa at levels above their sensory detection threshold, especially the monomer (-)-epicatechin (Natsume et al., 2000; Stark et al., 2006; Ziegleder, 2009), and its epimer (+)-catechin (Kothe et al., 2013), and oligomers such as procyanidin B2, which are also present above their taste threshold concentrations (Stark et al., 2006). However, approximately 35% of the polyphenol content of cacao is made up of the (-)-epicatechin monomer alone (Cooper et al., 2007). Interestingly, epicatechin appears to

decrease during roasting, while catechin and some larger molecular weight polymers of epicatechin increase (Kothe et al., 2013; Stanley et al., 2018), as do melanoidins (Quiroz-Reyes et al., 2018), themselves containing phenolic compounds (Morales et al., 2012). Overall, changes in these compounds during roasting appear to lead to an overall reduction in bitterness and astringency (Aprotosoaie et al., 2016a).

6.1.3.3 2,5-Diketopiperazines

2,5-Diketopiperazines (DKPs) are cyclic dipeptides that may be found in certain foods and beverages (Borthwick et al., 2017; Ginz et al., 2000), and that in the case of cacao, are created predominantly during roasting, and have a bitter taste (Bonvehí et al., 2000; Stark et al., 2006). The mechanism of formation is heat-induced cyclization of linear alpha-amino-acid oligomer precursors (i.e., peptides) (Rizzi, 1989). It has been claimed that the bitter taste of DKPs is augmented in the presence of theobromine (Borthwick et al., 2017), an interaction (Drewnowski, 2001) resulting in a characteristic cocoa-like bitterness (Pickenhagen et al., 1975). A variety of DKPs have been found in varying amounts in cacao after having undergone various roasting treatments (Bonvehí et al., 2000; Rizzi, 1989; Stark et al., 2006) as well as in a variety of other roasted foods (Bonvehí et al., 2000; Gautschi et al., 1997; Ginz et al., 2000; Roudot-Algaron et al., 1993). It has been claimed that the DKP cyclo(L-Pro-L-Val) is the most important bitter compound in roasted cocoa, based upon cacao sourced from Ghana with an unknown roast treatment (Stark et al., 2006), though this has apparently not yet been verified by

further research, nor for a variety of roast profiles or cacao origins/populations beyond the single sample of Ghana.

6.1.4 Causes of Bitterness Variation in Cacao

Factors that impact bitterness in cacao, and therefore chocolate, are numerous, including varietal, growing conditions (e.g., hours of sunlight and rainfall), ripeness at harvest, and post-harvest processing (Kongor et al., 2016). Geographical location, even within a single country, appears to play a role in overall methylxanthine concentration and theobromine to caffeine ratio, probably due to altitude of the specimens from which the cacao is derived (Carrillo et al., 2014). However, cacao varietal, ripeness at harvest, and the multi-stage post-harvest processing (i.e., fermentation and drying) have perhaps been most commonly noted as playing important roles (Afoakwa et al., 2008; Aprotosoaie et al., 2016a; Beckett et al., 2017; Kongor et al., 2016).

6.1.4.1 Varietal

Regarding cacao varietal, while ongoing research has led to a better understanding of the cacao genome and genetic variation in cacao generally (Johnson et al., 2009; Motamayor et al., 2008; Takrama et al., 2014; Zhang et al., 2011; Zhang et al., 2012), the sensory characteristics of the 10 identified genetic clusters of cacao (Motamayor et al., 2008) have yet to be thoroughly characterized, though an ongoing sensory evaluation project at Penn State will likely result in useful data on this topic (Hopfer & Brown, 2018, personal correspondence). Certainly, it is already clear that specific varietals show variation in bitterness (Kongor et al., 2016), with theobromine concentration (Bonvehí

et al., 2000) and ratios of theobromine to caffeine being significantly different across origins/varietals/populations (Timbie et al., 1978; Trognitz et al., 2013; Zoumas et al., 1980). Additionally, there appear to be differences in the DKP concentrations of the roasted cacao prepared from fermented and dried cacao of different origins (Rizzi, 1989).

6.1.4.2 Ripeness

Ripeness of the cacao fruit at the time of harvest is known to play an important role in cacao flavor in general (Afoakwa et al., 2008) and in bitter methylxanthine concentration in particular (Bonvehí et al., 2000; Timbie et al., 1978) as caffeine and theobromine both increase substantially between 12 weeks from flowering and fruit harvest at 5-6 months (Senanayake et al., 1971). Cacao flavonoid profile, including flavan-3-ols, also depends upon ripeness (Nazaruddin et al., 2006; Rusconi et al., 2010), as unripe cacao has 29% less epicatechin than ripe cacao (Payne et al., 2010).

6.1.4.3 Post-Harvest Processing

6.1.4.3.1 Fermentation

While post-harvest processing, which consists of fermentation and drying, is a well-known cause of bitterness reduction in cacao (Afoakwa et al., 2008; Aprotosoaie et al., 2016a; Beckett et al., 2017), it is a complex one with many unknown factors of its own (John et al., 2019). It is known that during cacao fermentation in the presence of oxygen and polyphenol oxidase, flavonoid oxidation and polymerization occur, leading to browning, and reduction of bitterness and astringency (Ziegleder, 2017). Additionally,

fermentation may result in an up to 40% reduction in bitter methylxanthines due to their diffusion through the cacao seed coat and loss in fermentation runoff/waste (Bonvehí et al., 2000), and flavonoids are also likely to be lost in runoff (Cooper et al., 2007). It has also been noted that the ratio of catechin to epicatechin shifts during fermentation, increasing during the process (Porter et al., 1991), likely due to increases in (-)-catechin as (-)-epicatechin concentration decreases (Cooper et al., 2007; Payne et al., 2010). Overall, there is a >80% decrease in total catechin monomers (i.e., epicatechin and catechin) during fermentation (Payne et al., 2010).

6.1.4.3.2 Drying

Well-managed drying practices result in further reductions in bitterness via additional flavonoid oxidation (Barišić et al., 2019; Ziegleder, 2009). Interestingly, and perhaps unexpectedly, very low concentrations of certain DKPs have been found in fermented and dried, but *unroasted*, cacao (Rizzi, 1989).

6.1.4.4 Roasting

Roasting, considered by some to be the most important step in processing cacao (Aprotosoaie et al., 2016a), results in the creation of bitter diketopiperazines (DKPs) from peptides (Rizzi, 1989; Ziegleder, 2017), and darker roasts, particularly at higher temperatures, appear to increase DKP levels the most (Bonvehí et al., 2000), whereas unroasted cocoa contains virtually no DKPs (Bonvehí et al., 2000). However, roasting is required to alter the harsh, unpleasant flavor of raw cacao, and to develop characteristic cocoa aroma (Ziegleder, 2009). Additionally, roasting alters the concentrations of

epicatechin, and its epimers and oligomers (Kothe et al., 2013; Stanley et al., 2018), compounds which are both bitter and astringent (Stark et al., 2006), sometimes in unexpected ways seemingly related to varietal (Kothe et al., 2013). Loss of epicatechin at temperatures over 70°C occurs, and at a roasting temperature of 120°C catechin content has been seen to increase by approximately 650% in fermented cacao (Payne et al., 2010). This is probably due to epimerization of epicatechin to catechin (De Taeye et al., 2014; Oracz et al., 2015; Payne et al., 2010) and decomposition of procyanidins (Oracz et al., 2015), probably first to epicatechin followed by epimerization (Zhu et al., 2002). For this reason, the epicatechin/catechin ratio helps to understand the previous processing of cacao (Payne et al., 2010). Interestingly, roasting also leads to production of high molecular weight melanoidins that bind to polyphenols such as catechin, epicatechin, and procyanidin B2 (Oracz et al., 2019). As for the epimerization of epicatechin during roasting, it appears to be from (-)-epicatechin to (-)-catechin, even though the naturally present form of catechin is (+)-catechin (De Taeye et al., 2014; Hu et al., 2016; Hurst et al., 2011; Payne et al., 2010). Cooper et al. (2007), was the first to show this (-)-epicatechin to (-)-catechin epimerization due to cacao processing, using a chiral column with UPLC (Cooper et al., 2007), and this was later confirmed by Kothe et al. (2013). Cooper et al. (2007) also showed that there is a strong linear relationship between epicatechin and procyanidins, as well as (+)-catechin, so that they can all be predicted from the epicatechin concentration, while (-)-catechin, however, cannot be predicted in the same fashion (Cooper et al., 2007). Finally, it has been known for decades that a small loss of methylxanthines from the cotyledons to the cacao shell is

expected during roasting, apparently increasing with degree of roast (Wadsworth, 1922).

6.1.5 Further Context

While this research project began due to a significant interest in understanding the relationship between roasting (i.e., application of heat at specific temperatures for specific lengths of time) of cacao and consumer perception of chocolate made therefrom, especially as it pertains to Bitterness and Liking, we also wished to relate the concentrations of bitter compounds in cacao and chocolate with the Bitterness perception of consumers of that very same chocolate. It is primarily for this reason that we collected quantitative chemical data on six of the most important bitter compounds in chocolate (Stark et al., 2006) for each treatment (see Chapter 4 for full details) that would be assessed by consumers (see Chapter 5 for full details). This desire even drove our decision making, in part, regarding the origin selection of cacao from Ghanaian, Peruvian, and Madagascan farms, in that we had hoped there would be significant variation in sensory and chemical characteristics between these three geographically and likely highly genotypically distinct (Fang et al., 2014; Motamayor et al., 2008; Takrama et al., 2014; Zhang et al., 2011) cacao origins, so as to minimize the chances of too many large correlations between the various compounds, making successful analysis more probable. Analyzing the treatments in this context is particularly interesting, because it means considering the Bitterness of certain compounds within the context of actual chocolate samples, a food that appears to contain 40,000 or more chemical compounds (Milev, Patras, Dittmar, Vrancken, & Kuhnert, 2014), making it the most

complex food yet investigated, even prior to roasting (Milev et al., 2014). And with any food so complex, there are bound to be significant interactions between perceived aroma, taste, and somatosensory characteristics that could lead to intriguing findings. We have already mentioned some of these findings in previous chapters. For example, the perception of chocolate aroma is responsible for an incredibly large and significant increase in the perception of all sensory characteristics measured, including Bitterness (Chapter 5). But other surprises could await. For example, there are known natural correlations between the increase and decrease of certain bitter compounds in chocolate (Cooper et al., 2007); how would these relate to changes in Bitterness? With a wealth of potential in the overall dataset based upon the 4970 observations of 8 sensory and hedonic characteristics of those 8 treatments per origin, we hope that despite the likelihood of correlations of specific bitter compounds in the data causing complications, that our findings will add enough new insight to the field in the more accurate understanding of Bitterness in cacao. As for our analytical approach, it consists of several methods, with details further described in the Results & Discussion section of this chapter: PCA- and PLS-based methods, and Mixed-Model Linear Regression. These methods are quite a common part of the modern sensory analysis toolbox (Lê et al., 2014) with mixed-model linear regression and type III ANOVA, specifically, being part of the “classic bread and butter technique” (Lawless, 2013; Næs et al., 2010).

6.2 Materials & Methods

6.2.1 Materials

6.2.1.1 Chemicals

High-purity standard compounds were purchased from several manufacturers (Table 6.1).

Table 6. 1: Standard compound information

Compounds	Manufacturer	Lot #	Part #	Purity Factor
Theobromine	Sigma Aldrich	BCBM9560V	T4500	0.987
Caffeine	Sigma Aldrich	MKBW1243V	W222402	1.000
(+)-Catechin*	Sigma Aldrich	WXBC5812V	C1251	0.868
(-)-Epicatechin	Sigma Aldrich	BCBW4134	E1753	0.951
Procyanidin B2	Cayman	0506392-7	0506392-7	1.000
cyclo(Pro-Val)	BAChem	1026889	1026889	0.999

*COA: (+)-Catechin hydrate 99%, adjusted for water weight (12.35%) to 86.8%

The following solvents were obtained from Fisher Scientific: HPLC Water Lot# 189570, HPLC Methanol Lot# 187803, 88% Formic Acid Lot# 180909, 0.5 M NaOH in methanol Ricca Lot# 4802L81, and 0.5 M HCl Ricca Lot# 4803D64. The following solvent was obtained from Millipore Sigma: N,N-Dimethylformamide Lot# SHBL1229.

Mobile phase mixtures used for HPLC analysis consisted of an aqueous phase (Channel A), containing 1140 µL 88% formic acid, and 50 mL methanol, made to volume in a 1 L volumetric flask with HPLC water, and an organic phase (Channel B), containing 1140 µL 88% formic acid, made to volume in a 1 L vol flask with HPLC methanol.

Solvent for standard mixtures (SMS) and sample extraction solvent (ES) consisted of 700 mL N,N-Dimethylformamide mixed with 340 mL HPLC water, which gives an

approximate volume of 1000 mL once the mixed solvent temperature is at equilibrium with room temperature.

6.2.1.2 Additional Analytical Materials & Equipment

HPLC analysis was performed with an Agilent 1200 series HPLC with Chemstation for LC Rev. B.02.01 [244] software, with Kinetex C18 and Zorbax Eclipse XDB-C18 columns, both 4.6 mm x 250 mm, run in sequence, with a photodiode array detector (DAD).

Ghanaian chocolate liquor prepared from roasted cacao was obtained from Patric Chocolate (Columbia, MO, USA).

Cacao liquor was ground with a whirling-blade coffee grinder (Hamilton Beach model #80335R), and extraction carried out using an ultrasonic water bath (Kendal model # HB-S-23DHT), vortexer (Vortex-Genie model # 12-812-v1), motorized pipette filler (RPI Corp model # 248646) paired with borosilicate 25 mL volumetric pipette, and a centrifuge with 6-slot fixed-angle rotor (Eppendorf model #5804, rotor model # F-34-6-38). 50 mL conical centrifuge tubes (Celltreat #229421) were also used. Subsamples of ground chocolate liquor and all standard compounds were massed with an analytical balance (DeltaRange model # AG204).

6.2.1.3 Cacao

All cacao from each of three origins (i.e., Madagascar, Ghana, Peru), was fermented and dried at origin before being packed into approximately 65 kg burlap sacks, themselves packed into steel shipping containers, and then transported to cacao warehouses in the United States by cargo ship.

6.2.1.3.1 Madagascar:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from three separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total (i.e., 7 kg per lot): Lots #2182432 (2018 harvest), #2672711 (2018 harvest), #1427425 (2018 harvest). This cacao was obtained from Guittard Chocolate (Burlingame, CA) Upon receipt, all three lots of this origin were blended into a homogenous composite, and hand-sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.), unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro (Concord, MA) Supergrain Premium RT bags with high vapor and gas barrier characteristics at <65% RH and <27 °C until roasted (approximately one month or less).

6.2.1.3.2 Ghana:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from three separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total (i.e., 7 kg per lot): Lot #19003223 (2018/2019 main harvest), #482019 (2018/2019 main harvest), #3729 (2017/2018 main harvest). This cacao was obtained from Guittard Chocolate (Burlingame, CA). Upon receipt, all three samples of this origin were blended into a homogenous composite, and hand-sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.),

unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro (Concord, MA) Supergrain Premium RT bags with high vapor and gas barrier characteristics at <65% RH and <27 °C until roasted (approximately one month or less).

6.2.1.3.3 Peru:

Once in the United States, cacao was stored in warehouses in standard conditions until sampling. Cacao was sampled across multiple bags per lot, from two separate lots of at least 12.5 MT each, with each lot consisting of cacao from a large number of trees. 21 kg of cacao was sampled in total from two separate lots covering multiple months as follows: 7 kg was sampled from a lot consisting of the January through early March 2018 harvest, and 14 kg (7 kg x 2) was sampled from a lot consisting of the late March to late June 2018 harvest. This cacao was obtained from Marañon Cacao (San Diego, CA). Upon receipt, all three samples of this origin were blended into a homogenous composite, and hand-sorted to remove dust, broken shell and beans, multiple bean clusters (i.e., doubles, triples, etc.), unfilled beans, and foreign objects such as leaves, stones, or burlap twine. Composited cacao was stored in sealed Grainpro (Concord, MA) Supergrain Premium RT bags with high vapor and gas barrier characteristics at <65% RH and <27 °C until roasted (approximately one month or less).

6.2.1.3.4 Cacao Quality Parameters

Cacao quality parameters related to phenotype, ripeness, fermentation, and drying, of a representative portion (180 g) of each 21 kg composite are detailed in Table 6.2.

Measurements were made just prior to conducting all roasting treatments. All three samples fall within acceptable parameters for good quality cacao according to the International Cocoa Organization (ICCO) (2015).

Table 6. 2: **Preliminary statistics on composite samples from each origin**

Origin:	Ghana	Madagascar	Peru
Bean-Count:	89 beans per 100 g	86 beans per 100 g	74 beans per 100 g
Slaty Color:	1%	0%	0%
Fully Purple:	1%	0%	26%
Partly Purple:	17%	6%	18%
Total Slaty or Purple:	19%	6%	44%
Light Brown:	0%	27%	7%
Medium Brown:	58%	62%	27%
Dark Brown:	23%	5%	10%
Total Brown:	81%	94%	44%
White Beans:	0%	0%	12%
Moisture Content:	6.13 %	6.68%	5.83%
Internal Mold:	0%	0%	0%
Infested:	1%	0%	1%

6.2.1.4 Cacao-Processing-Related Equipment

The roasting oven consisted of a Binder GmbH (Tuttlingen, Germany) model # FD56 120 V precision laboratory oven with forced air convection. This model of oven was also used for holding chocolate liquor in a just-melted state prior to the sensory analysis experiment. Roasting temperatures were monitored and recorded with an Omega (Norwalk, CT) model #RDXL4SD 4-channel datalogger with Type-K thermocouples (part #5sc-tt-k-30-36). Room temperature and humidity during roasting were monitored with

a SensorPush (Brooklyn, NY) #HT1 device with humidity calibrated with Boveda (Minnetonka, MN) Low-RH One-Step 32% calibration kit model #CAL32-SM, and temperature and humidity verified with a Bacharach (New Kensington, PA) Sling Psychrometer model #0012-7012.

Preliminary cacao cooling was performed by a Lasko (West Chester, PA) 20" box fan model #B20200, placed under ¼" stainless mesh bolted to Cambro (Huntington Beach, CA) food bin. Cacao cracking was performed with a custom CrankandStein (Atlanta, GA) 305 mm 3-roll cocoa cracker, and winnowing was performed using a custom food-grade winnower with 4" clear FDA-approved for food contact PVC, with air flow powered by a Grizzly Industrial, Inc (Springfield, MO) model #:G0441 3HP cyclone.

Chocolate liquor was ground using a Spectra 11 (Tamil Nadu, India) Stone Wet Grinder model #101, and strained through a Kitchenaid (Benton Harbor, MI) model # KES161OHOBW fine-mesh strainer. Cooled and solidified chocolate liquor was wrapped in aluminum foil and vacuum sealed in FoodSaver (Oklahoma City, OK) brand multi-layer vacuum bag with nylon vapor barrier. Portioning of chocolate liquor for the sensory evaluation experiment was carried out using positive displacement pipettes (Eppendorf (Hamburg, Germany) Repeater M4 pipette and Eppendorf Repeater E3 pipette) with disposable pipette tips (Eppendorf Combitips Advanced Lot# I183098J).

Food-grade ingredients were used for training calibration solutions prior to chocolate liquor assessments, consisting of tannic acid obtained from Spectrum Chemicals (New

Brunswick, NJ), citric acid and caffeine obtained from Sigma-Aldrich (St. Louis, MO), and C&H (Crockett, CA) Pure Cane Sugar (sucrose) obtained from a local grocery store.

6.2.2 Methods

6.2.2.1 Analytical Chemical Methods

6.2.2.1.1 Standard Curve Preparation

Stock solutions of all compounds were prepared by quantitatively transferring accurately massed quantities of standard compounds into a single 100 mL volumetric flask and making to volume with SMS (section 6.2.1.1) as per the amounts listed in Table 3.2. Calculated stock concentrations, adjusted to account for standard purity, are included (Table 6.3).

Table 6. 3: Stock solution information for each compound

Compounds	Expected Concentration (mg/g) in cacao*	Expected Concentration (mg/mL) in 50 mL extraction of 1 g cacao	Target Stock Concentration (mg/mL)	Accurate Mass in 100 mL Vol Flask (mg)	Purity-Adjusted Stock Concentration (mg/mL)
Theobromine	11.4519445	0.22903889	0.3	30.1	0.297
Caffeine	1.01328342	0.020265668	0.1	10.0	0.100
(+)-Catechin	0.68617464	0.013723493	0.1	10.4	0.090
(-)-Epicatechin	2.50000938	0.050000188	0.15	14.9	0.142
Procyanidin B2	1.20505716	0.024101143	0.1	10.0	0.100
cyclo(Pro-Val)	1.7423075	0.03484615	0.1	11.0	0.110

***Based upon (Stark et al., 2006)**

After stock preparation was complete, stock was diluted in SMS to prepare an external standard curve consisting of 5 approximately equidistant dilutions for each standard compound (Table 6.4).

Table 6. 4: Standard curve pure compound concentrations

Compounds	1 (mg/mL)	2 (mg/mL)	3 (mg/mL)	4 (mg/mL)	5 (mg/mL)
Theobromine	0.297	0.249	0.200	0.152	0.104
Caffeine	0.084	0.067	0.051	0.035	0.020
(+)-Catechin	0.061	0.046	0.032	0.018	0.0005
(-)-Epicatechin	0.096	0.072	0.050	0.028	0.001
Procyanidin B2	0.067	0.051	0.035	0.020	0.0010
cyclo(Pro-Val)	0.074	0.056	0.038	0.022	0.0005

Prepared external standard solutions in amber-tinted 2 mL HPLC autosampler vials were analyzed via HPLC. Mobile phase (section 6.2.1.1), was purged with high-purity helium for 15 minutes. The system was then purged of mobile phase already present in the lines for 5 minutes at 4 mL/min to avoid potential issues with gas bubble formation. Inlet frits in the mobile phase containers were also visually checked at this time to be sure that no air bubbles were trapped behind them. Initial conditions for the HPLC method (i.e., 0.75 mL/min flow rate at 100% Channel A) were then run for at least 20 minutes until backpressure had stabilized. Finally, the HPLC method was run in full, consisting of a gradient (Table 6.5) with 10-minute post-time for re-equilibration of the columns at 100% Channel A conditions.

Table 6. 5: HPLC gradient for standard quantification

Time (min)	% Channel A	% Channel B
0	100	0
25	70	30
50	50	50
55	0	100
Post (10 min)	100	0

A 0.75 mL/min flowrate was maintained throughout, with column temperature held at 30 °C. Injection volume was 10 µL, with post-injection methanol needle wash between injections. All wavelengths between 190 nm and 300 nm were monitored at 1 nm increments. Ultraviolet absorbances of all standards were ultimately used for quantification purposes using either 205 nm or 280 nm wavelengths (Table 6.6). An initial blank was run to assure that separation conditions would be the same for each standard injection. Prepared standards were run in order from least concentrated solution to most concentrated solution for each compound.

Table 6. 6: Wavelengths used for quantification of standard compounds

Compounds	UV Wavelength (nm)
Theobromine	280
Caffeine	280
(+)-Catechin	205
(-)-Epicatechin	280
Procyanidin B2	205
cyclo(Pro-Val)	205

Once the peak area related to UV absorbance at specific wavelengths was determined for each peak of each compound (Table 6.6), standard curves were plotted with known analyte concentration on the x-axis and peak size on the y-axis, and the line describing each curve derived as a mathematical equation allowing for the estimation of the concentration of each compound in each sample. Coefficients of determination (R^2) were also calculated for each curve, with values greater than 0.99 being deemed acceptable (tables 6.7 through 6.13).

Table 6. 7: Standard curve equations and R^2 values

Analyte	Standard Curve Equation	Coefficient of Determination (R^2)
Theobromine @ 280 nm	y = 34690x + 548.01	0.9952
Caffeine @ 280 nm	y = 32039x + 20.99	0.9996
(-)-Epicatechin @ 280 nm	y = 9489.7x - 1.5051	0.9999
(+)-Catechin @205 nm	y = 148382x + 276.06	0.9947
Procyanidin B2 @ 205 nm	y = 108918x + 40.658	0.9999
Cyclo(Pro-Val) @ 205 nm	y = 25110x + 14.485	0.9987

Y= Area count and X= analyte concentration (mg/mL)

Table 6. 8: Theobromine standard curve concentrations and area counts

Concentration mg/mL	Area (280 nm)
0.297	**10632.0**
0.249	**9339.0**
0.200	**7691.8**
0.152	**5803.2**
0.104	**4015.2**

Table 6. 9: Caffeine standard curve concentrations and area counts

Concentration mg/mL	Area (280 nm)
0.084	2682.0
0.067	2204.1
0.051	1659.5
0.035	1145.9
0.020	650.7

Table 6. 10: (-)-Epicatechin standard curve concentrations and area counts

Concentration mg/mL	Area (280 nm)
0.096	908.0
0.072	680.4
0.050	469.0
0.028	266.8
0.001	13.5

Table 6. 11: (+)-Catechin standard curve concentrations and area counts

Concentration mg/mL	Area (205 nm)
0.061	9027.2
0.046	7227.2
0.032	5225.5
0.018	3114.2
0.0005	72.5

Table 6. 12: Procyanidin B2 standard curve concentrations and area counts

Concentration mg/mL	Area (205 nm)
0.067	7368.5
0.051	5595.9
0.035	3849.5
0.020	2270.0
0.0010	114.7

Table 6. 13: cyclo(Proline-Valine) standard curve concentrations and area counts

Concentration mg/mL	Area (205 nm)
0.074	**1867.0**
0.056	**1402.1**
0.038	**1026.0**
0.022	**561.2**
0.0005	**14.58**

6.2.2.2 Roasting Experimental Design

For the roasting experimental design (Table 6.14), temperature range (i.e., 24°C to 171°C) and time range (i.e., 0 to 80 minutes) were chosen as reasonable ranges for potential modification of bitterness based upon literature regarding common roasting temperatures (Afoakwa et al., 2008; Ziegleder, 2017) and the advice of chocolate professionals; this includes a raw treatment at 24°C (approximate room temperature) for 0 minutes as a control. Combinations of time and temperature that are impossible, judged so extreme so as to be burnt, or that in essence are repeats of other combinations (e.g., 0 minutes and any temperature aside from 24°C, or 24°C and any time combination aside from 0 minutes), were excluded. All time and temperature combinations are repeated for each of the three origins (Madagascar, Ghana, and Peru). Roast order was randomized. The specific time and temperature combinations within these ranges were chosen as part of a Response Surface Methodology (RSM) approach, using JMP 14.0.0 software (Cary, NC) and an I-Optimal algorithm, which minimizes average variance of prediction for potential model coefficients (Jones et al., 2012; Myers

et al., 2016; Oyejola et al., 2015), while at the same time seeking to minimize covariance of model coefficients (Oyejola et al., 2015). The resulting design is an irregularly shaped non-rotatable design, akin to a central composite design (CCD), with a duplicate centerpoint to allow for pure error estimation. In addition to the previously mentioned reasons, this model was chosen because standard CCDs have already been commonly and successfully used in roasting optimization experiments (Farah et al., 2012; Kahyaoglu, 2008; Lee et al., 2001; Madihah et al., 2012; Mendes et al., 2001; Özdemir et al., 2000), while it has also been noted that computer-based optimal design algorithms are effective at helping to choose designs for complex problems that cannot easily be solved with standard design types, such as in the example of irregular experimental regions (Goos et al., 2011; Jones et al., 2012).

Table 6 1: Randomized modified I-optimal experimental design for roasting

Minutes	Temperature °C	Origin	Randomized Roast Order
40	114	Peru	1
40	114	Ghana	2
54	151	Madagascar	3
80	84	Peru	4
0	24	Peru	5
20	171	Ghana	6
80	84	Ghana	7
20	171	Madagascar	8
54	151	Peru	9
55	64	Peru	10
40	114	Peru	11
0	24	Ghana	12
40	114	Ghana	13
40	114	Madagascar	14
11	105	Peru	15
55	64	Madagascar	16
11	105	Madagascar	17
54	151	Ghana	18
11	105	Ghana	19
40	114	Madagascar	20
80	135	Peru	21
80	135	Ghana	22
0	24	Madagascar	23
80	84	Madagascar	24
55	64	Ghana	25
80	135	Madagascar	26
20	171	Peru	27

6.2.2.3 Roasting and Winnowing Method

For roasting and winnowing, 410 g of each cacao origin, specified by the experimental design roast order (Table 6.14), was weighed into each of two stainless steel mesh roasting trays (820 g total). The cacao was spread into an evenly distributed single-bean

layer on each tray. A thermocouple was placed in a hole made with a steel needle in a bean of approximate average size in the middle of each tray. The thermocouples were then connected to the datalogger, and the datalogger was set to record. The oven was preheated to the setpoint required in the roasting design (Table 6.14), with the convection fan on. Once the setpoint was reached, 10 additional minutes were allowed to elapse to assure preheating completion. Once pre-heating was completed, the timer was set to the required roasting time (Table 6.14), and the two trays were loaded into the oven with one tray right above the fan and the other right below it. The oven door was immediately closed (i.e., in less than approximately 5 seconds) and the timer was started.

Once the roast was complete, the oven was turned off, and the door was immediately opened. Forced-air cooling of the roasted cacao was employed to minimize carry-over roasting through immediate winnowing (i.e., breakage and removal of cocoa shell), augmented with preliminary fan cooling for higher temperature roasts (i.e., roasts of greater than 114°C in Table 5.2), to attain sufficient temperature decreases for all treatments (i.e., a drop to 50°C or less within 5 minutes). The end-result of this cooling and winnowing process, when augmented with manual removal of any remaining visible shell pieces, was room temperature cacao nibs, the main ingredient of chocolate liquor.

Each roast was completed in duplicate on the same day, and nibs from the roasted and winnowed duplicate treatments blended until homogeneous. Therefore, 1640 g of cacao was roasted in total for each treatment. Of this 1640 g, 1000 g of blended roasted

cacao nibs were allocated for chocolate liquor production and placed in a food-grade polypropylene bag, with air substantially removed, and labeled with all identifying information. The remaining 640 g were saved in case of future need. Nibs were stored in a room with temperature below 19.5 °C with relative humidity (RH) at approximately 40% or less and turned into chocolate liquor within 48 hours.

6.2.2.4 Chocolate Liquor Production Method

For chocolate liquor production, the stone bowl and grinding stones of the Spectra 11 wet grinder were preheated to approximately 49°C to 54°C overnight in a warming cabinet. Once the stone grinder components were mounted to the machine, it was immediately turned on and 1000 g of nibs were added slowly, over a period of 20 minutes, one large spoon (20 ml) at a time until a thick paste began to form with the lid then being replaced, until the paste became less viscous. Nibs were added in this way until the paste began to appear glossy and less viscous, allowing for faster addition. Once all nibs were added, the internal elements of the wet grinder were scraped with a silicone food-grade spatula, and this scraping process was repeated at 30-minute intervals three additional times, for four scrapings total. The chocolate liquor was refined to a smooth texture for a total of 8 hours from the time that all nibs were added to each batch at a temperature of approximately 50°C, maintained by production of frictional heat from the refining process itself. The final chocolate liquor was then poured through a fine-mesh culinary strainer, into a poly food-storage container using a silicone spatula to remove as much liquor as possible from the wet grinder. A yield of

approximately 900 g was achieved in all cases. The chocolate liquor was then covered with an air-tight lid, placed in a storage room at or below 19.5 °C with RH% at approximately 40% or less. All chocolate liquor batches were allowed to solidify at storage temperature, then unmolded, separated into two approximately equivalent portions—one each for chemical and sensory analysis—with each wrapped individually with an aluminum sheet, vacuum sealed in a multi-layer vacuum bag with nylon vapor barrier, and stored at or below 19.5 °C with RH% at approximately 40% or less until they were prepared for further analysis (i.e., within approximately 60 days for the chemical analysis, and within approximately 90 days for the sensory analysis).

6.2.2.5 Sensory Analysis Method

6.2.2.5.1 Chocolate Liquor Sample Sensory Preparation

Chocolate liquor samples were shipped on October 21st, 2019 during cool weather from Columbia, Missouri to State College, Pennsylvania. The transit took three days, and the intact package was received immediately upon delivery, the morning of October 24th, 2019 by Sensory Evaluation Center staff in the Food Science department at The Pennsylvania State University.

The samples were stored unopened at a cool room temperature below approximately 20°C for six days until October 30th, 2019, two days prior to final preparation of the sensory experiment, when the chocolate liquor samples were transferred to labeled, wide-mouth quart glass jars, sealed by covering the mouth of the jars first with aluminum foil followed by canning jar lids with rings, which were tightened to disallow

air exchange (Figure 6.1). Then, on the same day, all sealed jars were placed in a warming oven (Binder GmbH (Tuttlingen, Germany) model # FD56 120 V precision laboratory oven with forced air convection) and held at 44°C with convection for 24 hours to melt, before having the temperature lowered to 39°C on Thursday, October 31, 2019 for another 24 hours to maintain the melted state of the liquor while minimizing any changes of flavor characteristics.

Figure 6. 1: Warming oven filled with jars of solidified chocolate liquor with tight lids

On Friday, 11/01/2019, the day that final sample preparation began, the work room was at 44% relative humidity (RH) and approximately 22.8°C as measured by a Bacharach

(New Kensington, PA) red-spirit sling psychrometer model #12-7012. As the viscosity of the chocolate liquor appeared somewhat high, which would make portioning consumer samples most difficult, the holding temperature of the oven was increased to 42°C in the morning of that same day. Portioning of chocolate liquor was carried out using precision positive displacement pipettes (Eppendorf Repeater M4 pipette and Eppendorf Repeater E3 pipette), with disposable pipette tips (Eppendorf Combitips Advanced Lot# i183098j), with pipettes set to 300 µL in order to obtain equal sample quantities of approximately 0.3 g in the shape of small chocolate disks, the mass of which was verified each time a new pipette tip was used (i.e., each time a new chocolate liquor sample was sampled). These 0.3 g quantities were deposited on parchment paper on a metal tray, and then chilled in a commercial refrigerator (TRUE Manufacturing, Inc (O'Fallon, MO) Refrigerator serial# 1-4596848) with thermostat set to approximately 4°C for approximately 1 hour until the chocolate liquor disks had solidified. The chocolate disks were then transferred to labeled stainless steel food storage containers covered with double-thick heavy-duty aluminum foil prior to being returned to the refrigerator to maintain the texture of the disks until testing in order to have a consistent mouthfeel and appearance (figures 6.2 and 6.3).

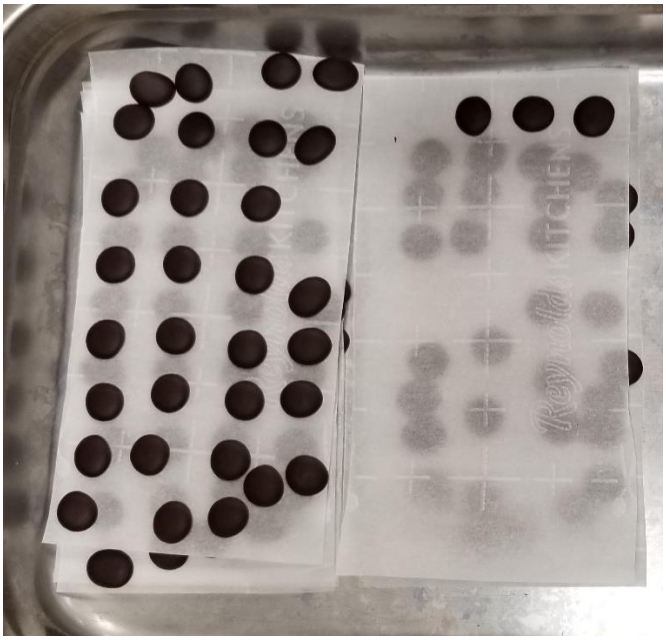

Figure 6. 2: Hotel pan filled with chocolate liquor disks

Figure 6. 3: Labeled hotel pans filled with chocolate liquor disks in a refrigerator

All 27 samples to be analyzed were prepared in this way over the course of approximately 36 hours from Friday, 11/1/2019 through 11/2/2019. As the first section of the study was fully recruited according to the experimental design (see Table 6.14 below) with 160 consumers, an excess of disks totaling approximately 200 for each of the 27 samples were prepared. After all disks were prepared for the first section of the sensory study, the warming ovens were turned off and all jars of liquor were cooled back to room temperature. Just prior to consumer analysis on the morning of Monday, November 4th, 2019, first day of the trial, approximately 8 disks of prepared chocolate liquor were transferred into each of 4 small, labeled plastic soufflé cups with lids (Figure 6.4) per sample, for all 27 samples. These cups were kept in the refrigerator until needed, and only one cup of each sample # was removed for immediate use. These cups were prepared fresh daily to avoid excessive scalping of aroma.

Figure 6. 4: Soufflé cups with pre-portioned chocolate-liquor disks for each day

The chocolate liquor sample preparation method described above was repeated identically for Section 2 of the sensory study to prepare sufficient samples to carry out the relevant experimental design in full (Table 6.15), likewise starting chocolate liquor disk preparation 3 days prior to the start of the trial, with only the total number of chocolate disks prepared being adjusted according to the design (i.e., an excess of 240 disks for each sample, as opposed to 200 for Section 1).

6.2.2.5.2 Consumer Evaluation Method

The sensory evaluation experiment took place in two sections. Section 1 consisted of a consumer evaluation study where all 27 chocolate liquor samples (i.e., all treatments across all 3 origins of cacao) were presented across 5 days without noseclips, allowing all consumers to taste and smell the samples in a way that is analogous to normal food consumption. Section 2, which took part 14 days later, required consumers to asses a subset of 9 of the 27 samples (i.e., all of the roasting treatments for the Ghanaian cacao only) across 3 days while wearing noseclips to block sensory input related to volatile aroma compounds, so that the interaction between aroma and all other sensory characteristics that were being measured could potentially be estimated for the Ghanaian samples. 145 consumers (aged 18-65, 38 males) were recruited for Section 1 (160 minus 15 participants who ultimately did not participate), and 100 consumers (aged 18-65, 25 males) were recruited for section 2 (108 minus 8 participants who ultimately did not participate) from an in-house database of interested consumers. All consumer assessment took place in the Sensory Evaluation Center (SEC) within the Food

Science Department of The Pennsylvania State University in University Park, PA. The SEC consists of a commercial-grade food preparation area, used for sample assembly, a meeting area, used to check in consumers, and a testing room, used for product assessments. The testing room consists of 12 individual tasting booths equipped with computerized data collection via Compusense Cloud software (Academic Consortium, Guelph, ONT), and sample serving hatches through which all samples are passed to the participants. Lighting is adjustable, and red light was used for both sections of the study to mask color differences between the samples.

Consumers were scheduled for their sample evaluation in blocks of 12 at a time, 20-min apart. Each testing day, consumers evaluated 5 different samples and also underwent a brief sensory training on the first testing day for both Sections 1 and 2, prior to completing assessments. Small disks (0.3 g) of chocolate liquor (see section 6.2.2.5.1 above for preparation details) were served on small plastic tasting spoons on placemats showing the sample code and order of assessment (see figures 6.7 and 6.8 below). Research procedures were exempted from institutional review board review by the Penn State Office of Research Protections under the wholesome foods exemption in 45 CFR 46.101(b) (protocol number 33164). Consumer participants were compensated for their time according to the IRB protocol, and in Section 1 were given $5 incentives on days 1 and 2, $10 incentives on days 3 and 4, and $15 on day 5 for completing the entire trial. Consumers in Section 2 were given $5 incentives on days 1 and 2, and a $10 incentive on day 3 for completing the entire trial. Samples were assessed by consumers

according to the sensory evaluation designs developed for Sections 1 and 2 (see Tables 6.14 and 6.15 below).

6.2.2.5.3 Consumer Evaluation Experimental Design

6.2.2.5.3.1 Consumer Evaluation Experimental Design

Section 1

For Section 1, a Williams 27-present-27 design was created, consisting of 27 rows and 27 columns, assuring that each of the 27 samples would be presented in the design exactly 27 times, resulting in a design with positional and pairwise balance, which was important for achieving a near-balanced design for first-order carryover to the extent possible given various constraints. Because 160 consumers were initially recruited to taste 5 samples per day for each of 5 days, 4000 slots (160 x 5 x 5) were needed in the design. Since the Williams 27 design has only 729 slots (27 x 27) the design was repeated as needed to fill up the additional slots, therefore the full design was repeated approximately 5.5 times (4000 / 729), meaning that in the full 5-day, 160-consumer study, each sample would be tasted an average of 148 times (5.5 x 27) by an average of 104 consumers, with an average of 35 consumers tasting each sample two or more times. Summary statistics of this design, including the exact number of times each sample appears, and the number of consumers intended to taste each sample is shown in Table 6.14.

Table 6. 14: Consumer sensory design Section 1 summary statistics

Sample #	# Times Sample to be Tasted	# People to Taste 1 Time Only	# People to Taste 2 Times	# People to Taste 3 Times	# People to Taste 1 or more Times
1	151	65	25	12	102
2	151	69	26	10	105
3	151	64	27	11	102
4	151	63	29	10	102
5	150	68	29	8	105
6	148	72	23	10	105
7	148	76	21	10	107
8	148	77	19	11	107
9	148	75	23	9	107
10	148	73	27	7	107
11	148	77	25	7	109
12	148	80	25	6	111
13	148	80	25	6	111
14	148	75	26	7	108
15	148	70	27	8	105
16	147	69	28	6	104
17	146	70	27	6	104
18	145	69	26	8	103
19	145	66	26	9	101
20	145	71	25	8	104
21	145	71	25	8	104
22	146	68	27	8	103
23	147	68	26	9	103
24	147	62	24	11	98
25	149	59	25	12	97
26	151	59	25	14	98
27	153	58	28	13	99
Averages	**148**	**69**	**26**	**9**	**104**
Std Dev	**2**	**6**	**2**	**2**	**4**

6.2.2.5.3.2 Consumer Evaluation Experimental Design Section 2

For Section 2, a Williams 9-present-9 design was created, consisting of 9 rows and 9 columns, assuring that each of the 9 samples would be present in the design exactly 9 times, resulting in a design with positional and pairwise balance, which was important for achieving a near-balanced design for first-order carryover to the extent possible given various constraints. Because we recruited 108 consumers to taste 5 samples per day for each of 3 days, 1620 slots (108 x 5 x 3) were needed in the design. Since the Williams 9 design has only 81 slots (9 x 9) the design was repeated as needed to fill up the additional slots, therefore the full design was repeated approximately 20 times (1620 / 81), meaning that in the full 3-day, 108-consumer study, each sample would be tasted 180 times (20 x 9) by a total of 108 consumers, with 72 consumers tasting each sample twice. The exact number of times each sample appears in the design, and the number of consumers intended to taste them both appear in Table 6.15.

Table 6. 15: Consumer sensory design Section 2 summary statistics

Sample #	# Times Sample to be	# People to Taste 1 Time	# People to Taste 2 Times	# People to Taste 1 or more Times
1	180	36	72	108
2	180	36	72	108
3	180	36	72	108
4	180	36	72	108
5	180	36	72	108
6	180	36	72	108
7	180	36	72	108
8	180	36	72	108
9	180	36	72	108
Averages	**180**	**36**	**72**	**108**

It may be noted that the average number of people who were to taste each sample in Section 1 is 104 (Table 6.14), which is quite close to the number of people who were to taste each sample in Section 2 (i.e., 108) (Table 6.15). This was intentional, so that estimates of the effect of noseclips in Section 2 would be based upon ratings from approximately the same number of people as in Section 1 where samples were rated without noseclips.

6.2.2.5.4 Recruitment of Consumers for the Study

The SEC at The Pennsylvania State University, where the sensory portion of this study was conducted, maintains a secure electronic list of previous and potential participants for consumer and other sensory trials. This meant that many consumers receiving the recruitment email had previously been part of other consumer sensory trials, and even chocolate-related ones, and were familiar with the process of tasting samples and responding to questions using a computerized system (Compusense Cloud, Academic Consortium, Guelph, ONT). For recruitment, a screener was sent out to the SEC database to screen for eligible participants with regards to gender, age, food allergies, existence of taste defects, current medication, pregnancy or breastfeeding status, whether they smoked, had swallowing issues, or mouth piercings. Additionally, the screener included questions about chocolate preferences and frequency of consumption for several types of chocolate or chocolate-containing products. Table 6.16 shows the responses that were required for a consumer participant to be accepted into the study.

Table 6. 16: Recruitment requirements for this consumer study

Qualifying Characteristic	Requirement
Gender	Any
Age	18-65 years inclusive
Allergies	No allergies noted
Taste Defects	No taste defects
Medications	No medications taken
Pregnant	Not pregnant
Breast Feeding	Not breast feeding
Smoker	Non-smoker
Swallowing Problems	No swallowing problems
Mouth Piercings	No mouth piercings
Frequency of Milk Chocolate Consumption	Once a month or more frequently
Frequency of Dark Chocolate Consumption	Once a month or more frequently
Frequency of Chocolate Praline Consumption	Once every 3 months or more
Chocolate Preference	Either milk or dark chocolate

In this way, 160 consumers were recruited for Section 1 of the study, and 108 consumers were recruited for Section 2 of the study.

6.2.2.5.5 Consumer Assessment Daily Process and Ballot

6.2.2.5.5.1 Section 1

On day one of the Section 1, 5-day study, consumers signed in at the front desk of the SEC and were given a tag with a number designating their sample set number, from 1 to 160. This number was noted in a sign-in manual and this specific code was not changed for the remainder of the 5-day trial. Once the consumer participant had received their numbered tag, they were shown into the assessment room, and picked one of the available empty individual tasting booths, with red lighting turned on. Once the consumer had signed into the Compusense software, they were asked to turn off and put away their cellphone, to be courteous to fellow participants, and to pass their

numbered tag through the hatch in front of them, whereupon they digitally provided Implied Informed Consent, and received an explanation about the number of tests in the section (i.e., one per day for 5 days), compensation amounts (i.e., consumers in section 1 were given $5 incentives on days 1 and 2, $10 incentives on days 3 and 4, and $15 on day 5 for completing the entire trial), that they would evaluate 5 dark chocolate samples per day on each of the 5 days, and that they would first receive training on rating-scale usage and attribute-rating calibration (i.e., sweet, sour, bitter, astringent, chocolate flavor). Participants were then asked to read the instructions on each page and respond to the best of their ability. The training then began.

For the training, participants were first given an explanation of how to use the generalized Labeled Magnitude Scale (gLMS) in this study as follows:

> "The scale you will use today starts at *'no sensation'* (NS) on the left and ends at the *'strongest sensation of any kind'* on the right.
>
> Think of what the strongest sensation is for you and remember it throughout this session. The top of the scale should not change, regardless of the sensation or quality you are rating.
>
> Adjectives are placed along the scale. You should use these to help make your ratings, but feel free to click anywhere along the scale.
>
> When you are using the scale be sure to separate how *intense* something is from how much you *like or dislike* it. For example, if something is weakly bitter and you dislike bitter, don't be tempted to rate it as more bitter because you don't like it.

Finally, remember that *the top is the strongest sensation of any kind which represents the most intense sensation you might experience across any type of sensation*. What you perceive to be the strongest sensation of any kind should be consistent across items. It is very important that the same sensation is at the top of the scale for each sensation you rate."

Next, the scale training began, with questions asked as seen in figure 6.5.

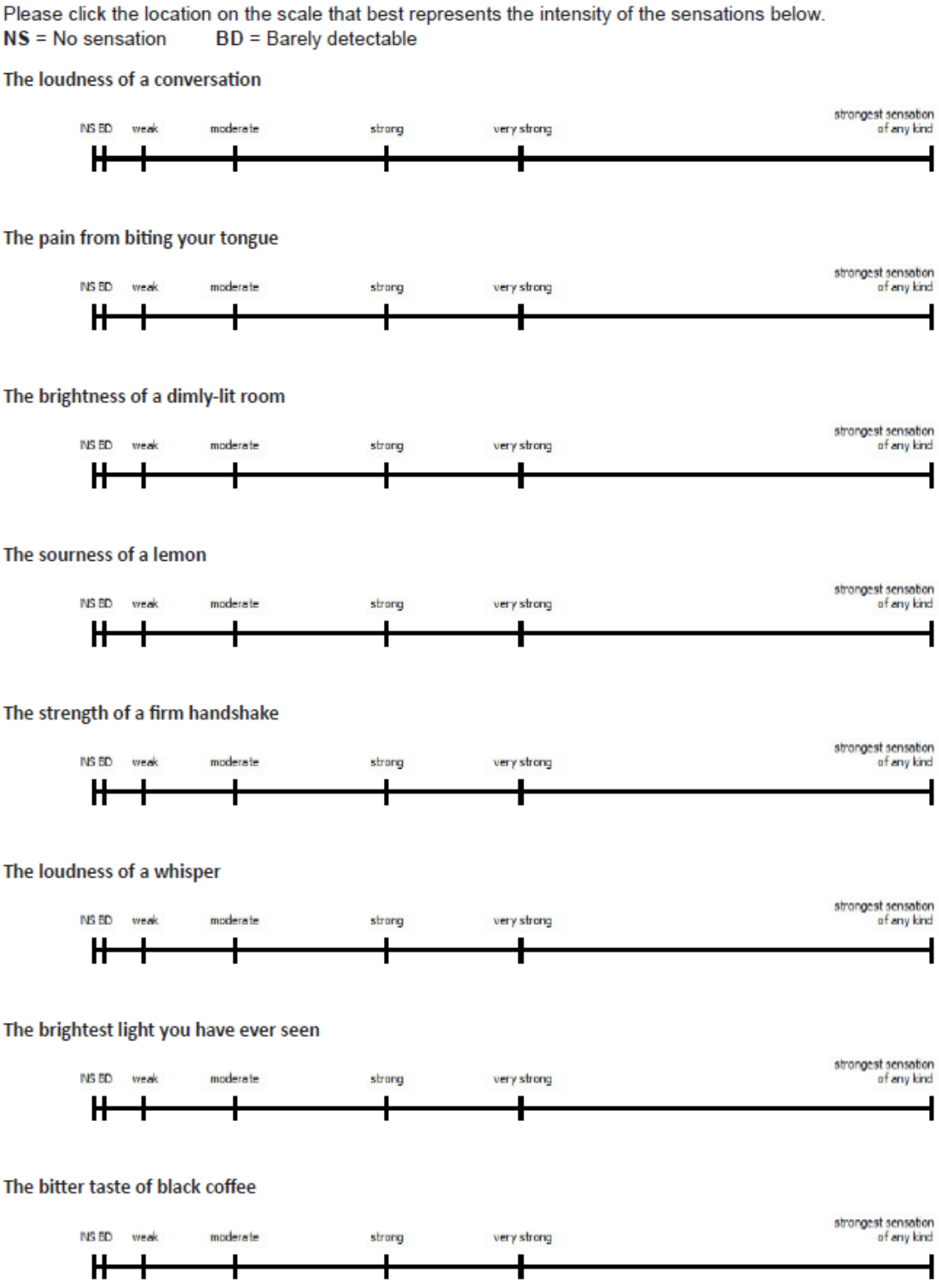

Figure 6. 5: Scale training questions with the generalized Labeled Magnitude Scale (gLMS).

Next in the training, a set of calibration samples was served to the participants to familiarize them with the attributes that they would be rating in the chocolate liquor samples. These attributes were, in order: astringent, sour, bitter, sweet, and Cocoa/Dark Chocolate. Each training solution was provided in a small, coded medicine cup (Figure 6.6), except for the cocoa/dark chocolate attribute which was described verbally only. Concentrations and descriptions of the calibration solutions are found in Table 6.17.

Figure 6. 6: Training calibration solutions

Table 6. 17: Training calibration solutions used in the sensory evaluation.

Calibration Name	Concentration and Preparation	Description
Astringent	41.5 g/L tannic acid in DI water	astringent (puckering, drying, roughing) sensations
Sour	1.5 g/L citric acid in DI water	sour sensations
Bitter	1.0 g/L caffeine in DI water	bitter sensations
Sweet	30 g/L sucrose in DI water	sweet sensations
Cocoa/Dark Chocolate	Verbal description only. No solution.	The intensity and richness of deep dark chocolate and cocoa flavors. For example, a piece of dark chocolate, or the smell of freshly baked chocolate brownies

Participants were asked to taste a solution monadically, swish it in the mouth for 5 seconds, expectorate, and consider that it characterized a particular named sensation, for example, “astringency (puckering, drying, roughing sensation),” and then they were asked to cleanse their palate with water, expectorate once more, and then move on to the next calibration sample. Each consumer panelist proceeded in this way through each characteristic until all solutions were tasted. The final part of this training was for the “cocoa/dark chocolate” attribute. Rather than tasting a sample, the participants were asked to think of the “intensity and richness of deep dark chocolate and cocoa flavors. For example, “a piece of dark chocolate, or the smell of freshly baked chocolate brownies.” At this point, the participants passed back their training samples, were reminded once more how to use the rating scale for which they had already been trained, and then they were passed a cafeteria tray containing the five coded samples that they would taste for the day’s session (figures 6.7 and 6.8)

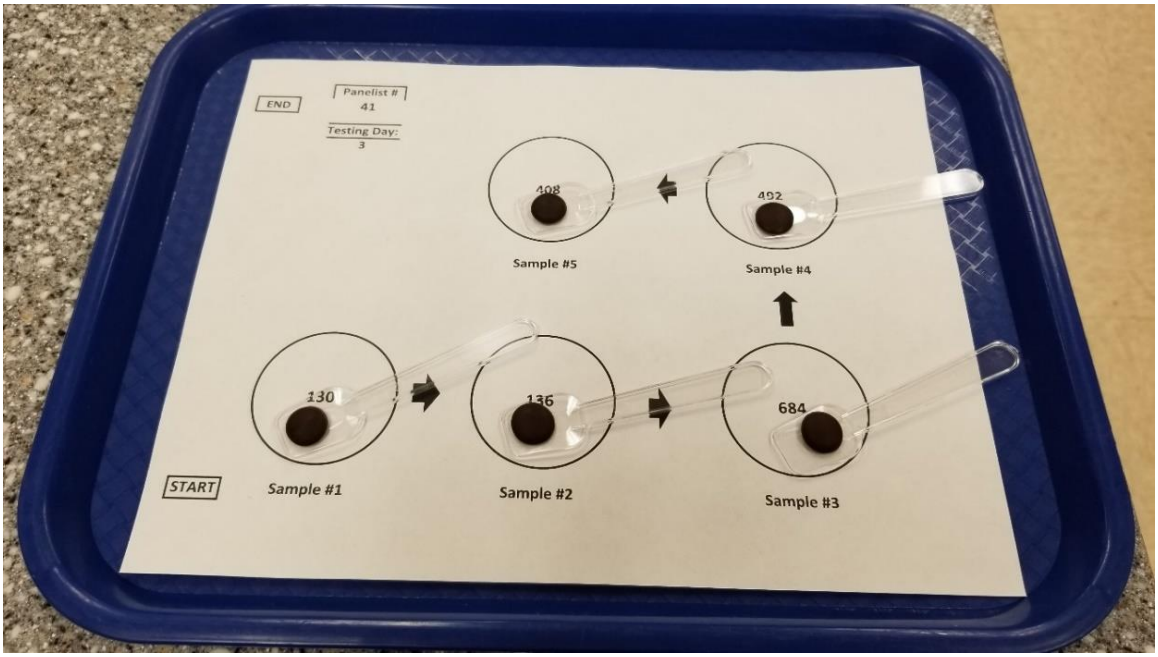

Figure 6. 7: Cafeteria tray with sample-serving placemat, plastic spoons, and samples.

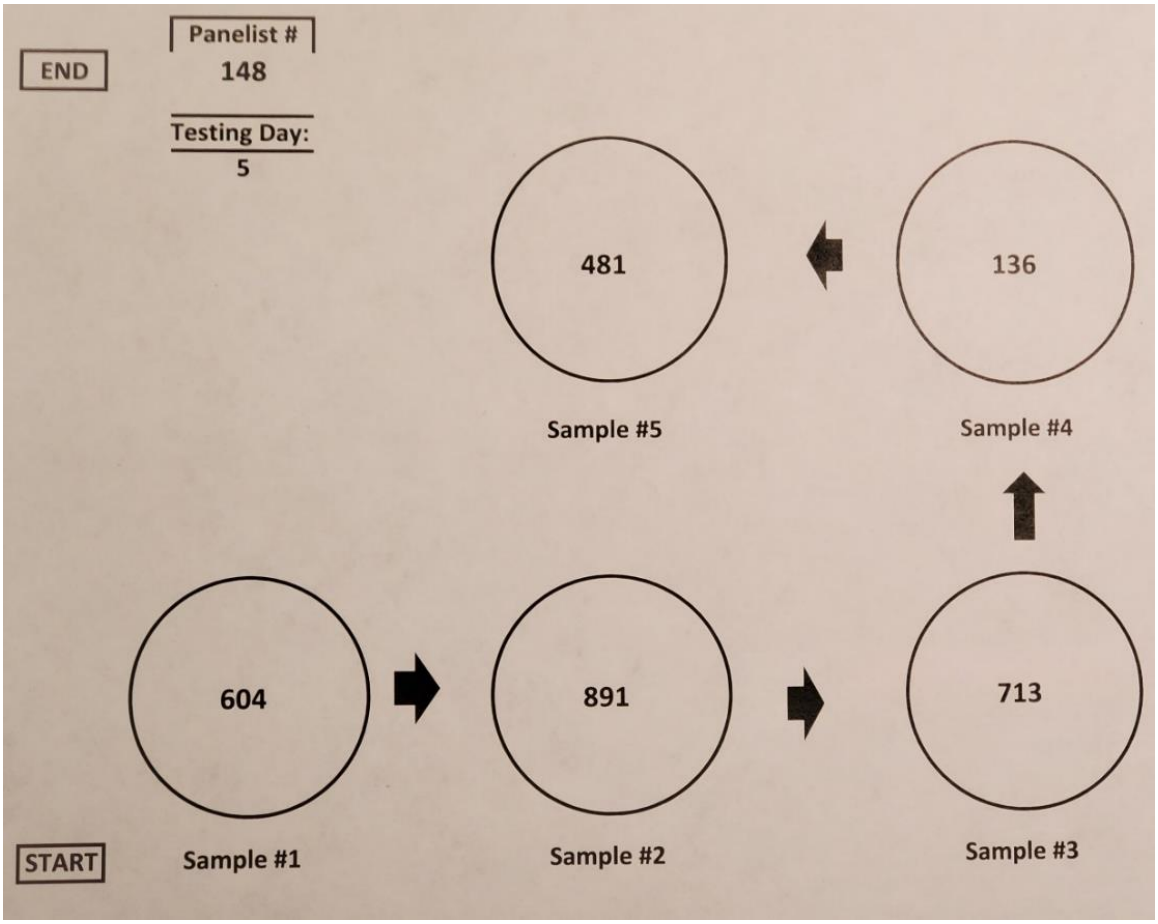

Figure 6. 8: Sample-serving placemat showing order with start and end locations, sample coding for each of the five samples, panelist #, and testing day

Once they received their samples, consumer participants were told to click “next” on the screen, and then to taste Sample #1, which had a three-digit code matching the one seen on the placemat (see example in Figure 6.8). Participants were asked to place the entire 0.3 g disk of chocolate liquor in their mouth, and to immediately rate how much they liked it, using a hedonic 9-point category scale (Figure 6.9).

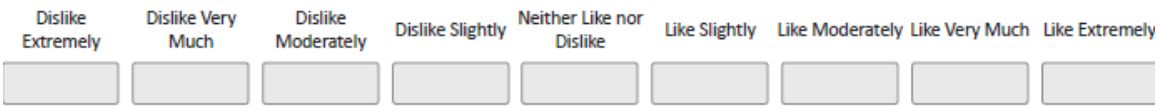

Figure 6. 9: The 9-point hedonic category scale used in this study.

Consumer participants were then immediately asked to rate, while the sample was still in their mouth, the intensities of the other characteristics (i.e., sourness, bitterness, astringency, sweetness, cocoa flavor, and other flavors and textures) (Figure 6.10). Once all attributes had been rated, participants had a mandatory 2-minute break where they were asked to rinse their mouth with room temperature water, swish, swirl and expectorate into a provided disposable polystyrene cup, before they proceeded to the remaining samples in the same fashion, one at a time, again with a 2-minute break and palate cleansing between each sample assessment. On the first tasting day of each session, after the last assessment, participants were asked about their chocolate preference (Figure 6.11), chocolate consumption frequency (Figure 6.12), gender, age, and ethnicity (Figure 6.13). For both sessions 1 and 2, participants returned for 4 or 2 more days respectively to evaluate the remaining samples in the same fashion.

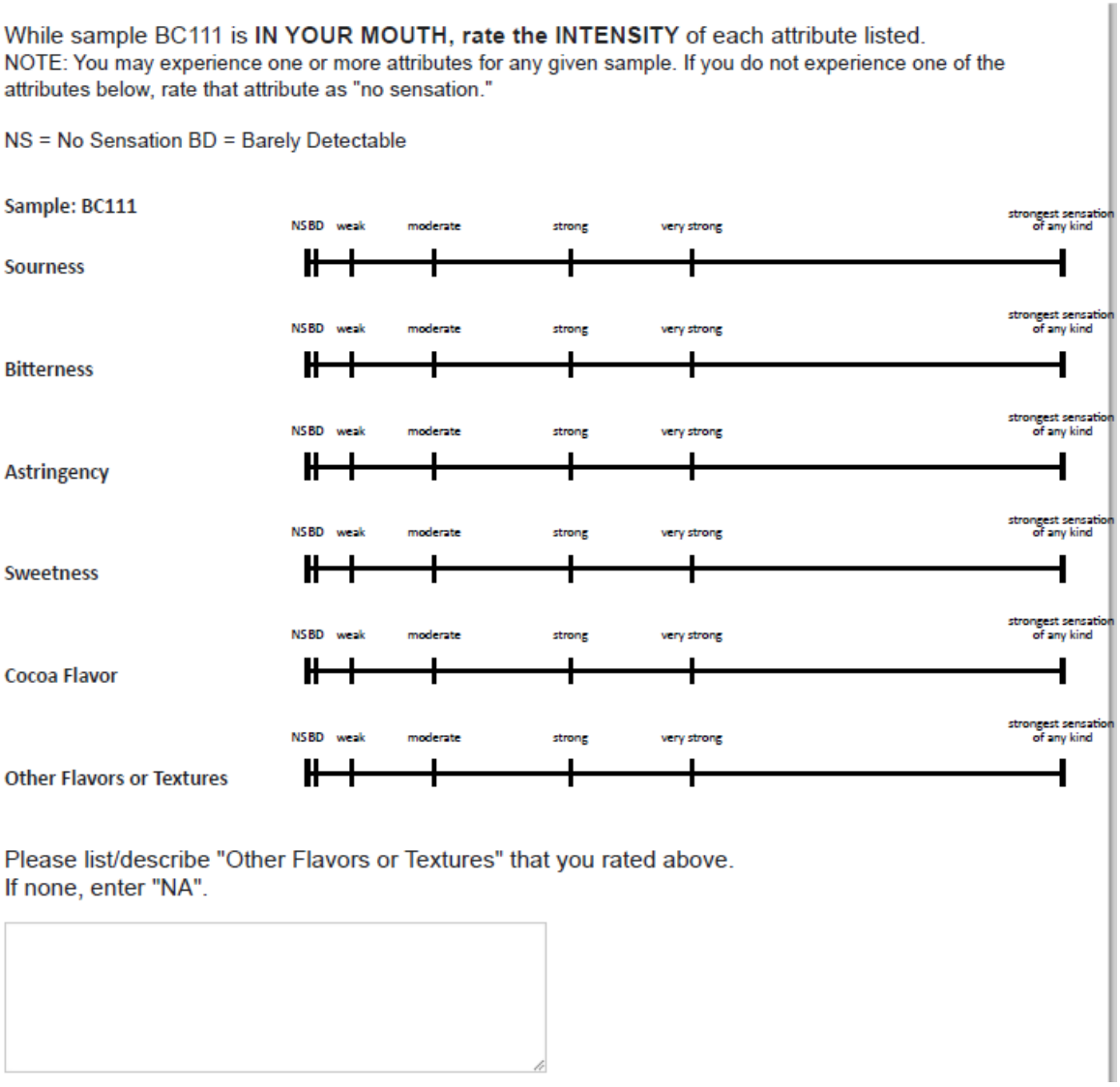

Figure 6. 10: Ballot for all chocolate liquor samples used in both sensory sessions.

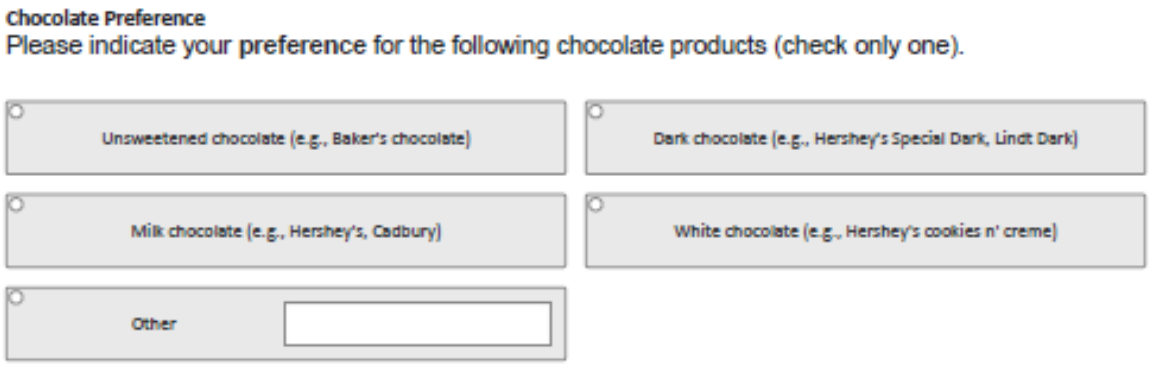

Figure 6. 11: Chocolate preference question asked on day 1 in both sessions.

Figure 6. 12: Chocolate consumption frequency questions

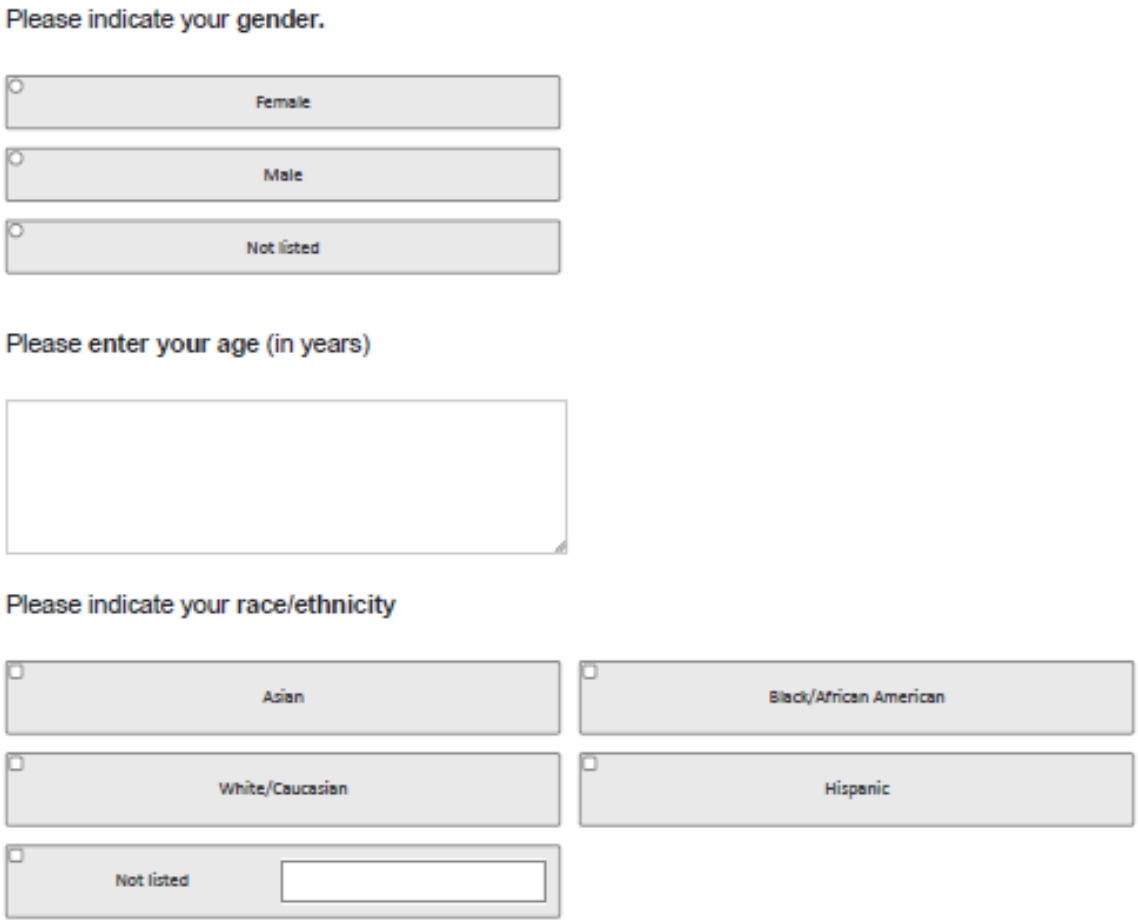

Figure 6. 13: Additional demographic information collection

6.2.2.5.5.2 Section 2

Section 2 of the Consumer Trial was carried out exactly as in Section 1, with three differences. Firstly, in Section 2, all training and sample evaluations were carried out by participants while wearing noseclips (Figures 6.14 and 6.15). The following instructions were added regarding noseclip usage prior to any calibration or evaluation. Any bolded words were also bold in the electronic ballot:

> "Before you proceed, please **put on your nose clip now**.
>
> Put the rubber feet over your nostrils and ensure that the clip fits tightly. You should be completely unable to breathe through your nose. Please do not remove the nose clip until we instruct you to do at the end of the test!"

Proper noseclip usage can be seen in Figure 6.15. Secondly, sample evaluations occurred on only 3 days, and consumers in Section 2 therefore had a slightly different incentive structure for this reason (i.e., $5 incentives on days 1 and 2, and a $10 incentive on day 3 for completing the entire Section 2 trial). The third and final difference between Section 1 and Section 2 was that the experimental design, while based upon a Williams design, just like in Section 1, was necessarily smaller, given the smaller number of samples being assessed (i.e., 9). Summary statistics of the Section 2 experimental design are found in Table 6.15 above. Otherwise, recruitment conditions, participant training, ballot questions, serving of samples, number of samples per day, and assessment conditions such as red lighting and timing all remained the same.

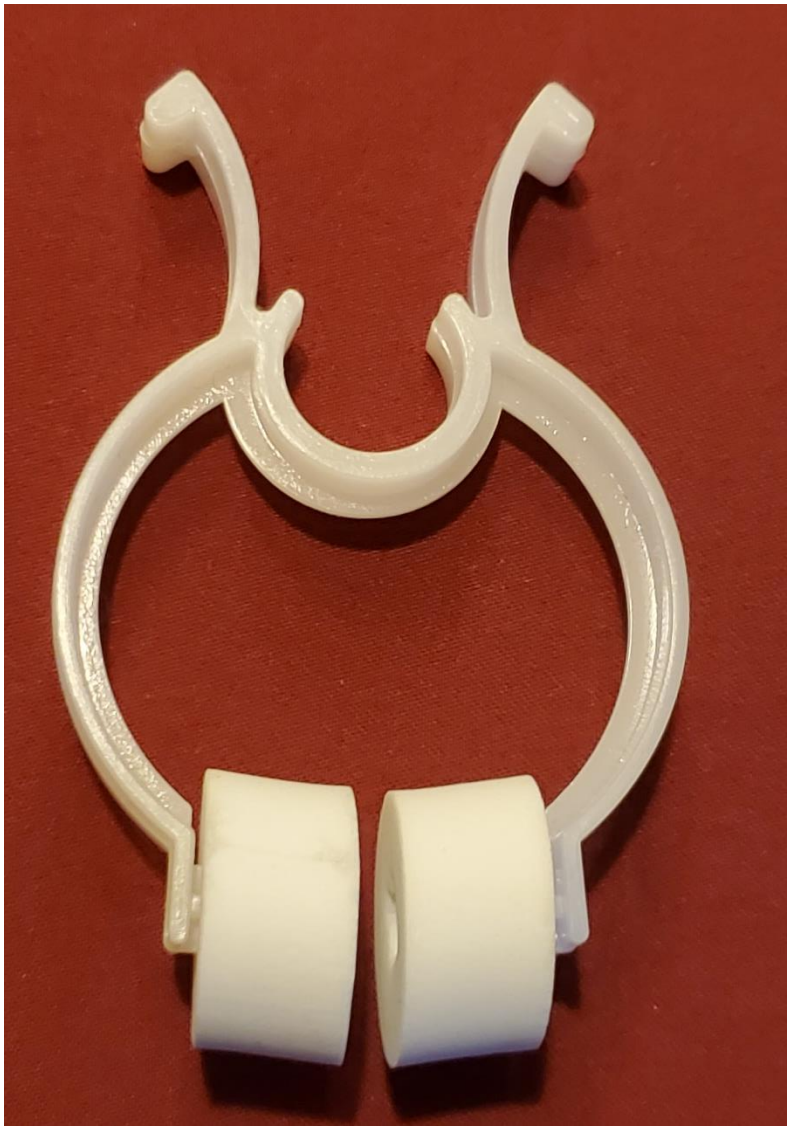

Figure 6. 14: The noseclips used in Section 2 of the consumer evaluation

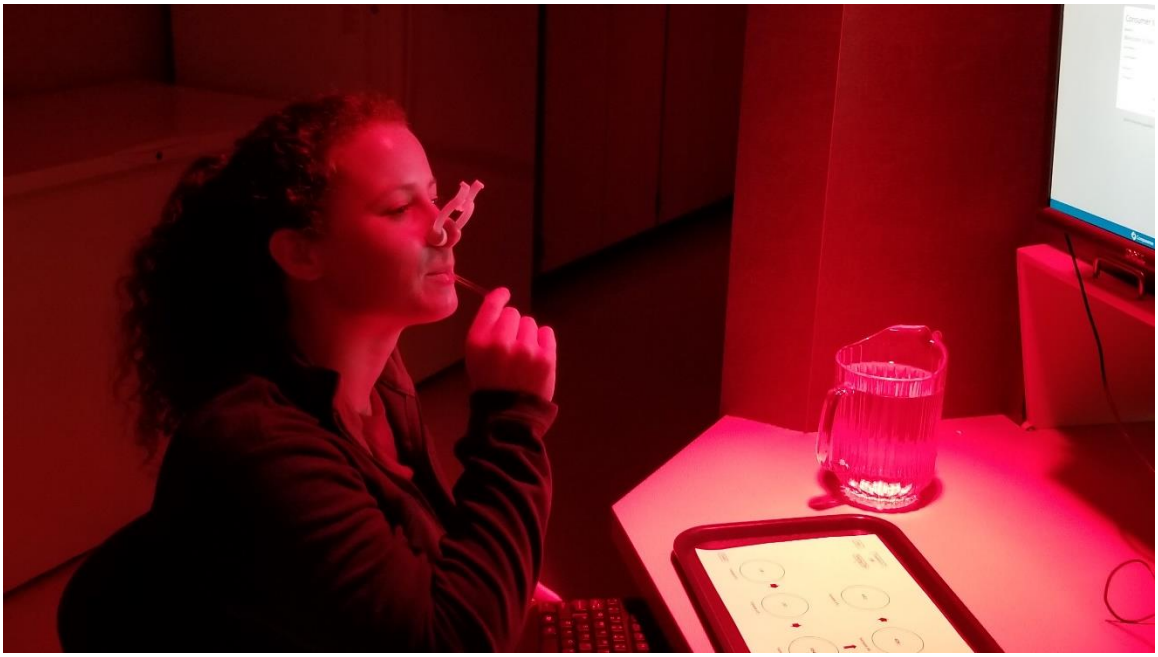

Figure 6. 15: A booth in the SEC testing room with an employee modeling a portion of the assessment with noseclips. Note the red lighting, which was used for Sections 1 and 2.

6.2.2.6 Statistical Analysis of Data

6.2.2.6.1 Methodology

Principal components analysis (PCA), Partial Least Squares Regression (PLSR) and External Bitterness Mapping were carried out on consumer and chemical data. Additional details of each analysis are included in the results and discussion sections below.

6.2.2.6.1.1 Software:

Analysis of the sensory data took place in RStudio v. 1.2.1334, build 1379(f1ac3452) running R version 3.6.0 Patched (2019-06-04 r76666), ggplot2(v.3.3.2: Wickham, 2016)

for PCA biplots, and FactoMineR (v. 2.3: Le, Josse, and Husson, 2008) and SensoMineR (v.1.26: Husson, Le, and Cadoret, 2020) for bitterness maps and related plots.

6.3 Results & Discussion

Once quantitative chemical and consumer sensory data on all of the roasted treatments was obtained (Chapters 4 and 5), exploration of the simplest relationships between the potential chemical predictors of Bitterness began (i.e. Theobromine, Caffeine, Epicatechin, Procyanidin B2, Catechin, and Cyclo(Pro-Val)). As seen in the table of Pearson correlations (Table 6.18), significant positive correlation exists in our data between Epicatechin and Procyanidin B2 (i.e., R=0.97) which was expected given previous work (Cooper et al., 2007). However, a seemingly new finding was that Epicatechin, and therefore Procyanidin B2, were both highly negatively correlated with cyclo(Pro-Val) (i.e., R= -0.756 and R= -0.735 respectively). The magnitude of this correlation was slightly surprising, given the lack of a direct chemical relationship between 2,5-diketopiperazines (e.g., cyclo(Pro-Val), and flavan-3-ols (e.g., Epicatechin, Procyanidin B2, and Catechin). The same could be said of the strong *positive* correlation between Catechin and cyclo(Pro-Val) (Table 6.18) (i.e., R=0.693). These correlations in potential predictor values presented a challenge for linear regression methods. Therefore, we initially pursued methods that are known to be more effective at analyzing data with correlated variables, specifically, Principal Components Analysis (PCA) and Partial Least Squares (PLS) (Lê et al., 2014).

Table 6. 18: Pearson correlations between bitter chemicals (ratings without noseclips)

	Theo	Caff	Epi	Cat	ProB2	cPV
Theo	1.0000	-0.07	0.077	0.136	0.161	0.004
Caff		1.000	0.506	0.344	0.408	0.035
Epi			1.000	-0.459	0.968	-0.756
Cat				1.000	-0.550	0.693
ProB2					1.000	-0.735
cPV						1.000

6.3.1 Principal Components Analysis (PCA)

Figure 6.16 is the first of two biplots that we considered as part of a Principal Components analysis of the data. This biplot shows the first two principal components (i.e., PC1 and PC2), which explain 69.3% of the variation in the "without noseclip data" across all three origins. We did not include the noseclip data in the analysis so as not to introduce significant bias, given that Bitterness ratings using noseclips exist only for the Ghana origin in our dataset (see Chapter 5 for noseclip effect results, and section 6.3.2 below for a separate analysis of chemicals including the "with noseclip" data).

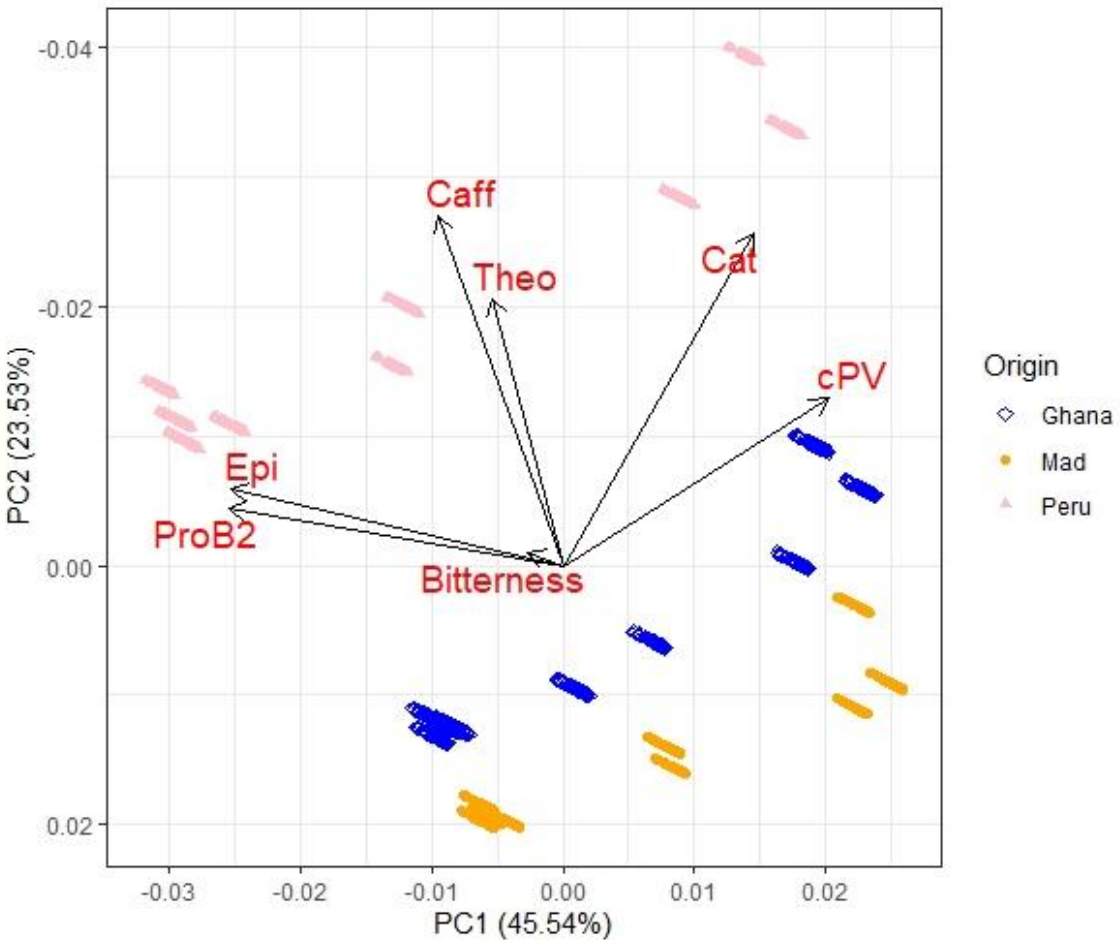

Figure 6. 16: PCA Biplot of six bitter compounds concentrations and Bitterness, all scaled. Principal components 1 and 2. 69.07% variation in the data explained. Epi=Epicatechin, ProB2=ProcyanidinB2, Caff=Caffeine, Theo=Theobromine, Cat=Catechin, cPV=cyclo(Pro-Val)

There are several interesting characteristics in Figure 6.16. First, as noted in previous work (Cooper et al., 2007) and as shown in Table 6.18, a significant positive correlation exists between Epicatechin and Procyanidin B2. This makes sense, not only because they are both in the flavan-3-ol chemical class, but also because they were both previously shown to decrease with roasting (Chapter 4). Given that Bitterness also decreases with roasting (Chapter 5), the positive correlation between these two flavan-3-ol compounds and Bitterness is also not surprising. Next, it can be seen that Caffeine and Theobromine

are both positively correlated, which is somewhat surprising given that Table 6.18 did not show a large correlation. What could account for this? They are both related compounds in the methylxanthine class, and also have quite stable concentrations during roasting, meaning that the majority of the variation in their values is due to cross-origin variation (Chapter 4). Perhaps that accounts for the correlation we see here in PC1 and PC2, especially given that Peru has both the highest levels of Caffeine, and approximately the highest level of Theobromine—though similar to Ghana—especially given the position of the Peruvian treatments in the PC1 x PC2 space. Next, we can see that Catechin, and to an even greater extent cyclo(Pro-Val) are negatively correlated with Bitterness. This seems slightly perplexing, because one clear assumption is that known important bitter compounds in cacao (i.e., Theobromine, Caffeine, Epicatechin, Catechin, ProcyanidinB2, and cyclo(Pro-Val) (Stark et al., 2006)), should all lead to increased Bitterness as their concentrations increase, and not the inverse. This being the case, one wonders how it is possible for Bitterness to decrease with the increase of Catechin and cyclo(Pro-Val), with the latter compound even being quite likely to have a variety of other unmeasured bitter diketopiperazines (DKPs) increasing right along with it (Bonvehí et al., 2000). The most likely answer is that the effect of Bitterness *increase* by these compounds is weak enough to be overshadowed by one or more other factor effects related to Bitterness *decrease*. One reasonable supposition would be that as compounds such as Epicatechin and Procyanidin B2 polymerize, along with other unmeasured Procyanidin dimers, trimers, etc. (Aprotosoaie et al., 2016a), and since they polymerize to somewhat less bitter, or even tasteless, compounds (Stark et al., 2006), in

so doing, they have a powerful enough effect on Bitterness decrease to hide a weaker effect of Bitterness increase by Catechin, cyclo(Pro-Val) and even related DKPs. This especially makes sense given the much higher mean concentrations of Epicatechin and Procyanidin B2 compared to Catechin (i.e., just over 7:1 and 3:1 respectively across all origins and treatments), and over the experimental region, the ratio of mean *decrease* of Epicatechin to *increase* of Catechin alone is approximately 3.4:1, and there is a reasonable likelihood that Epicatechin and Catechin have similar amounts of Bitterness given their close chemical relationship (Stark et al., 2006). Another item of interest in Figure 6.16 is related to the fact that the Peruvian cacao in our study has the highest mean values of Epicatechin, Procyanidin B2, Caffeine, and Catechin, with Theobromine values that approximately match those of the Ghanaian cacao; indeed the biplot reflects most of these details quite well. Additionally, when considering the mean origin-wide values of Bitterness (Table 6.19), Peru is rated as more bitter than either Ghanaian or Madagascan chocolate, and this is reflected in the biplot as well. But the biplot of PC1 x PC2 does not tell the whole story of the relationship between these six bitter compounds and Bitterness. For a more complete picture, consideration of a second biplot is also useful (Figure 6.17).

Table 6. 19: Mean Bitterness Ratings for All Origins (no noseclips)

Origin	Mean Bitterness
Ghana	29.314
Madagascar	30.831
Peru	32.663

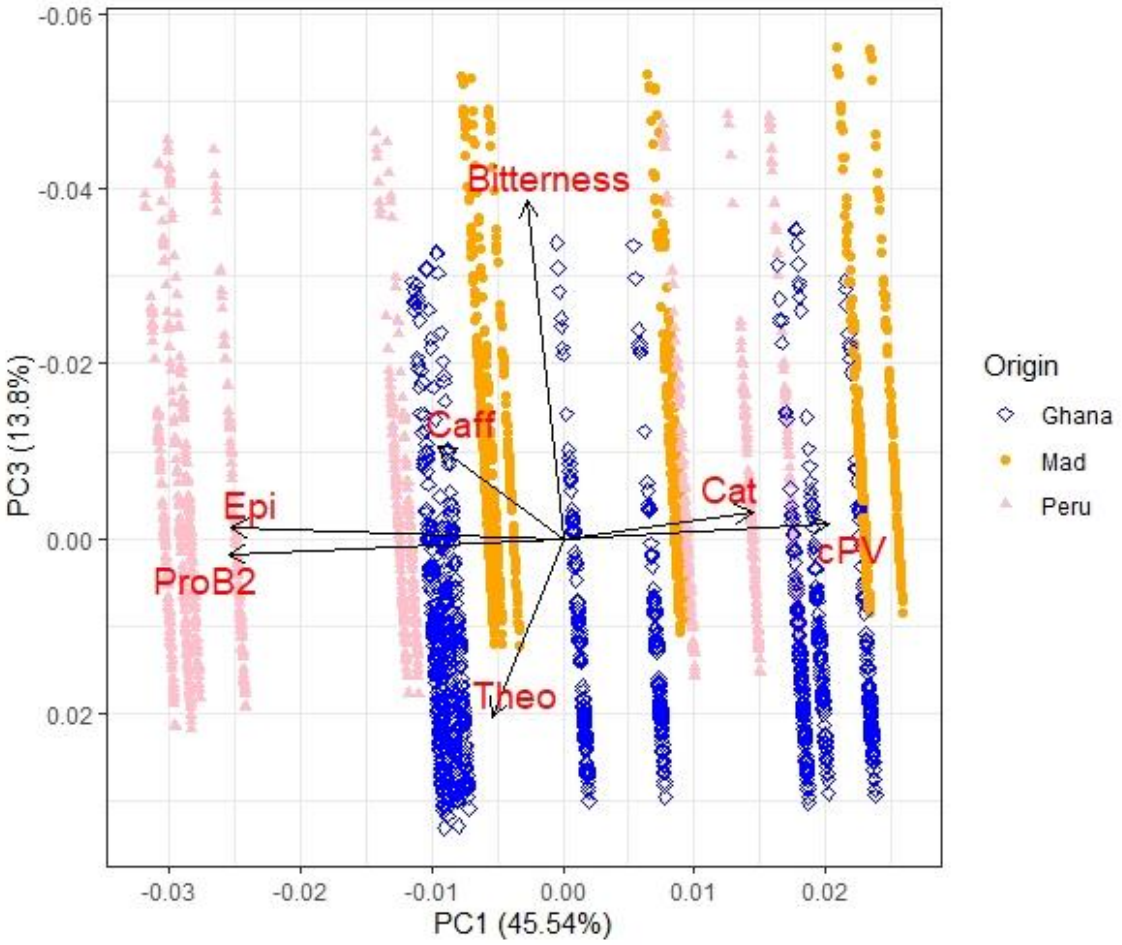

Figure 6. 17: PCA Biplot of six bitter compound concentrations and Bitterness, all scaled. Principal components 1 and 3. 59.34% variation in the data explained.

Figure 6.17 shows another biplot from the same analysis, based upon PC1 x PC3. Between these two principal components, 59.34% of the variation in the dataset is explained, including 13.80% of variation not accounted for in figure 6.16. Between these two figures, 82.87% of the variation in the data is explained, giving a relatively full picture of the relationships of all variables involved. In that context, Figure 6.17 is interesting because the strong positive correlations between Epicatechin and Procyanidin B2 remain, but Catechin and cyclo(Pro-Val) are even more correlated, and

even more negatively correlated to Epicatechin and ProcyanidinB2 than in figure 6.16, which could support the supposition that a decrease in Epicatechin, Procyanidin B2, and probably other procyanidins, might reduce Bitterness to an extent that could overwhelm effects of the Catechin and cyclo(Pro-Val) increases on Bitterness. We also see in Figure 6.17 that Caffeine, of all compounds, is most positively correlated with Bitterness, followed by Epicatechin to a much lesser degree, and then Theobromine is quite negatively correlated with Bitterness. At first, this seems somewhat surprising, given that Theobromine is present at a higher concentration than any other compound in all three origins (Chapter 4). However, this provides useful information, given that Theobromine switched from a positive to negative correlation with Bitterness between Figures 6.16 and 6.17. There is something in the chemical data that could account for this. Table 6.19 shows that for the order from most to least mean Bitterness by origin, we see Peru, Madagascar, and then Ghana. However, Peru and Ghana have close to the same mean value of Theobromine, while Madagascar has much less than both (i.e., 71% and 68% of the Peruvian and Ghanaian values respectively). So, if only the relationship between Madagascar and Peru were to be considered, Theobromine appears as if it has a powerful effect on increasing Bitterness, but if only Madagascar and Ghana were to be considered, the story is quite the opposite. Because the only large change in Theobromine in our study is across origins, and because the origins with the lowest and highest mean values of Bitterness, Ghana and Peru respectively, have approximately the same amount of Theobromine, within a few percent, it is not difficult to imagine that Theobromine, while likely contributing to the mean value of Bitterness in some way,

may simply not be responsible for a large amount of the *difference* in mean Bitterness between the origins, allowing its effect to perhaps be overshadowed by changes in concentrations of the other compounds, such as Epicatechin and/or the procyanidins. While data from additional origins will help to clarify the accuracy of this supposition, one might still wonder how this could be true, given that Theobromine is present at much higher concentrations than all other bitter compounds in our study (Chapter 4), and in fact, present at a higher concentration than all of them combined, at a ratio across all origins and treatments of about 2.16:1 Theobromine : (Caffeine+Epicatechin+ProcyanidinB2+Catechin+cyclo(Pro-Val)). The reason could be as simple as the unique chemical characteristics of Theobromine compared to the other compounds in our study. For example, Theobromine is not particularly water soluble compared to the other compounds (Martin et al., 1981; Spiller, 1997; Stark et al., 2006; Wadsworth, 1920), and saliva being a water-based solution, is not likely to dissolve a large amount of the Theobromine present in the chocolate. Without dissolution, it would remain in the masticated chocolate and out of reach of bitter receptors in the mouth. Additionally, there is some question of the kinetics of dissolution of Theobromine, which during our method development for quantitative chemical analysis (Chapter 3) was seen to dissolve incredibly slowly in water compared to Caffeine and the other bitter compounds when mixed into even relatively dilute stock solutions (<0.33 mg/mL), to the extent that even with an overnight rest in the volumetric flask, it was still not dissolved by morning, and required sonication for completion. For these reasons, especially within the fat-based chocolate matrix, which one could imagine

might protect non-dissolved molecules from ready contact with saliva, it is reasonable to imagine that only a small portion of the large concentration of the Theobromine present in the chocolate might be tasted. If this is the case, then the inconsistency in apparent correlation of Theobromine to Bitterness between the biplots (Figures 6.16 and 6.17) would be avoidable only with further data from additional origins covering much greater variation in Theobromine concentrations.

In summary, PCA analysis, in the context of previously seen Pearson correlations between the potential predictors of Bitterness, leads us to several conclusions. Of the quantitated compounds in our dataset, the correlated (R~0.97) Epicatechin and Procyanidin B2 and then Caffeine appear to be most correlated with Bitterness, and because of the apparent negative correlations of both Catechin and cyclo(Pro-Val) to Bitterness, which we know is misleading given that bitter compounds should be positively correlated with Bitterness, the correlation between Epicatechin, Procyanidin B2 and Bitterness might even be stronger in reality than it appears. Considering this relationship, and what is already known of the importance of the Bitterness of these compounds in cacao (Stark et al., 2006), it is not unreasonable to imagine that this group of flavan-3-ols that decreases with roasting (i.e., Epicatechin, Procyanidin B2, along with other non-quantitated procyanidins (Cooper et al., 2007)) could be the most important compounds influencing Bitterness in the transition from raw to roasted cacao. The true correlation of Caffeine with Bitterness, which should reflect to some degree the causal relationship of the two, however, is much less clear, in part because

of the relative lack of variation of Caffeine during roasting. So while in our data Caffeine is clearly correlated positively with Bitterness, whether this correlation is coincidental, due to the origins chosen and their mean Bitterness levels irrespective of Caffeine, or whether the mean Bitterness levels are, at least in part, because of an effect of Caffeine's presence, this dataset alone cannot tell us. Still, we do know already that the presence of Caffeine in cacao plays a role in Bitterness in general (Stark et al., 2006), and so certainly with data from additional origins, covering more variation in Caffeine concentrations, a stronger statement about the correlation of Caffeine and Bitterness should be possible.

6.3.1.1 External Bitterness Perception Map (PLSR-Based)

But we need not wait for additional data to say at least a bit more about the relationship of these six bitter compounds and Bitterness ratings in these 27 chocolate treatments. For a slightly different approach, we turned to a novel use of PLSR and Consumer rating mapping borrowed from the previously used (Chapter 5) external preference mapping (Bowen, Blake, Tureček, & Amyotte, 2019; Lê et al., 2014). This method of analysis relates all chemical and mean Bitterness rating loadings (PLSR components 1 and 2) to data of full consumer ratings of Bitterness on a contour plot that includes all 27 treatments. Together, these elements produce a complete External Bitterness Perception map (Figures 6.18, 6.19, and 6.20) with regions color coded by level of Bitterness (i.e., dark red for most bitter and dark blue for least bitter). The PLSR loadings superimposed on the map show the relationship of bitter compounds to

consumer perception of Bitterness. Cross-validation (CV) was carried out (Table 6.20) prior to fitting the PLSR model, and it shows that the first two components have a cross-validated adjusted root mean squared error of prediction (CV-RMSEP) of 0.04975, which is only very slightly higher than the lowest CV-RMSEP of 0.04592 for 4 components. Additionally, the first two components explain approximately 55.35% of the variation in mean Bitterness ratings and 67.38% of the variation in the chemical predictor data in the CV training set. It is not surprising that these six chemical compounds do not allow for a 2-component description of Bitterness that is greater than 55.35%, as there are many bitter compounds not specifically accounted for in our model (i.e., at least 37 more (Stark et al., 2006)), but also the categorical effect of the sensation of chocolate aroma, which results in mean Bitterness increase (Chapter 5), and other taste characteristics interacting with Bitterness as well (e.g., Sweetness)(Lawless et al., 2010b) combined with general correlation with other sensory characteristics (i.e., Sourness and Astringency (Chapter 5)). With these considerations in mind, 55.35% variation of Bitterness explained seems quite substantial.

Table 6. 20: Cross-Validation results for PLSR

```
VALIDATION: RMSEP
Cross-validated using 10 random segments.
     (Intercept)  1 comps 2 comps  3 comps 4 comps 5 comps 6 comps
CV         0.0664  0.0533  0.0501  0.0477  0.0462  0.0472  0.0508
adjCV      0.0664  0.0530  0.0497  0.0474  0.0459  0.0468  0.0501

TRAINING: % variance explained
            1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
X             50.98    67.38    85.15    98.29    99.76   100.00
Bitterness    43.88    55.35    59.56    62.57    64.35    64.38
```

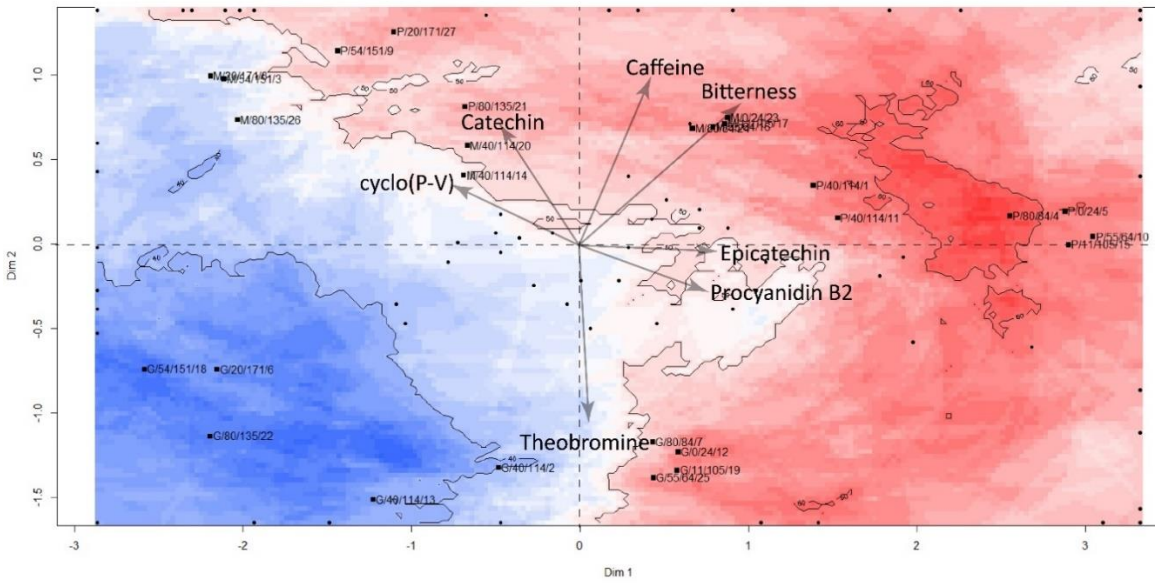

Figure 6. 18: External Bitterness Perception Map: Includes roasted chocolate treatments in PLSR loading space (Dimensions 1 and 2), with color-coded contours consisting of Consumer Bitterness ratings (highest bitterness in dark red and lowest bitterness in dark blue). Loadings of the chemical and Bitterness variables produced during the PLSR analysis are slightly enlarged for visibility, but were kept to scale.

This External Bitterness Rating map results in a largely recognizable scene, as the loadings for Caffeine and the flavan-3-ols Epicatechin and Procyanidin B2 are most positively correlated with Bitterness, while Catechin is slightly less so, approaching orthogonality, and cyclo(P-V) and Theobromine both appear to be slightly negatively correlated with Bitterness, a situation that we have also previously seen and discussed based upon Figures 6.16 and 6.17 PCA biplots for cyclo(Pro-Val) and Theobromine respectively.

In the right half of the External Bitterness Rating Map (Figure 6.19), we can see several Peruvian treatments in or near the region of greatest Bitterness. These treatments and

those in the region of next highest Bitterness are either raw or lightly roasted (i.e., 0 min/24°C, 11 min/105°C, 55 min/64°C, 80 min/84°C) for all origins, along with the Peru 40min/114°C treatment, which is the approximate centerpoint of our experimental roasting region, and so perhaps could be deemed to be moderately roasted for the region. These findings align with our Chapter 5 estimates of roasting effect on Bitterness. Additionally, Caffeine and Bitterness are generally more associated with Madagascar and Peruvian treatments, which is the case (Chapters 4 and 5), and Epicatechin and Procyanidin B2 are most associated with Peruvian treatments, which is also the case (Chapters 4 and 5). Finally, Theobromine is most associated with Ghana, and less so with Peruvian treatments. While this is true, the difference between them is quite small (i.e., Peruvian treatments have a mean value of Theobromine that is 96% as much as that of Ghana), but again, Ghanaian treatments have a significantly lower mean Bitterness than Peruvian treatments (Chapter 5), and given the importance of Bitterness in this model as the response, this could yet again show that the potential effect of Theobromine on Bitterness in our dataset is overwhelmed by origin effects (e.g., mean Bitterness differences between the three origins).

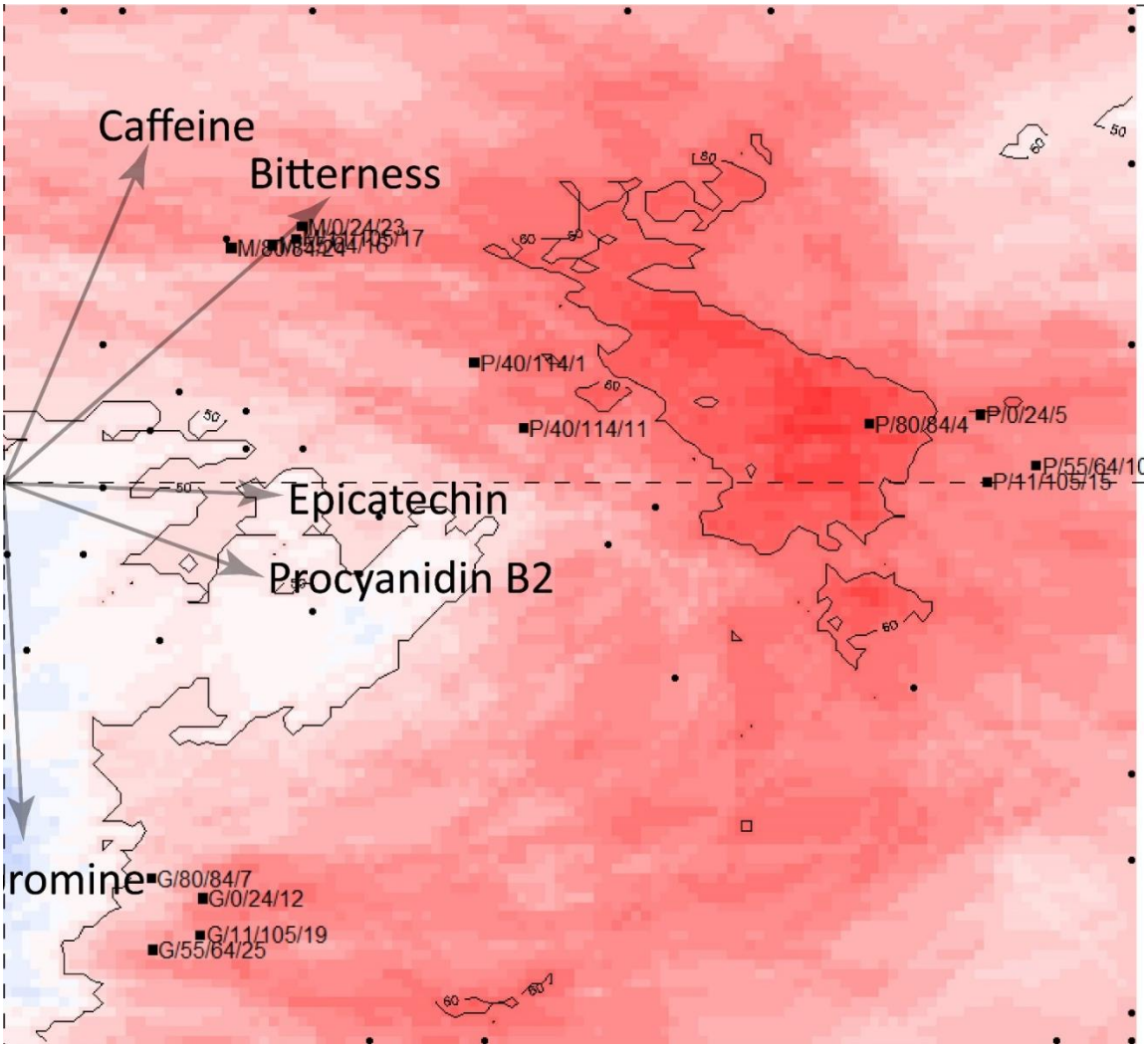

Figure 6. 19: The right half of figure 6.18, showing the region of most Bitterness, the loadings most related to that region, and the treatments within that region. Coding of Treatments is Origin/Time/Temperature/Treatment#. Therefore G/55/64/25 is Ghana roasted for 55 minutes at 64°C, and with treatment code #25. Darker red is more bitter. Blue is less bitter.

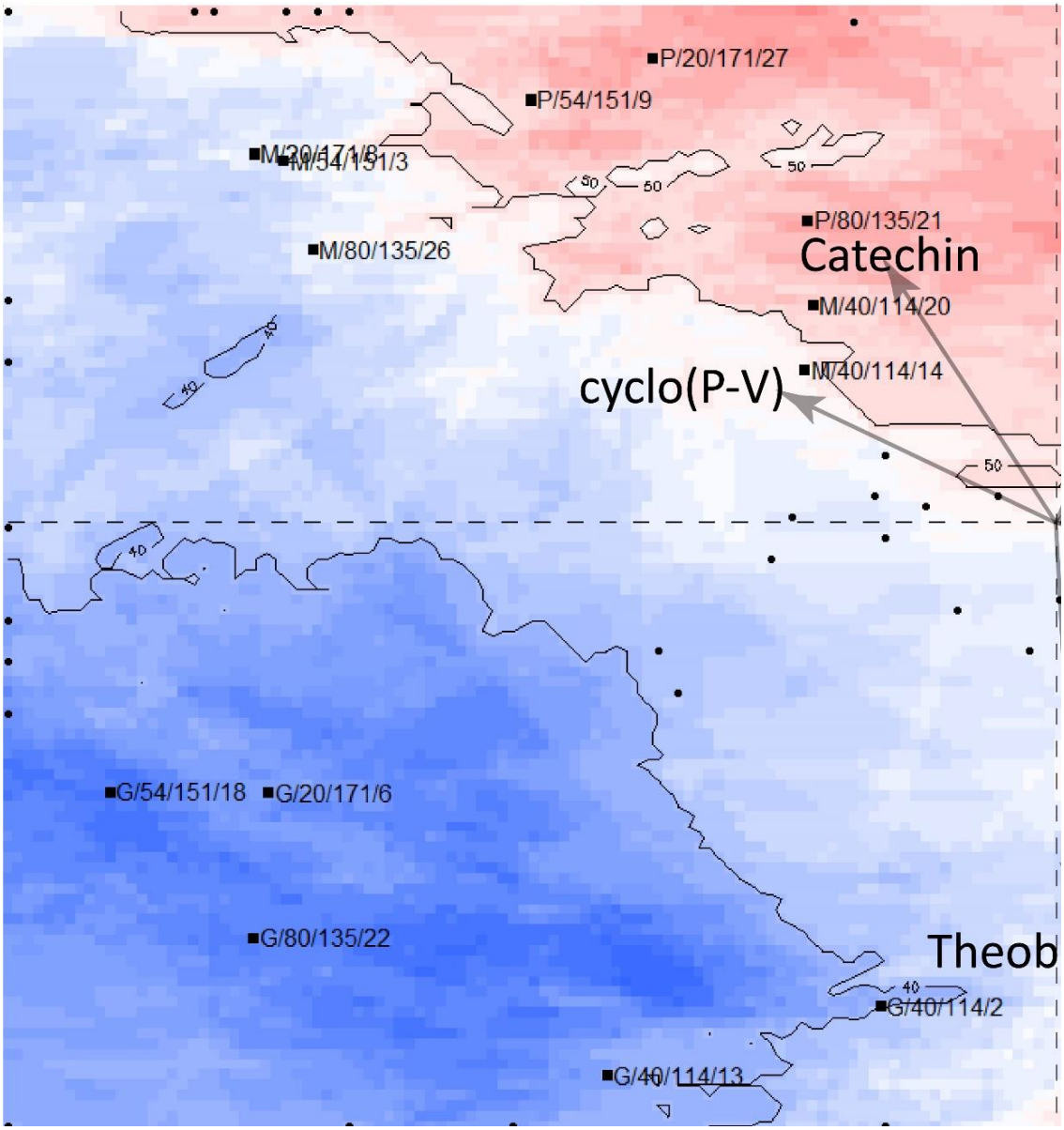

Figure 6. 20: The left half of figure 6.19, showing the region of least Bitterness, the loadings most related to that region, and the treatments within that region Coding of Treatments is Origin/Time/Temperature/Treatment#. Therefore G/20/171/6 is Ghana roasted for 20 minutes at 171°C, and with treatment code #6. Red is more bitter. Dark blue is less bitter.

As for Figure 6.20, showing the left half of Figure 6.18, the most roasted treatments (i.e., 20 min/171°C, 54 min/151°C, 80 min/135°C) are in the region of lowest Bitterness, and

as expected, based upon Chapter 5 results, they are all Ghanaian. The treatments most associated with Catechin are Peruvian and to a lesser extent Madagascan, which is correct, and the treatments most associated with cyclo(Pro-Val) are Madagascan, which is also correct by a small margin (Chapter 4). Overall, this External Bitterness Perception Map supports assumptions held due to consideration of the Pearson correlations and the PCA analysis above, such as the probable importance of Epicatechin and Procyanidin B2 in terms of chocolate Bitterness, the potential importance of Caffeine, and the problematic status of the Theobromine variable, at least in our data set, and the consistently weak correlation of both Catechin and cyclo(Pro-Val) with Bitterness.

6.3.2 Mixed-Model Linear Regression

Given these PCA and PLSR findings, there is an apparent opportunity that exists to solve the problem of correlated bitter compounds in the data in order to perform linear regression after all. Given what we know of the strong positive correlation between Epicatechin and Procyanidin B2 (R=0.97), as well as their negative correlations with Catechin and cyclo(Pro-Val), it makes little sense to maintain them all as separate predictors in a linear model. Rather, if some of them are linearly combined, by adding the concentration values of two pairs of chemical variables for each treatment to produce two new sensible variables, correlation is dramatically reduced (i.e., ≤0.57 in all cases, rather than the previous high values of 0.97 and 0.75), therefore reducing multicollinearity and potential variance inflation issues. The combinations used were Epicatechin + ProcyanidinB2 = EpiProB2, and Catechin + cyclo(Pro-Val)= CatcPV. To

perform mixed-model linear regression, first the predictor variables were mean-centered and scaled given that they are otherwise on quite different scales in the experimental region. Next, a mixed model was fit with Bitterness perception ratings as the response variable using a full potential model (i.e., all relevant first order and second-order effects and their interactions excluding cubic effects and beyond, as well as random effects). Interactions were deemed reasonable as they have previously been suggested in cacao for Bitterness (Pickenhagen et al., 1975). The full potential model is as follows:

Bitterness ~ Theo + Caff + EpiProB2 + CatcPV + Theo:CatcPV + Caff:CatcPV + Caff:EpiProB2 + EpiProB2:CatcPV + Theo:EpiProB2 + Theo:Caff + I(Theo^2) + I(Caff^2) + I(EpiProB2^2) + I(CatcPV^2)+ CatcPV:I(Theo^2) + CatcPV:I(Caff^2) + CatcPV:I(EpiProB2^2) + EpiProB2:I(CatcPV^2) + EpiProB2:I(Caff^2) + EpiProB2:I(Theo^2) + Theo:I(EpiProB2^2) + Theo:I(Caff^2) + Theo:I(CatcPV^2)+ Caff:I(CatcPV^2)+ Caff:I(Theo^2) + Caff:I(EpiProB2^2) +

(1|OAorder) + (1|Day) + (1|OAConsumer:Day) + (1|OAConsumer) + (1|AgeBin) + (1|ChocPref) + (1|WDorder)

Model Predictor Notes: Origin was ultimately not included as a potential factor, as during preliminary exploratory modeling, models containing Origin were generally found to have either extremely large VIF values for some model predictors (exceeding 100 and sometimes 200), worse AIC values, or both, as compared to the best models without Origin. Theo is Theobromine, Caff is

Caffeine, EpiProB2 is a variable consisting of the summed Epicatechin and ProcyanidinB2 values for the specific treatment, and CatcPV is a variable consisting of summed Catechin and cyclo(Pro-Val) values for the specific treatment. Terms in parentheses are random effects, and consist of the following: OAConsumer is the individual consumer, OAorder is the overall sample order across the entire section (all days), Day is the specific day of the section (such as first, second, third, etc.), AgeBin is the full range of ages from 18-65 split into 5 bins and analyzed as categorical factors, ChocPref is one of six chocolate preferences chosen by each consumer as described in Materials & Methods, WDorder is within-day sample tasting order.

Model selection began with backward stepwise selection with R package lmerTest (see Materials and Methods for additional information) which first selects the random effects structure, with single-term deletions based upon p-values calculated with the REML-likelihood ratio test. Next, the fixed-effect structure is chosen with single-term deletions relying upon Satterthwaite's method for calculating degrees of freedom and p-values, while respecting term hierarchy. Residuals of the selected model were then subjected to diagnostic plots. With this data set, residuals of the potential models for predicting Bitterness were found to have a non-skewed distribution with some lack of normality in the tails, but a square-root transformation of the response substantially corrected the issue. After transformation, backward-stepwise-regression-based model selection was again performed, and the residuals were once again inspected via

diagnostic plots to confirm that no significant problems remained. Because there were so many potential fixed-effect predictor terms (>25), with many being combinations of the main effects (i.e., interactions), thereby increasing correlation of predictors that already contained some amount correlation, it was thought wise to carry out backward stepwise regression using not simply the full potential model noted above, but also a variety of subsets for starting models as well. Because this resulted in a number of potential models—all with the same random-effect structure but different fixed effect structures—the one that had both the best (smallest) AIC value was ultimately chosen. Type III ANOVA for unbalanced data was then performed for the selected model to obtain estimates of p-values and coefficients for each fixed-effect term and p-values for each random-effect term in the model. The selected model for best predicting sqrt(Bitterness) was as follows:

Sqrt(Bitterness) ~ Caff + EpiProB2 + CatcPV + Caff:EpiProB2 + I(EpiProB2^2) +

(1|OAorder) + (1|Day) + (1|OAConsumer:Day) + (1|OAConsumer)

Given the presence of random effects, conventional R^2 values could not be computed for the model, so a pseudo R^2 value was instead computed. For this model, the pseudo R^2 is 0.6569, suggesting that the combination of fixed and random effects in the model account for approximately 65.7% of the variance in the square root of the response. Given the nature of the data (i.e., psychophysical data based upon sensory analysis) R^2 values over 0.25 are considered to be large (Cohen, 1988; Hemphill, 2003).

Table 6. 21: Summary of calculated model statistics, including effect estimates

```
Fixed effects:

               Estimate Std. Error        df t value   Pr(>|t|)
(Intercept)      5.1351     0.2568    6.3913   20.00 0.00000053 ***
Caff             0.0954     0.0391 3074.8641    2.44     0.0148 *
EpiProB2         0.1217     0.0488 3019.9415    2.49     0.0127 *
CatcPV          -0.0998     0.0418 3014.7619   -2.39     0.0170 *
I(EpiProB2^2)    0.0877     0.0293 3008.5546    3.00     0.0028 **
Caff:EpiProB2   -0.1681     0.0397 2969.1332   -4.23 0.00002369 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 6. 22: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

              Sum Sq Mean Sq NumDF DenDF F value   Pr(>F)
Caff            8.13    8.13     1  3075    5.95   0.0148 *
EpiProB2        8.48    8.48     1  3020    6.21   0.0127 *
CatcPV          7.79    7.79     1  3015    5.71   0.0170 *
I(EpiProB2^2)  12.26   12.26     1  3009    8.98   0.0028 **
Caff:EpiProB2  24.47   24.47     1  2969   17.92 0.000024 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 6. 23: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                    npar logLik   AIC LRT Df          Pr(>Chisq)
<none>                11  -6047 12116
(1 | OAorder)         10  -6055 12130  16  1        0.0000604003 ***
(1 | Day)             10  -6065 12149  35  1        0.0000000029 ***
(1 | OAConsumer:Day   10  -6147 12314 200  1 <0.0000000000000002 ***
(1 | OAConsumer)      10  -6338 12695 582  1 <0.0000000000000002 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Variance Inflation Factors (VIFs) for first-order fixed effects are slightly higher than 2.5, as might be expected for a dataset with some predictor correlation, even given that additive variable combination to create simpler and less correlated variables was carried out. Still, all VIF values are below the reasonable cutoff of 10 commonly recommended

(Montgomery et al., 2012; Zuur et al., 2010) (i.e., 3.7 for Caffeine, 6.4 for EpiProB2, and 4.2 for CatcPV). Of the fixed-effects terms (Table 6.22), all terms are significant (p-value <0.05), except for Caff:EpiProB2 which is highly significant (i.e., p-value <0.001). As for the random-effects terms (Table 6.23), they are all highly significant. In Table 6.21 it can be seen, perhaps most notably, that Theobromine was not selected as an important predictor of Bitterness. Chemical and sensory data from additional chocolate origins would help to elucidate exactly why, and help to understand if, perhaps due to known chemical characteristics of Theobromine, the compound may be less important for understanding Bitterness in chocolate than other, more dilute compounds. Table 6.21 also shows that of the coefficient estimates based upon scaled predictors, their signs are approximately as expected, i.e., Caffeine and EpiProB2 are positively correlated with Bitterness, and CatcPV is negatively correlated with Bitterness, which as previously discussed could reasonably be considered to be due to relatively small respective effects of Catechin and cyclo(Pro-Val) being concealed by the more substantial effects of other bitter compound changes, in other words, some amount of known correlation, with Epicatechin and procyanidin compounds being likely possibilities. As for the coefficient estimates, it would probably not be wise to place too much emphasis on their exact values when trying to estimate the potential causal relationship between these compounds and Bitterness, but it is still interesting to note that EpiProB2 and its quadratic effect are both positively correlated with Bitterness, and so in future study, change of concentration of these compounds, decreases specifically, should be considered as a potential main effect on Bitterness reduction during roasting. Finally,

the interaction between Caffeine and EpiProB2 is interesting, given that it is the most significant effect in the model, and also larger in magnitude than any of the three main effects. It suggests that as Caffeine and EpiProB2 simultaneously increase, their correlation with Bitterness actually decreases. This might be due to the fact that Epicatechin and Procyanidin B2 are both astringent in addition to being bitter, and given the correlation that we have previously seen between Astringency and Bitterness (Chapter 5), it is not surprising that a mixture of compounds, where one is *only* bitter and the others are *both* bitter and astringent, may result in some significant interesting behavior in the Bitterness ratings. However, there is perhaps a more compelling explanation in that origin effects, which in this case would consist of large differences in mean Bitterness ratings between origins with significantly different but relatively constant Caffeine concentrations, and also EpiProB2 concentrations that do not fully overlap across origins, may be contributing to this behavior (Figure 6.21). Therefore, this supposed interaction should be kept in mind during the design of any future experiments, and selection of additional origins, especially with a variety of Caffeine:Epicatechin ratios should help to clarify the issue. So this model based on all treatments rated for Bitterness without noseclips, across all three origins, suggests that, not surprisingly, EpiProB2 (i.e., Epicatechin and ProcyanidinB2) and Caffeine might be, of all bitter compounds that we have studied, the most important bitter chemicals for explaining Bitterness in chocolate.

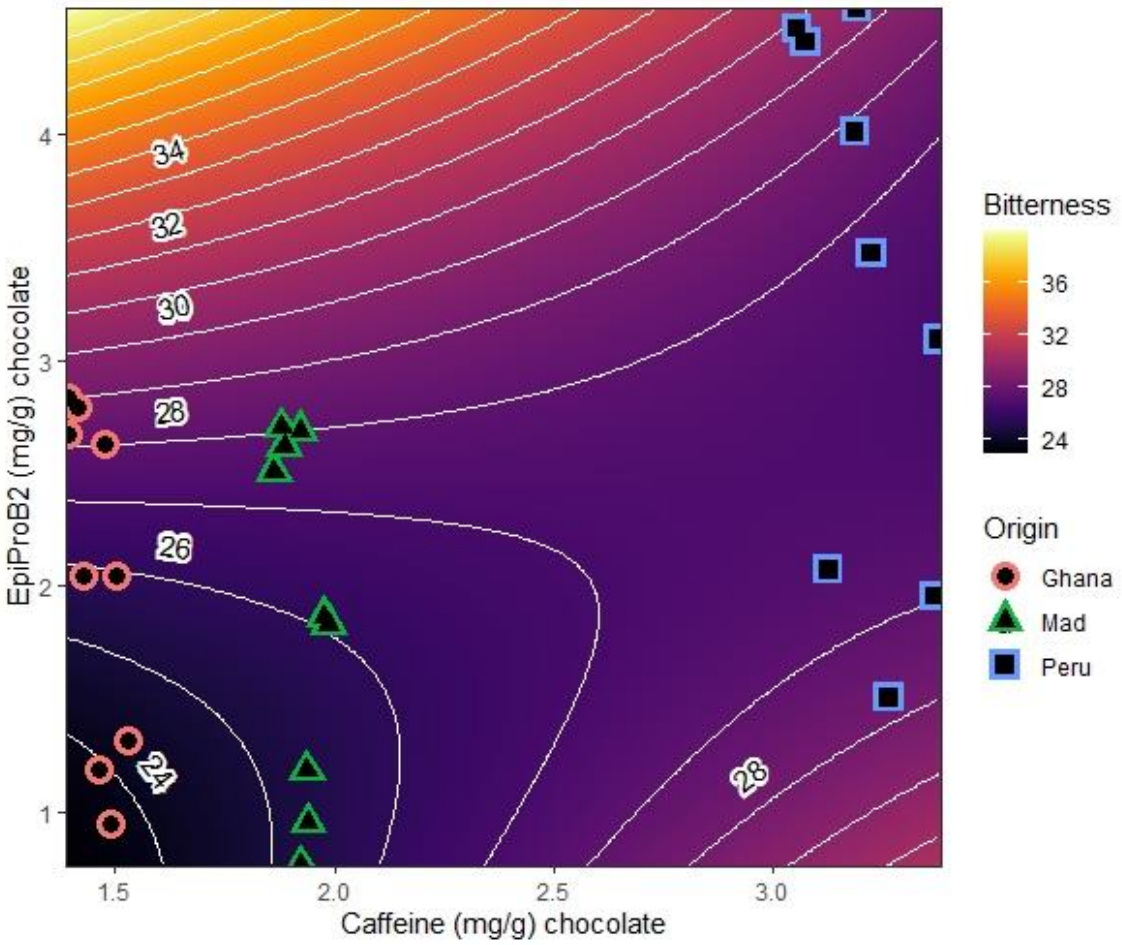

Figure 6. 21: Contour plot of Caffeine by EpiProB2, showing EpiProB2^2 and Caff:EpiProB2 effects with Bitterness as response, and suggesting that origin effects might be responsible for at least some the Caff:EpiProB2 effect in the model discussed above.

However, there is one additional approach that may help to confirm whether the predictors so far discussed are indeed important for understanding Bitterness in chocolate. By modeling the correlation of bitter compound concentrations and Bitterness ratings in the Ghana dataset only, both with and without noseclips, we can at least verify whether similar results are obtained as those already discussed just above. It

is reasonable to consider, after all, that the perception of Bitterness in chocolate when tasted with noseclips might give us a better sense for the contribution of various tasted bitter chemicals without the contribution of any enhancement or suppression effects due to aroma perception. While this may be mostly of theoretical interest, since consumers do not generally eat chocolate without perceiving its aroma, such an understanding of the perception of Bitterness as it relates to the chemical compounds in chocolate without aroma present may still provide information that could allow for the design of more comprehensive experiments in the future. As for model selection, since with a subset of the data that includes only Ghana the degrees of freedom were reduced substantially, it was not possible to perform backward-stepwise selection on the entire very large potential model as was done above. Instead, the full potential model here was chosen to consist of all the terms judged significant in the model just discussed, plus interactions of all main effects, plus Nose, and interactions of Nose and all potential model effects. Therefore, the full potential model being considered is:

Sqrt(Bitterness) ~ Caff + EpiProB2 + CatcPV +

Caff:EpiProB2 + Caff:CatcPV + CatcPV:EpiProB2 + I(EpiProB2^2) +

Nose + Nose:Caff + Nose:EpiProB2 + Nose:CatcPV +

Nose:Caff:EpiProB2 + Nose:Caff:CatcPV + Nose:CatcPV:EpiProB2 +

Nose:I(EpiProB2^2) +

(1|OAorder) + (1|Day) + (1|OAConsumer:Day) + (1|OAConsumer)

Model Predictor Notes: Caff is Caffeine, EpiProB2 is a variable consisting of the summed Epicatechin and ProcyanidinB2 values for the specific treatment, and CatcPV is a variable consisting of summed Catechin and cyclo(Pro-Val) values for the specific treatment. Terms in parentheses are random effects, and consist of the following: OAConsumer is the individual consumer, OAorder is the overall sample order across the entire section (all days), Day is the specific day of the section (such as first, second, third, etc.).

Model selection began by carrying out backward stepwise selection with R package lmerTest (see Materials and Methods for additional information) which first selects the random effect structure, with single-term deletions based upon p-values calculated with the REML-likelihood ratio test. Next, the fixed-effect structure is chosen with single-term deletions relying upon Satterthwaite's method for calculating degrees of freedom and p-values, while respecting term hierarchy. Residuals of the selected model were then subjected to diagnostic plots. With this data set, residuals of the potential models for predicting Bitterness were found to have a non-skewed distribution with some lack of normality in the tails, but a square-root transformation of the response substantially corrected the issue. After transformation, backward-stepwise-regression-based model selection was again performed, and the residuals were once again inspected via diagnostic plots to confirm that no significant problems remained. However, because many of the potential models had very high VIF values for their predictors, it was

thought wise to carry out backward stepwise regression using not simply the full potential model noted above, but also a variety of subsets for starting models as well. Because this resulted in a number of potential models—all with the same random-effect structure but different fixed effect structures—the one that had both the best (smallest) AIC value *and* BIC value was ultimately chosen. Diagnostic plots showed no significant issues, and Variance Inflation Factors (VIFs) for all model predictors were below 2.5. Therefore, the final selected model was:

Sqrt(Bitterness) ~ Caff + EpiProB2 + Nose + Nose:EpiProB2 +

(1|OAorder) + (1|Day) + (1|OAConsumer:Day) + (1|OAConsumer)

Tables 6.24, 6.25, and 6.25 show the relevant model statistics.

Table 6. 24: Summary of calculated model statistics, including effect estimates

```
Fixed effects:
                Estimate Std. Error    df t value            Pr(>|t|)
(Intercept)        5.382      0.408   68.00   13.17 <0.0000000000000002 ***
Caff               0.958      0.453 2310.53    2.11             0.03474 *
EpiProB2           0.254      0.047 2214.56    5.33          0.00000011 ***
Nose1             -0.498      0.049 2342.77  -10.17 <0.0000000000000002 ***
EpiProB2:Nose1    -0.125      0.036 2184.11   -3.46             0.00054 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 6. 25: ANOVA Table (Type III tests)

```
Type III Analysis of Variance Table with Satterthwaite's method

              Sum Sq Mean Sq NumDF DenDF F value              Pr(>F)
Caff             6.6     6.6     1  2311    4.46             0.03474 *
EpiProB2        42.1    42.1     1  2215   28.38          0.00000011 ***
Nose           153.3   153.3     1  2343  103.36 <0.0000000000000002 ***
EpiProB2:Nose   17.8    17.8     1  2184   12.00             0.00054 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Table 6. 26: ANOVA-Like Table for Random Effects (Single-Term Deletions)

```
                     npar logLik  AIC LRT Df           Pr(>Chisq)
<none>                 10  -4686 9392
(1 | OAorder)           9  -4689 9397   7  1                0.009 **
(1 | Day)               9  -4698 9414  24  1          0.00000097 ***
(1 | OAConsumer:Day)    9  -4734 9487  97  1 < 0.0000000000000002 ***
(1 | OAConsumer)        9  -4878 9773 383  1 < 0.0000000000000002 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Given the presence of random effects, conventional R^2 values could not be computed for the model, so a pseudo R^2 value was instead computed. For this model, the pseudo R^2 is 0.6347, suggesting that the combination of fixed and random effects in the model account for approximately 63.5% of the variance in the square root of the response. Given the nature of the data (i.e., psychophysical data based upon sensory analysis) R^2 values over 0.25 are considered to be large (Cohen, 1988; Hemphill, 2003).

All fixed effects (Table 6.25) were highly significant (p-value≤0.001), with the exception of Caffeine, which was significant (p-value≤0.05). All random effects terms (Table 6.26) were highly significant, with the exception of OAorder which was significant. As for effects estimates (Table 6.24), it is interesting to note that even though Caffeine does not change significantly across the region of interest (i.e., from 1.37 mg/g to 1.53 mg/g

chocolate), this model suggests that an increase in Caffeine of one standard deviation (i.e., 0.053 mg/g chocolate) has the greatest effect on an increase in Bitterness, with a mean effect size on the square root of Bitterness that is approximately 3.8 times greater (0.9588/0.2549) than the effect of a one standard deviation increase in EpiProB2 (i.e., 0.68 mg/g chocolate). Still, because EpiProB2 changes much more (i.e., from 0.97 to 2.79 mg/g chocolate) over the region of interest, its overall effect is still quite important (Figures 6.22 and 6.23).

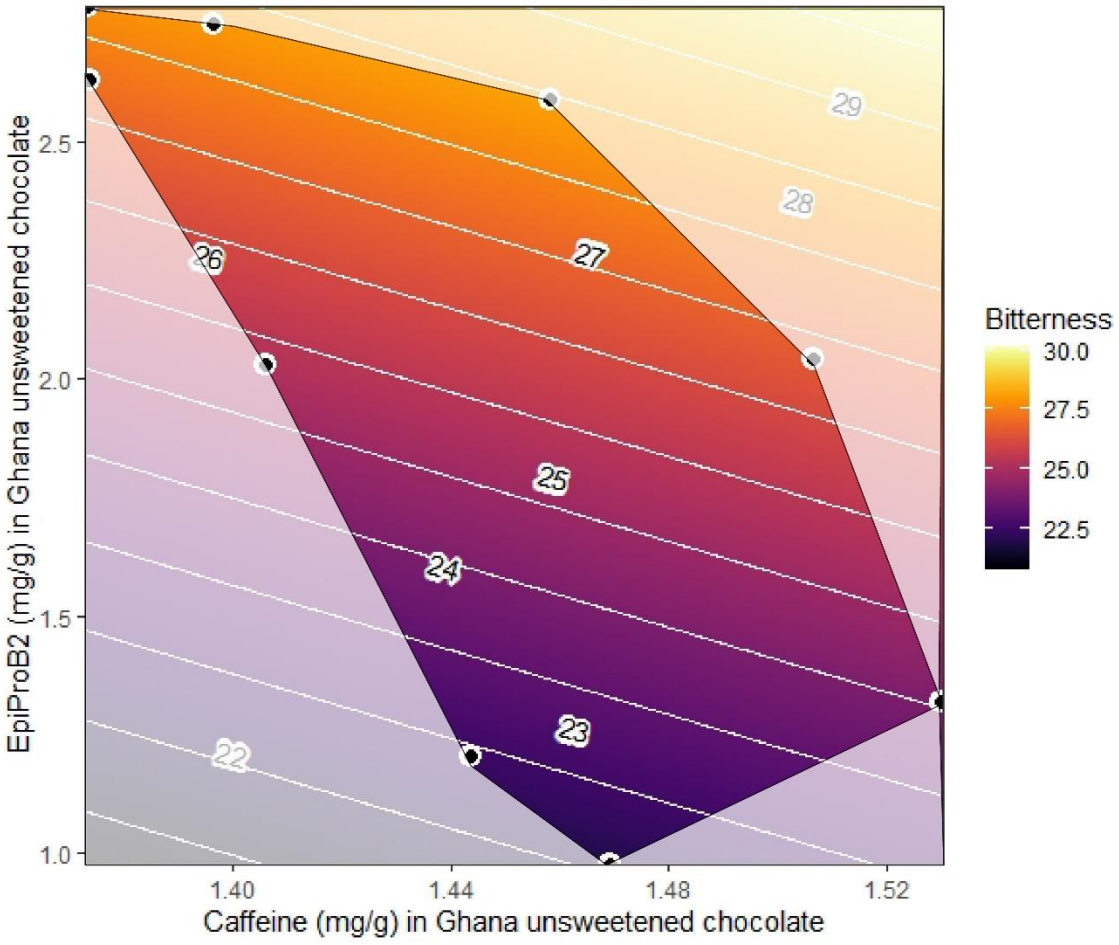

Figure 6. 22: Estimated Bitterness in the region of interest for Ghana WITHOUT noseclips, based upon the selected model above with Caffeine and EpiProB2 predictors.

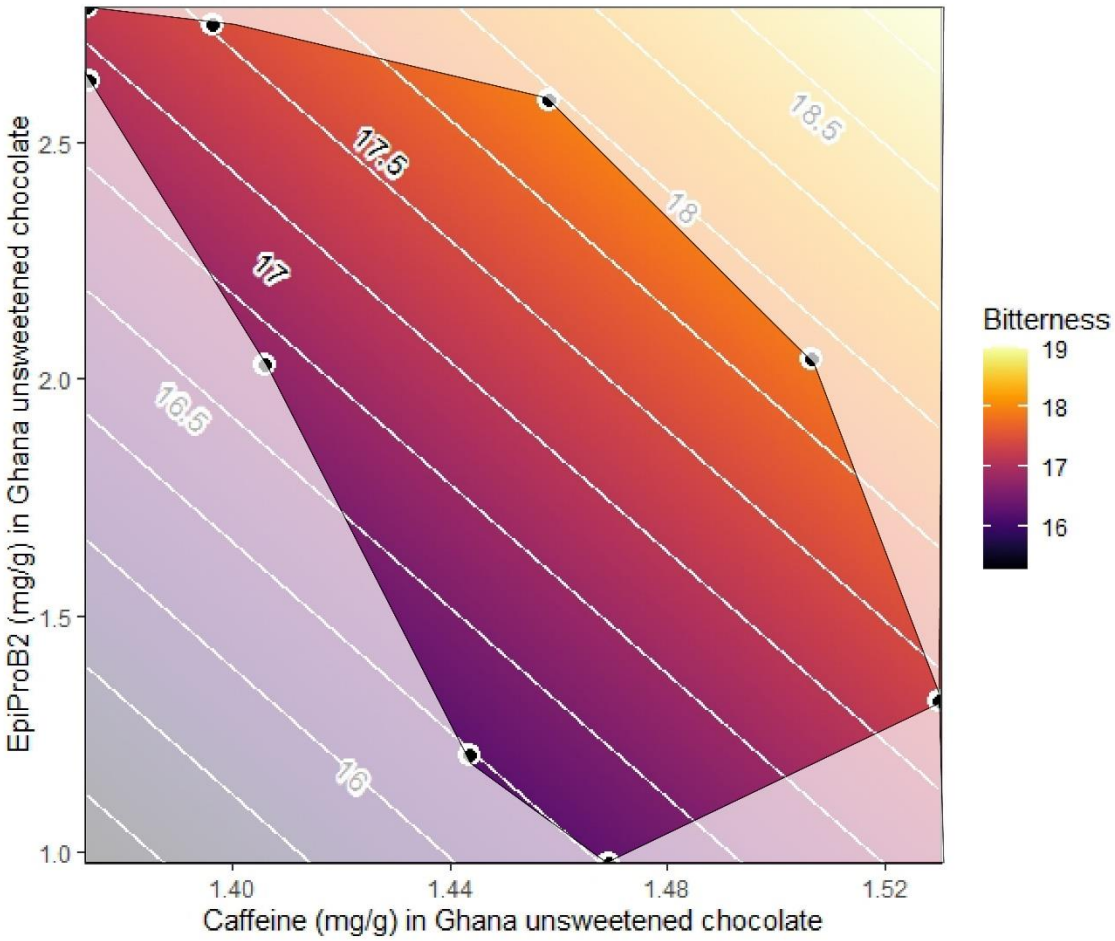

Figure 6. 23: Estimated Bitterness in the region of interest for Ghana WITH noseclips, based upon the selected model above with Caffeine and EpiProB2 predictors.

Figures 6.22 and 6.23 show the general trend of Bitterness increase with both EpiProB2 and Caffeine increase. We also see the expected, given chapter 5 results, decrease in mean Bitterness when noseclips are worn (Figure 6.23). Finally, we see (Figure 6.23) that the interaction Nose:EpiProB2 (Table 6.24) describes an approximately 50% smaller correlation between EpiProB2 and Bitterness when noseclips are worn. It may seem bizarre that the perception of aroma could somehow enhance the Bitterness of EpiProB2 but not Caffeine, but if we recall that Epicatechin and ProcyanidinB2, the components of EpiProB2, both decrease substantially with roasting, at the same time

that many new aroma compounds are being produced, it is easier to imagine that since EpiProB2 changes substantially over the roasting region of interest together with aroma, that any Bitterness enhancement effect of aroma perception on EpiProB2 would be more easily detected and quantified. If we look at the same model, but with an added Caff:Nose effect, we see that as expected, the interaction effect is negative (i.e., -0.5715), and quite large in magnitude, but its p-value is 0.21981, and therefore not significant, and so it was not selected to remain in the model. It is not unreasonable to consider that with data for many more origins and roast profiles, that the significance and magnitude of that potential Nose:Caff interaction could change. Even without additional data, however, we can already say that this model supports the chapter 5 conclusion that the perception of aroma has an enhancement effect on Bitterness and suggests potential interactions that may be chemical-specific. Furthermore, the suggestion made previously several times in this chapter that EpiProB2 and Caffeine are significantly correlated with Bitterness in our chocolate treatments is again supported. Importantly, because this analysis was performed on data from Ghana treatments only, the significance of Caffeine, or EpiProB2 for that matter, cannot have been biased by origin effects, as may have been possible in the previous mixed-model analysis in this chapter.

6.3.3 Summary

Ultimately, to determine a probable quantitative causal relationship between bitter compounds in chocolate and the perception of bitterness by chocolate consumers, it would require a dataset with significant variation for all, or at least most, of the potential predictor chemicals, and to avoid all confounding would require that each predictor be present at all levels for each of the corresponding levels of all other predictors. This quickly becomes unwieldy with chocolate, as there are at least 43 known bitter taste compounds (Stark et al., 2006), as well as the substantial effect of perceived aroma, (Chapter 5), and potentially many specific aroma compounds with effects as well, of which there could be even tens of thousands (Milev et al., 2014). Designing an experiment to consider even a small fraction of these compounds would not realistically be possible. However, through the power of PCA and PLS, as well as mixed-model linear regression, combined with knowledge gained from surface response analyses (chapters 5 and 6), we have been able to come to some useful conclusions as a foundation for future research, with the significant benefit of carrying out the analyses on real chocolate samples likely to contain most of the compounds relevant to Bitterness, even if they are not all yet known. And so though it is not possible to unequivocally show the causal relationship between these 6 bitter compounds and Bitterness, the relationships that do occur in the data between consumer Bitterness ratings of 27 chocolate treatments made from samples of three different origins of cacao, and 6 of the most important bitter compounds in cacao (Stark et al., 2006) have

still been quite illustrative of some of the potential, and even probable, underlying mechanisms of Bitterness change in cacao and chocolate.

6.4 Conclusions

We have little evidence suggesting that Theobromine concentration is strongly correlated to Bitterness in chocolate, either across origins or due to changes during roast. On the other hand, there is much stronger evidence that Caffeine concentration due to roast and origin is correlated with the amount of Bitterness in chocolate, but the relative stability of caffeine during roasting combined with this project's goal of specifically studying roasting effects, means that a more precise determination of Caffeine's ultimate significance and coefficient size as a predictor of Bitterness will require further study still. When considering Epicatechin and Procyanidin B2, little can be said about them separately based upon our data, due to their R=0.97 correlation, but as a group, most probably along with other correlated procyanidins (Cooper et al., 2007), they are, of all the bitter compounds analyzed in this study, the most clearly consistently correlated with changes in Bitterness across all treatments, including all three origins. Notably, Epicatechin has previously been shown to be a more important contributor to Bitterness than higher molecular weight procyanidins (Stark et al., 2006), and if true, its overall importance in chocolate, of all the compounds we have studied, could be greatest. In contrast, Catechin and cyclo(Pro-Val), either separately or as a group, do not appear likely to be particularly important for changes in Bitterness, given their estimated negative correlations to Bitterness. This raises additional questions

about the importance of diketopiperazines (DKPs) as a class as they pertain to Bitterness in chocolate. Given that DKP concentrations increase with roasting (Bonvehí et al., 2000) as Bitterness decreases, there is a chance that individually, or as a class, they are not as responsible for chocolate Bitterness as had been previously suggested (Pickenhagen et al., 1975; Stark et al., 2006). Specifically, we have found no evidence supporting the claim (Stark et al., 2006) that cyclo(Pro-Val) is the most important bitter compound in cacao or chocolate. Again, study of other origins and roasting treatments in the future will be required to further elucidate this complex issue.

CHAPTER 7

7 OVERALL CONCLUSIONS

This research project, which began development as early as 2015, finally came to full fruition in 2020, and in doing so, resulted in findings covering a significant range of chocolate-related topics, from chemical analysis, to sensory analysis, and more. A new method for simultaneous extraction and analysis of perhaps the most important bitter compounds in cacao and chocolate was developed. A custom response-surface methodology (RSM)-based design to study roasting was created with particular emphasis on I-optimality for the irregularly shaped experimental region. Multiple lots of three cacao origins were sourced, roasted, and made into chocolate. Finally, chemical and sensory analysis of the treatments were carried out, followed by significant data analysis and interpretation. While full discussion of such results was included as part of each chapter, as well as detailed conclusions, in this chapter, we summarize a selection of perhaps the most important conclusions of the entire project.

7.1 Quantitative Chemical Analysis Method Development

The aqueous 70% N,N-Dimethylformamide solvent system and HPLC method that were developed for simultaneous extraction of important bitter compounds from three different chemical classes (i.e., methylxanthines, flavan-3-ols, and 2,5-diketopiperazines) functioned successfully, resulting in acceptable standard curves for all compounds, good separation for analytes of interest, and excellent precision with only slightly high RSD% (i.e., 5%< RSD <10%) for cyclo(Pro-Val)). Additionally, recovery of all compounds, despite

their very different solubilities in water, was quite good, without the need for the multiple extraction solvents/methods used in previous studies. For this reason, this method of extraction and analysis is suggested for use in the study of bitter compounds in cacao in the future, especially when fast and efficient extraction and separation is a necessity.

7.2 Chemical Analysis

While our work generally supports previous findings as regards changes during roasting of compounds in the three classes of bitter compounds studied (i.e., methylxanthines, 2,5-diketopiperazines, and flavan-3-ols), further study must occur to better understand the production of diketopiperazines in cacao during roasting, perhaps with a specific eye on cyclo(Pro-Val). This is especially the case because the previously stated importance of cyclo(Pro-Val) for cacao bitterness (Stark et al., 2006) was in large part dependent upon its concentration in a single analyzed sample with unknown fermentation or roasting protocol from Ghana. Given that our chemical data is based upon 8 treatments per origin of cacao—3 origins total including Ghana--across a roasting experimental region of reasonably large size, including a raw treatment, all with reasonable fermentation characteristics, and yet we still did not once confirm concentrations of cyclo(Pro-Val) similar to that of previous research, but rather much lower concentrations, it raises questions that must be investigated further.

7.3 Sensory Evaluation

We discovered that in chocolate treatments made from the samples of three origins that we sourced (i.e., Ghana, Peru, and Madagascar), reduction of Bitterness, Sourness, and Astringency is correlated with increased Liking. It therefore makes sense to choose raw material naturally low in these three characteristics (e.g., Ghana in our study), and to optimize roast to minimize these characteristics. Also, consumers appear to have a preference for increased Cocoa Intensity. So, choosing raw materials accordingly (e.g., Ghana in our study), and roasting to maximize this characteristic may be advised. Roast profiles that minimize and maximize each of these characteristics respectively can vary, at least slightly, by origin, and there is not a complete understanding of this variation within or between origins, but temperature combinations such as 20 minutes/171°C, 80 minutes/135°C, and 54 minutes/151°C, or times and temperatures that would be analogous to these with other roasting equipment, should be considered. Likewise, if maximizing Liking is a consideration, there is not currently support for using raw and lightly roasted cacao (i.e., 0 minutes at 24°C, 11 minutes at 105°C, or 55 minutes at 64°C or similar, based upon specific roasting equipment). Of course, as with any complex system, which chocolate certainly is, caveats do exist. Our study involved analysis of consumer ratings of unsweetened chocolate with multiple roasting treatments. Additions of sugar, salt, and other ingredients would likely introduce significant main and interaction effects that would be relevant to overall sensory characteristics and consumer Liking, especially given the correlation with Sweetness and Liking in chocolate noted in our models. Additionally, there are other aroma profiles that are relevant for

cacao aside from Cocoa, such as Floral, Fruity or Nutty (Aprotosoaie et al., 2016a), and these may also be impacted by ingredient addition in a variety of ways (i.e., mixture suppression or enhancement (Lawless et al., 2010b) relevant to Liking. Finally, it is now understood that perception of chocolate aroma plays a significant and large role in the perception of taste modalities (i.e., Bitterness, Sourness, Sweetness), and the somatosensation Astringency, as well as Liking in chocolate.

7.4 Analysis of Correlation of Bitterness Perception Ratings and Bitter Chemical Data

Regarding the relationship of the 6 specific bitter compounds and Bitterness perception ratings by consumer panelists across all three origins of chocolate treatments, we have little evidence to suggest that Theobromine concentration is strongly correlated to Bitterness in chocolate. There is much stronger evidence that Caffeine concentration changes play a role in the change of Bitterness in chocolate, though a trustworthy estimate of the magnitude of this potential increase is not currently possible. As for Epicatechin and Procyanidin B2, little can be said about them independently, due to their R~0.97 correlation, but as a group, most probably along with other correlated procyanidins (Cooper et al., 2007), they are, of all the bitter compounds analyzed in this study, the most clearly correlated with changes in Bitterness across all treatments, including all three origins. Notably, given that Epicatechin has previously been shown to be a more important contributor to Bitterness than higher molecular weight procyanidins (Stark et al., 2006), its overall importance in chocolate of all the

compounds in the flavan-3-ol class could be greatest. In contrast, Catechin and cyclo(Pro-Val), either separately or as a group, do not appear to be particularly important for changes in Bitterness. This raises additional questions about the importance of diketopiperazines (DKPs) as a class as they pertain to Bitterness in chocolate. Given that DKP concentrations increase with roasting (Bonvehí et al., 2000) as Bitterness decreases, there is a chance that individually or as a class, they are not as correlated with Bitterness as had been previously suggested (Pickenhagen et al., 1975; Stark et al., 2006). Specifically, we have found no evidence supporting the claim (Stark et al., 2006) that cyclo(Pro-Val) is the most important bitter compound in cacao or chocolate.

7.5 Final Thoughts

Given the complexity of chocolate, with a number of chemical compounds approaching 40,000 or more, making it the most complex food ever analyzed (Milev et al., 2014), answers to some of the more complicated questions that we have investigated may not come easily, and will require many more researchers dedicated to the study of chocolate, more advanced analytical instruments, and greater funding. Still, in the context of a world where companies are sometimes unsure about what people like most, and why, especially as these questions relate to potential health effects of cacao metabolites, some of which are bitter, there are now some helpful conclusions that have been determined. It now seems clear that Bitterness is generally not appreciated in chocolate by the average chocolate consumer, and that Liking is, on average,

negatively correlated with Bitterness. Additionally, it is clear that roasting helps to minimize Bitterness in chocolate, leading to maximal Liking, meaning that consumers do not prefer raw or lightly heat-treated chocolate of any origin we studied. Combined with previous findings that suggest the benefits of chocolate consumption for health (Mostofsky et al., 2017), and more specific findings on the positive health benefits of high-molecular-weight flavan-3-ols and melanoidins in chocolate made from roasted cacao (Bellesia et al., 2014; Dorenkott et al., 2014; Morales et al., 2012; Quiroz-Reyes et al., 2018; Stanley et al., 2018), it appears that taking advantage of optimized roasting to minimize Bitterness, Sourness, and Astringency, and maximize Liking is quite a reasonable approach, especially as it may allow for reduced addition of sugar in chocolate products.

FUTURE RESEARCH

It is suggested that future research on this topic take advantage of the newest analytical instrumentation, such as more highly selective LC-MS, GC-MS, and FTICR-MS instruments for chemical analysis in relation to even more substantial sensory evaluation across a variety of chocolate treatments with additional cacao sources/origins, and roast profiles, preferably that come from trees that have been genotyped, and post-harvest processes that have been well detailed. In this way, many more compounds within the complex chocolate matrix, whether identified or not, with much greater variation from natural and process-based differences may be modeled in relation to a variety of sensory characteristics to give a much fuller picture of chemical-by-sensory dose-response relationships in chocolate. A future step would be to consider integration of additional ingredients into chocolate samples prior to analysis, such as sugar and salt, which will undoubtedly lead to interesting new interactions in the models predicting sensory responses.

REFERENCES

Trading and shipping cocoa beans. [Website]. London, UK: International Cocoa Organization; 2015 [Accessed 2020 07/26/2020] Available from: https://www.icco.org/about-cocoa/trading-a-shipping.html.

Fda title 21: Requirements for specific standardized cacao products. United States Food and Drug Association; 2019 [Accessed 2019 9/10/2019] Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=163&showFR=1&subpartNode=21:2.0.1.1.39.2.

Adamson, G.E., Lazarus, S.A., Mitchell, A.E., Prior, R.L., Cao, G., Jacobs, P.H., Kremers, B.G., Hammerstone, J.F., Rucker, R.B., Ritter, K.A. (1999). Hplc method for the quantification of procyanidins in cocoa and chocolate samples and correlation to total antioxidant capacity. *Journal of Agricultural and Food Chemistry, 47*(10), 4184-4188.

Afoakwa, E.O., Paterson, A., Fowler, M., Ryan, A. (2008). Flavor formation and character in cocoa and chocolate: A critical review. *Critical Reviews in Food Science and Nutrition, 48*(9), 840-857.

Ahn, S., Shin, H.-W., Mahmood, U., Khalmuratova, R., Jeon, S.-Y., Jin, H.R., Choi, J.-S., Kim, H.-S., Kim, D.W. (2016). Chronic anosmia induces depressive behavior and reduced anxiety via dysregulation of glucocorticoid receptor and corticotropin-releasing hormone in a mouse model. *Rhinology, 54*(1), 80-87.

Alalor, C. (2015). Review paper advances in the technology of taste-making of unpleasant tasting oral dosage forms: A review. *Continental Journal of Applied Sciences, 10*(1), 18-27.

Andújar, I., Recio, M., Giner, R., Ríos, J. (2012). Cocoa polyphenols and their potential benefits for human health. *Oxidative Medicine and Cellular Longevity,* 1-23.

Aprotosoaie, A.C., Luca, S.V., Miron, A. (2016a). Flavor chemistry of cocoa and cocoa products—an overview. *Comprehensive Reviews in Food Science and Food Safety, 15*(1), 73-91.

Aprotosoaie, A.C., Miron, A., Trifan, A., Luca, V.S., Costache, I.-I. (2016b). The cardiovascular effects of cocoa polyphenols—an overview. *Diseases, 4*(4), 39.

Arts, I.C., Hollman, P.C. (1998). Optimization of a quantitative method for the determination of catechins in fruits and legumes. *Journal of Agricultural and Food Chemistry, 46*(12), 5156-5162.

Barchart (2019). *Cocoa*. The cmdty yearbook. Chicago: Barchart.com, Inc. pp. 38-42.

Barišić, V., Kopjar, M., Jozinović, A., Flanjak, I., Ačkar, Đ., Miličević, B., Šubarić, D., Jokić, S., Babić, J. (2019). The chemistry behind chocolate production. *Molecules, 24*(17), 3163.

Bartoshuk, L.M., Duffy, V.B., Miller, I.J. (1994). Ptc/prop tasting: Anatomy, psychophysics, and sex effects. *Physiology & Behavior, 56*(6), 1165-1171.

Basu, S., Yoffe, P., Hills, N., Lustig, R.H. (2013). The relationship of sugar to population-level diabetes prevalence: An econometric analysis of repeated cross-sectional data. *PloS one, 8*(2), e57873.

Beckett, S.T., Fowler, M.S., Ziegler, G.R. (2017). *Beckett's industrial chocolate manufacture and use* (5th ed). Chichester, UK: John Wiley & Sons.

Bellesia, A., Tagliazucchi, D. (2014). Cocoa brew inhibits in vitro α-glucosidase activity: The role of polyphenols and high molecular weight compounds. *Food Research International, 63*(2014), 439-445.

Bonvehi, J.S., Coll, F.V. (1997). Evaluation of bitterness and astringency of polyphenolic compounds in cocoa powder. *Food Chemistry, 60*(3), 365-370.

Bonvehí, J.S., Coll, F.V. (2000). Evaluation of purine alkaloids and diketopiperazines contents in processed cocoa powder. *European Food Research and Technology, 210*(3), 189-195.

Borthwick, A.D., Da Costa, N.C. (2017). 2, 5-diketopiperazines in food and beverages: Taste and bioactivity. *Critical Reviews in Food Science & Nutrition, 57*(4), 718-742.

Bowen, A.J., Blake, A., Tureček, J., Amyotte, B. (2019). External preference mapping: A guide for a consumer-driven approach to apple breeding. *Journal of Sensory Studies, 34*(1), e12472.

Bowser, S.M., Moore, W.T., McMillan, R.P., Dorenkott, M.R., Goodrich, K.M., Ye, L., O'Keefe, S.F., Hulver, M.W., Neilson, A.P. (2017). High-molecular-weight cocoa procyanidins possess enhanced insulin-enhancing and insulin mimetic activities in human primary skeletal muscle cells compared to smaller procyanidins. *The Journal of Nutritional Biochemistry, 39*(2017), 48-58.

Bray, G.A., Popkin, B.M. (2014). Dietary sugar and body weight: Have we reached a crisis in the epidemic of obesity and diabetes? *Diabetes Care, 37*(4), 950-956.

Calviño, A.M., García-Medina, M.R., Cometto-Muniz, J.E., Rodríguez, M.B. (1993). Perception of sweetness and bitterness in different vehicles. *Perception & Psychophysics, 54*(6), 751-758.

Carrillo, L.C., Londoño-Londoño, J., Gil, A. (2014). Comparison of polyphenol, methylxanthines and antioxidant activity in theobroma cacao beans from different cocoa-growing areas in colombia. *Food Research International, 60*(2014), 273-280.

Carvalho, F.R., Wang, Q.J., van Ee, R., Persoone, D., Spence, C. (2017). "Smooth operator": Music modulates the perceived creaminess, sweetness, and bitterness of chocolate. *Appetite, 108*(2017), 383-390.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second ed). New York, NY: Academic Press.

Cooper, K.A., Campos-Giménez, E., Jiménez Alvarez, D., Nagy, K., Donovan, J.L., Williamson, G. (2007). Rapid reversed phase ultra-performance liquid chromatography analysis of the major cocoa polyphenols and inter-relationships of their concentrations in chocolate. *Journal of Agricultural and Food Chemistry, 55*(8), 2841-2847.

Cooper, K.A., Donovan, J.L., Waterhouse, A.L., Williamson, G. (2008). Cocoa and health: A decade of research. *British Journal of Nutrition, 99*(01), 1-11.

Dagan-Wiener, A., Di Pizio, A., Nissim, I., Bahia, M.S., Dubovski, N., Margulis, E., Niv, M.Y. (2018). Bitterdb: Taste ligands and receptors database in 2019. *Nucleic Acids Research, 47*(D1), D1179-D1185.

De Camargo, M.C.R., Toledo, M.C.F. (1999). Hplc determination of caffeine in tea, chocolate products and carbonated beverages. *Journal of the Science of Food and Agriculture, 79*(13), 1861-1864.

De Taeye, C.d., Kankolongo Cibaka, M.-L., Jerkovic, V., Collin, S. (2014). Degradation of (−)-epicatechin and procyanidin b2 in aqueous and lipidic model systems. First evidence of “chemical” flavan-3-ol oligomers in processed cocoa. *Journal of Agricultural and Food Chemistry, 62*(36), 9002-9016.

Del Rio, D., Calani, L., Cordero, C., Salvatore, S., Pellegrini, N., Brighenti, F. (2010). Bioavailability and catabolism of green tea flavan-3-ols in humans. *Nutrition, 26*(11-12), 1110-1116.

Dorenkott, M.R., Griffin, L.E., Goodrich, K.M., Thompson-Witrick, K.A., Fundaro, G., Ye, L., Stevens, J.R., Ali, M., O’Keefe, S.F., Hulver, M.W. (2014). Oligomeric cocoa procyanidins possess enhanced bioactivity compared to monomeric and polymeric cocoa procyanidins for preventing the development of obesity, insulin resistance, and impaired glucose tolerance during high-fat feeding. *Journal of Agricultural and Food Chemistry, 62*(10), 2216-2227.

Drewnowski, A. (2001). The science and complexity of bitter taste. *Nutrition Reviews, 59*(6), 163-169.

Drewnowski, A., Gomez-Carneros, C. (2000). Bitter taste, phytonutrients, and the consumer: A review. *The American Journal of Clinical Nutrition, 72*(6), 1424-1435.

Edwin, J., Masters, W. (2005). Genetic improvement and cocoa yields in ghana. *Experimental Agriculture, 41*(04), 491-503.

Ellam, S., Williamson, G. (2013). Cocoa and human health. *Annual Review of Nutrition, 33*(2013), 105-128.

Fang, W., Meinhardt, L.W., Mischke, S., Bellato, C.M., Motilal, L., Zhang, D. (2014). Accurate determination of genetic identity for a single cacao bean, using molecular markers with a nanofluidic system, ensures cocoa authentication. *Journal of Agricultural and Food Chemistry, 62*(2), 481-487.

Farah, D., Zaibunnisa, A. (2012). Optimization of cocoa beans roasting process using response surface methodology based on concentration of pyrazine and acrylamide. *International Food Research Journal, 19*(4), 1355-1359.

Fischer, A., Gilad, Y., Man, O., Pääbo, S. (2005). Evolution of bitter taste receptors in humans and apes. *Molecular Biology and Evolution, 22*(3), 432-436.

Fischer, U., Noble, A.C. (1994). The effect of ethanol, catechin concentration, and ph on sourness and bitterness of wine. *American Journal of Enology and Viticulture, 45*(1), 6-10.

Gaudette, N.J., Pickering, G.J. (2013). Modifying bitterness in functional food systems. *Critical Reviews in Food Science and Nutrition, 53*(5), 464-481.

Gautschi, M., Schmid, J.P., Peppard, T.L., Ryan, T.P., Tuorto, R.M., Yang, X. (1997). Chemical characterization of diketopiperazines in beer. *Journal of Agricultural and Food Chemistry, 45*(8), 3183-3189.

Ginz, M., Engelhardt, U.H. (2000). Identification of proline-based diketopiperazines in roasted coffee. *Journal of Agricultural and Food Chemistry, 48*(8), 3528-3532.

Goos, P., Jones, B. (2011). *Optimal design of experiments: A case study approach*. Chichester, UK: John Wiley & Sons.

Gu, Y., Hurst, W.J., Stuart, D.A., Lambert, J.D. (2011). Inhibition of key digestive enzymes by cocoa extracts and procyanidins. *Journal of Agricultural and Food Chemistry, 59*(10), 5305-5311.

Hemphill, J.F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist, 58*(1), 78-79.

Hockfield, H., Fullom, C., Roper, G., Sheeley, R., Hurst, W., Martin, R. (1982). Thermochemical investigations of the dimerization of theobromine. Proceedings of the American Chemical Society (MARM)

Middle Atlantic Region, USA.

Hu, F.B. (2013). Resolved: There is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases. *Obesity Reviews, 14*(8), 606-619.

Hu, S., Kim, B.-Y., Baik, M.-Y. (2016). Physicochemical properties and antioxidant capacity of raw, roasted and puffed cacao beans. *Food Chemistry, 194*(2016), 1089-1094.

Hurst, W.J., Krake, S.H., Bergmeier, S.C., Payne, M.J., Miller, K.B., Stuart, D.A. (2011). Impact of fermentation, drying, roasting and dutch processing on flavan-3-ol stereochemistry in cacao beans and cocoa ingredients. *Chemistry Central Journal, 5*(1), 53.

Ishaq, S., Jafri, S. (2017). Biomedical importance of cocoa (theobroma cacao): Significance and potential for the maintenance of human health. *Matrix Science Pharma, 1*(1), 01-05.

John, W.A., Böttcher, N.L., Aßkamp, M., Bergounhou, A., Kumari, N., Ho, P.-W., D'Souza, R.N., Nevoigt, E., Ullrich, M.S. (2019). Forcing fermentation: Profiling proteins, peptides and polyphenols in lab-scale cocoa bean fermentation. *Food Chemistry, 278*(2019), 786-794.

Johnson, E.S., Bekele, F.L., Brown, S.J., Song, Q., Zhang, D., Meinhardt, L.W., Schnell, R.J. (2009). Population structure and genetic diversity of the trinitario cacao (l.) from trinidad and tobago. *Crop Science, 49*(2), 564-572.

Jones, B., Goos, P. (2012). I-optimal versus d-optimal split-plot response surface designs. *Journal of Quality Technology, 44*(2), 85-101.

Kahyaoglu, T. (2008). Optimization of the pistachio nut roasting process using response surface methodology and gene expression programming. *LWT-Food Science and Technology, 41*(1), 26-33.

Kahyaoglu, T., Kaya, S. (2006). Determination of optimum processing conditions for hot-air roasting of hulled sesame seeds using response surface methodology. *Journal of the Science of Food and Agriculture, 86*(10), 1452-1459.

Kappes, S., Schmidt, S., Lee, S.Y. (2006). Color halo/horns and halo-attribute dumping effects within descriptive analysis of carbonated beverages. *Journal of Food Science, 71*(8),

Kay, J., Haywood, P. (1946). The determination of theobromine in cocoa residues—an examination of the wadsworth method. *Analyst, 71*(841), 162-166.

Keast, R.S., Bournazel, M.M., Breslin, P.A. (2003a). A psychophysical investigation of binary bitter-compound interactions. *Chemical Senses, 28*(4), 301-313.

Keast, R.S., Breslin, P.A. (2003b). An overview of binary taste–taste interactions. *Food Quality and Preference, 14*(2), 111-124.

Keast, R.S.J. (2008). Modification of the bitterness of caffeine. *Food Quality and Preference, 19*(5), 465-472.

Keen, C.L., Holt, R.R., Oteiza, P.I., Fraga, C.G., Schmitz, H.H. (2005). Cocoa antioxidants and cardiovascular health. *The American Journal of Clinical Nutrition, 81*(1), 298S-303S.

Kim, H., Keeney, P. (1983). Method of analysis for (-)-epicatechin in cocoa beans by high performance liquid chromatography. *Journal of Food Science, 48*(2), 548-551.

Kim, H., Keeney, P. (1984). (-)-epicatechin content in fermented and unfermented cocoa beans. *Journal of Food Science, 49*(4), 1090-1092.

Kongor, J.E., Hinneh, M., Van de Walle, D., Afoakwa, E.O., Boeckx, P., Dewettinck, K. (2016). Factors influencing quality variation in cocoa (theobroma cacao) bean flavour profile—a review. *Food Research International, 82*(2016), 44-52.

Kothe, L., Zimmermann, B.F., Galensa, R. (2013). Temperature influences epimerization and composition of flavanol monomers, dimers and trimers during cocoa bean roasting. *Food Chemistry, 141*(4), 3656-3663.

Kreiser, W.R., Martin Jr, R. (1978). High pressure liquid chromatographic determination of theobromine and caffeine in cocoa and chocolate products. *Journal-Association of Official Analytical Chemists, 61*(6), 1424-1427.

Langer, S., Marshall, L.J., Day, A.J., Morgan, M.R. (2011). Flavanols and methylxanthines in commercially available dark chocolate: A study of the correlation with nonfat cocoa solids. *J Agr Food Chem, 59*(15), 8435-8441.

Langner, E., Rzeski, W. (2014). Biological properties of melanoidins: A review. *International Journal of Food Properties, 17*(2), 344-353.

Lawless, H.T. (2013). *Quantitative sensory analysis: Psychophysics, models and intelligent design* (First ed). West Sussex, UK: John Wiley & Sons.

Lawless, H.T., Heymann, H. (2010a). *Physiological and psychological foundations of sensory function*. Sensory evaluation of food: Principles and practices. 2nd ed. New York, NY: Springer Science & Business Media. pp. 19-56.

Lawless, H.T., Heymann, H. (2010b). *Sensory evaluation of food: Principles and practices* (Second ed). New York, NY: Springer Science & Business Media.

Lê, S., Worch, T. (2014). *Analyzing sensory data with r* (First ed). Boca Raton, FL: Chapman and Hall/CRC Press.

Lee, S.-Y., Yoo, S.-S., Lee, M.-J., Kwon, I.-B., Pyun, Y.-R. (2001). Optimization of nibs roasting in cocoa bean processing with lotte-better taste and color process. *Food Science and Biotechnology, 10*(3), 286-293.

Ley, J.P., Blings, M., Paetz, S., Krammer, G.E., Bertram, H.-J. (2006). New bitter-masking compounds: Hydroxylated benzoic acid amides of aromatic amines as structural analogues of homoeriodictyol. *Journal of Agricultural and Food Chemistry, 54*(22), 8574-8579.

Li, Y., Feng, Y., Zhu, S., Luo, C., Ma, J., Zhong, F. (2012). The effect of alkalization on the bioactive and flavor related components in commercial cocoa powder. *Journal of Food Composition and Analysis, 25*(1), 17-23.

Madihah, K.Y.K., Zaibunnisa, A.H., Norashikin, S., Rozita, O., Misnawi, J. (2012). Optimization of roasting conditions for high-quality robusta coffee. *APCBEE Procedia, 4*(2012), 209-214.

Maehashi, K., Huang, L. (2009). Bitter peptides and bitter taste receptors. *Cellular and Molecular Life Sciences, 66*(10), 1661-1671.

Margolis, G., Chiovini, J., Pagliaro, F.A. (1989). Removal of xanthine stimulants from cocoa. Google Patents.

Martin, A., Paruta, A., Adjei, A. (1981). Extended hildebrand solubility approach: Methylxanthines in mixed solvents. *Journal of Pharmaceutical Sciences, 70*(10), 1115-1120.

Matsui, N., Ito, R., Nishimura, E., Yoshikawa, M., Kato, M., Kamei, M., Shibata, H., Matsumoto, I., Abe, K., Hashizume, S. (2005). Ingested cocoa can prevent high-fat diet-induced obesity by regulating the expression of genes for fatty acid metabolism. *Nutrition, 21*(5), 594-601.

Mendes, L.C., de Menezes, H.C., Aparecida, M., Da Silva, A. (2001). Optimization of the roasting of robusta coffee (c. Canephora conillon) using acceptability tests and rsm. *Food Quality and Preference, 12*(2), 153-162.

Mennella, J.A., Pepino, M.Y., Reed, D.R. (2005). Genetic and environmental determinants of bitter perception and sweet preferences. *Pediatrics, 115*(2), e216-e222.

Milev, B.P., Patras, M.A., Dittmar, T., Vrancken, G., Kuhnert, N. (2014). Fourier transform ion cyclotron resonance mass spectrometrical analysis of raw fermented cocoa beans of cameroon and ivory coast origin. *Food Research International, 64*(958-961.

Miller, K.B., Hurst, W.J., Flannigan, N., Ou, B., Lee, C., Smith, N., Stuart, D.A. (2009). Survey of commercially available chocolate-and cocoa-containing products in the united states. 2. Comparison of flavan-3-ol content with nonfat cocoa solids, total polyphenols, and percent cacao. *Journal of Agricultural and Food Chemistry, 57*(19), 9169-9180.

Miller, K.B., Hurst, W.J., Payne, M.J., Stuart, D.A., Apgar, J., Sweigart, D.S., Ou, B. (2008). Impact of alkalization on the antioxidant and flavanol content of commercial cocoa powders. *Journal of Agricultural and Food Chemistry, 56*(18), 8527-8533.

Min, S., Yang, H., Seo, S., Shin, S., Chung, M., Kim, J., Lee, S., Lee, H., Lee, K. (2013). Cocoa polyphenols suppress adipogenesis in vitro and obesity in vivo by targeting insulin receptor. *International Journal of Obesity, 37*(4), 584.

Mintel (2016). Chocolate confectionery 2016 executive summary. London, UK.

Montgomery, D.C., Peck, E.A., Vining, G.G. (2012). *Introduction to linear regression analysis* (Fifth ed). Hoboken, NJ: John Wiley & Sons.

Morales, F.J., Somoza, V., Fogliano, V. (2012). Physiological relevance of dietary melanoidins. *Amino Acids, 42*(4), 1097-1109.

Mostofsky, E., Johansen, M.B., Tjønneland, A., Chahal, H.S., Mittleman, M.A., Overvad, K. (2017). Chocolate intake and risk of clinically apparent atrial fibrillation: The danish diet, cancer, and health study. *Heart, 103*(15), 1163-1167.

Motamayor, J.C., Lachenaud, P., e Mota, J.W.d.S., Loor, R., Kuhn, D.N., Brown, J.S., Schnell, R.J. (2008). Geographic and genetic population differentiation of the amazonian chocolate tree (theobroma cacao l). *PLoS One, 3*(10), e3311.

Myers, R.H., Montgomery, D.C., Anderson-Cook, C.M. (2016). *Experimental designs for fitting response surfaces--ii*. Response surface methodology: Process and product optimization using designed experiments. 4th ed. Chichester, UK: John Wiley & Sons. pp. 451-542.

Næs, T., Brockhoff, P.B., Tomić, O. (2010). *Statistics for sensory and consumer science* (First ed). West Sussex, UK: John Wiley & Sons.

Natsume, M., Osakabe, N., Yamagishi, M., Takizawa, T., Nakamura, T., Miyatake, H., Hatano, T., Yoshida, T. (2000). Analyses of polyphenols in cacao liquor, cocoa,

and chocolate by normal-phase and reversed-phase hplc. *Bioscience, Biotechnology, and Biochemistry, 64*(12), 2581-2587.

Nazaruddin, R., Seng, L., Hassan, O., Said, M. (2006). Effect of pulp preconditioning on the content of polyphenols in cocoa beans (theobroma cacao) during fermentation. *Industrial Crops and Products, 24*(1), 87-94.

Negri, R., Di Feola, M., Di Domenico, S., Scala, M.G., Artesi, G., Valente, S., Smarrazzo, A., Turco, F., Morini, G., Greco, L. (2012). Taste perception and food choices. *Journal of Pediatric Gastroenterology and Nutrition, 54*(5), 624-629.

Neyraud, E., Dransfield, E. (2004). Relating ionisation of calcium chloride in saliva to bitterness perception. *Physiology & Behavior, 81*(3), 505-510.

Oracz, J., Nebesny, E., Żyżelewicz, D. (2015). Changes in the flavan-3-ols, anthocyanins, and flavanols composition of cocoa beans of different theobroma cacao l. Groups affected by roasting conditions. *European Food Research and Technology, 241*(5), 663-681.

Oracz, J., Nebesny, E., Żyżelewicz, D. (2019). Identification and quantification of free and bound phenolic compounds contained in the high-molecular weight melanoidin fractions derived from two different types of cocoa beans by uhplc-dad-esi-hr-msn. *Food Research International, 115*(2019), 135-149.

Oyejola, B., Nwanya, J. (2015). Selecting the right central composite design. *International Journal of Statistics and Applications, 5*(1), 21-30.

Özdemir, M., Devres, O. (2000). Analysis of color development during roasting of hazelnuts using response surface methodology. *Journal of Food Engineering, 45*(1), 17-24.

Pacyniak, B. (2014). Global state of the industry. Candy Industry. Troy, MI: BNP Media. pp. 26-31.

Pacyniak, B. (2016). Global state of the industry. Candy Industry Troy, MI: BNP Media. pp. 17-22.

Payne, M.J., Hurst, W.J., Miller, K.B., Rank, C., Stuart, D.A. (2010). Impact of fermentation, drying, roasting, and dutch processing on epicatechin and catechin content of cacao beans and cocoa ingredients. *Journal of Agricultural and Food Chemistry, 58*(19), 10518-10527.

Peluso, I., Palmery, M., Serafini, M. (2015). Effect of cocoa products and flavanols on platelet aggregation in humans: A systematic review. *Food & Function, 6*(7), 2128-2134.

Pickenhagen, W., Dietrich, P., Keil, B., Polonsky, J., Nouaille, F., Lederer, E. (1975). Identification of the bitter principle of cocoa. *Helvetica Chimica Acta, 58*(4), 1078-1086.

Porter, L., Ma, Z., Chan, B. (1991). Cacao procyanidins: Major flavanoids and identification of some minor metabolites. *Phytochemistry, 30*(5), 1657-1663.

Prodolliet, J., Bruelhart, M. (1993). Determination of aspartame and its major decomposition products in foods. *Journal of AOAC International, 76*(2), 275-282.

Quiroz-Reyes, C.N., Fogliano, V. (2018). Design cocoa processing towards healthy cocoa products: The role of phenolics and melanoidins. *Journal of Functional Foods, 45*(2018), 480-490.

Ramalakshmi, K., Raghavan, B. (1999). Caffeine in coffee: Its removal. Why and how? *Critical Reviews in Food Science and Nutrition, 39*(5), 441-456.

Risner, C.H. (2008). Simultaneous determination of theobromine,(+)-catechin, caffeine, and (−)-epicatechin in standard reference material baking chocolate 2384, cocoa, cocoa beans, and cocoa butter. *Journal of Chromatographic Science, 46*(10), 892-899.

Rizzi, G. (1989). *Heat-induced flavor formation from peptides*. In: T.H. Parliament RJM, C.-T. Ho, editor. Thermal generation of aromas. Washington, D.C.: American Chemical Society. pp. 172-181.

Roudot-Algaron, F., Le Bars, D., Einhorn, J., Adda, J., Gripon, J. (1993). Flavor constituents of aqueous fraction extracted from comté cheese by liquid carbon dioxide. *Journal of Food Science, 58*(5), 1005-1009.

Roy, G.M. (1997). *Modifying bitterness: Mechanism, ingredients, and applications* (First ed). Boca Raton, FL: CRC Press.

Rusconi, M., Conti, A. (2010). Theobroma cacao l., the food of the gods: A scientific approach beyond myths and claims. *Pharmacological Research, 61*(1), 5-13.

Scharbert, S., Hofmann, T. (2005). Molecular definition of black tea taste by means of quantitative studies, taste reconstitution, and omission experiments. *Journal of Agricultural and Food Chemistry, 53*(13), 5377-5384.

Senanayake, U., Wijesekera, R. (1971). Theobromine and caffeine content of the cocoa bean during its growth. *Journal of the Science of Food and Agriculture, 22*(5), 262-263.

Spiller, G.A. (1997). *Caffeine* (First ed). Boca Ration, FL: CRC Press.

Stanley, T.H., Van Buiten, C.B., Baker, S.A., Elias, R.J., Anantheswaran, R.C., Lambert, J.D. (2018). Impact of roasting on the flavan-3-ol composition, sensory-related chemistry, and in vitro pancreatic lipase inhibitory activity of cocoa beans. *Food Chemistry, 255*(1), 414-420.

Stark, T., Bareuther, S., Hofmann, T. (2005a). Sensory-guided decomposition of roasted cocoa nibs (theobroma cacao) and structure determination of taste-active polyphenols. *Journal of Agricultural and Food Chemistry, 53*(13), 5407-5418.

Stark, T., Bareuther, S., Hofmann, T. (2006). Molecular definition of the taste of roasted cocoa nibs (theobroma cacao) by means of quantitative studies and sensory experiments. *Journal of Agricultural and Food Chemistry, 54*(15), 5530-5539.

Stark, T., Hofmann, T. (2005b). Structures, sensory activity, and dose/response functions of 2, 5-diketopiperazines in roasted cocoa nibs (theobroma cacao). *Journal of Agricultural and Food Chemistry, 53*(18), 7222-7231.

Strat, K.M., Rowley, T.J., Smithson, A.T., Tessem, J.S., Hulver, M.W., Liu, D., Davy, B.M., Davy, K.P., Neilson, A.P. (2016). Mechanisms by which cocoa flavanols improve metabolic syndrome and related disorders. *The Journal of Nutritional Biochemistry, 35*(2016), 1-21.

Sukha, D.A., Umaharan, P., Butler, D.R. (2017). The impact of pollen donor on flavor in cocoa. *Journal of the American Society for Horticultural Science, 142*(1), 13-19.

Sun-Waterhouse, D., Wadhwa, S.S. (2013). Industry-relevant approaches for minimising the bitterness of bioactive compounds in functional foods: A review. *Food and Bioprocess Technology, 6*(3), 607-627.

Sun, M., Tao, L., Gu, Y., Lambert, J. (2015). Dietary cocoa improves hepatic mitochondrial function in obese mice. *The FASEB Journal, 29*(1 Supplement 271.7)

Takrama, J., Kun, J., Meinhardt, L., Mischke, S., Opuku, S., Padi, F.K., Zhang, D. (2014). Verification of genetic identity of introduced cacao germplasm in ghana using single nucleotide polymorphism (snp) markers. *African Journal of Biotechnology, 13*(21),

Tarka Jr, S.M., Hurst, W.J. (1998). *Chemistry, isolation, and biosynthesis of methylxanthines*. In: Spiller G, editor. Caffeine. First ed. Boca Raton, FL: CRC Press. pp. 13-23.

Terada, H., Suzuki, A., Tanaka, H., Yamamoto, K. (1992). Determination of catechins and methylxanthines in foodstuffs by semi-micro high performance liquid chromatography. *Food Hygiene and Safety Science, 33*(4), 347-354.

Timbie, D., Sechrist, L., Keeney, P. (1978). Application of high-pressure liquid chromatography to the study of variables affecting theobromine and caffeine concentrations in cocoa beans. *Journal of Food Science, 43*(2), 560-565.

Tokusoglu, Ö., Ünal, K. (2002). Optimized method for simultaneous determination of catechin, gallic acid, and methylxanthine compounds in chocolate using rp-hplc. *European Food Research and Technology, 215*(4), 340-346.

Trognitz, B., Cros, E., Assemat, S., Davrieux, F., Forestier-Chiron, N., Ayestas, E., Kuant, A., Scheldeman, X., Hermann, M. (2013). Diversity of cacao trees in waslala, nicaragua: Associations between genotype spectra, product quality and yield potential. *PLoS One, 8*(1), e54079.

Turkmen, N., Sari, F., Velioglu, Y.S. (2006). Effects of extraction solvents on concentration and antioxidant activity of black and black mate tea polyphenols determined by ferrous tartrate and folin–ciocalteu methods. *Food Chemistry, 99*(4), 835-841.

Van der Greef, J., Tas, A., Nijssen, L., Jetten, J., Höhn, M. (1987). Identification and quantitation of diketopiperazines by liquid chromatography—mass spectrometry, using a moving belt interface. *Journal of Chromatography A, 394*(1), 77-88.

Villeneuve, F., Cros, E., Vincent, J.-C., Macheix, J.-J. (1989). Recherche d'un indice de fermentation du cacao. Iii: Evolution des flavan-3-ols de la fève. *Café, Cacao, Thé, 33*(3), 165-170.

Vlachojannis, J., Erne, P., Zimmermann, B., Chrubasik-Hausmann, S. (2016). The impact of cocoa flavanols on cardiovascular health. *Phytotherapy Research, 30*(10), 1641-1657.

Vries, J.W., Johnson, K.D., Heroff, J.C. (1981). Hplc determination of caffeine and theobromine content of various natural and red dutched cocoas. *Journal of Food Science, 46*(6), 1968-1969.

Wadsworth, R. (1921). The estimation of theobromine in cocoa and its products. *Analyst, 46*(1921), 32.

Wadsworth, R.V. (1920). Note on the solubilities of theobromine. *Analyst, 45*(529), 133-134.

Wadsworth, R.V. (1922). The theobromine content of cacao-beans and cocoa. *Analyst, 47*(553), 152-163.

Whymper, R. (1921). *Cocoa and chocolate: Their chemistry and manufacture* (Second ed). London, UK: J & A Churchill.

Woskresensky, A. (1842). Sur la theobromine. *Journal de Pharmacie et de Chimie, 3*(1), 136-137.

Zhang, D., Gardini, E.A., Motilal, L.A., Baligar, V., Bailey, B., Zuñiga-Cernades, L., Arevalo-Arevalo, C.E., Meinhardt, L. (2011). Dissecting genetic structure in farmer selections of theobroma cacao in the peruvian amazon: Implications for on farm conservation and rehabilitation. *Tropical Plant Biology, 4*(2), 106-116.

Zhang, D., Martínez, W.J., Johnson, E.S., Somarriba, E., Phillips-Mora, W., Astorga, C., Mischke, S., Meinhardt, L.W. (2012). Genetic diversity and spatial structure in a new distinct theobroma cacao l. Population in bolivia. *Genetic Resources and Crop Evolution, 59*(2), 239-252.

Zhang, H., Zhang, H., Troise, A.D., Fogliano, V. (2019). Melanoidins from coffee, cocoa and bread are able to scavenge α-dicarbonyl compounds under simulated physiological condition. *Journal of Agricultural and Food Chemistry, 67*(39), 10921–10929.

Zhu, Q.Y., Holt, R.R., Lazarus, S.A., Ensunsa, J.L., Hammerstone, J.F., Schmitz, H.H., Keen, C.L. (2002). Stability of the flavan-3-ols epicatechin and catechin and related dimeric procyanidins derived from cocoa. *Journal of Agricultural and Food Chemistry, 50*(6), 1700-1705.

Ziegleder, G. (2009). *Flavour development in cocoa and chocolate*. In: Beckett ST, editor. Industrial chocolate manufacture and use. Singapore: John Wiley & Sons. pp. 169-191.

Ziegleder, G. (2017). *Flavour development in cocoa and chocolate*. In: T BS, editor. Beckett's industrial chocolate manufacture and use. 5th ed. Chichester, UK: John Wiley & Sons Ltd. pp. 185-215.

Zink, D.L. (1997). The impact of consumer demands and trends on food processing. *Emerging Infectious Diseases, 3*(4), 467.

Zosel, K. (1981). Process for the decaffeination of coffee. Google Patents.

Zoumas, B.L., Kreiser, W.R., Martin, R. (1980). Theobromine and caffeine content of chocolate products. *Journal of Food Science, 45*(2), 314-316.

Zuur, A.F., Ieno, E.N., Elphick, C.S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution, 1*(1), 3-14.

Żyżelewicz, D., Zakłos-Szyda, M., Juśkiewicz, J., Bojczuk, M., Oracz, J., Budryn, G., Miśkiewicz, K., Krysiak, W., Zduńczyk, Z., Jurgoński, A. (2016). Cocoa bean (theobroma cacao l.) phenolic extracts as ptp1b inhibitors, hepatic hepg2 and pancreatic β-tc3 cell cytoprotective agents and their influence on oxidative stress in rats. *Food Research International, 89*(2016), 946-957.

VITA

Alan Patrick McClure was born in Kansas City, Missouri, U.S.A. on the 18th of November 1978 to Linda S. McClure (née Peterson), BS in geology and Thomas L. McClure, MS in psychology. After graduation from high school, Alan first pursued the study of music at Webster University in St. Louis, Missouri, U.S.A., working as a camp counselor for Camp Wyman for underprivileged children during the summer. When he returned to Webster, he founded the Milarepa House for student diversity, which held meetings hoping to create a constructive environment within which students from different backgrounds could interact to better understand each other. During this time, Alan became increasingly captivated by eastern religious traditions such as Buddhism, and this led to an eventual transfer to the University of Missouri-Columbia, to pursue a B.A. in Religious Studies. While finishing his undergraduate career at MU, Alan tutored children in need in math and reading prior to spending approximately a year in the Beaujolais region of France, during which he was engaged for a semester in study at the Université Jean Moulin Lyon III. It was in France that Alan fell in love with craft food production, including chocolate, bread, cheese, and charcuterie. Upon returning to the University of Missouri, Alan graduated summa cum laude with a Bachelor of Arts degree in Religious Studies and departmental honors (2005), and soon thereafter founded what was to become a record-holding, award-winning, bean-to-bar chocolate manufacturing company in Columbia, Missouri (2006). Alan was also co-founder and chair of the national trade organization Craft Chocolate Makers of America from 2008-2014 and served for two years on the board of the Fine Chocolate Industry Association. After

many years of interest in the science of chocolate and food generally, Alan returned to the University of Missouri-Columbia in 2014 to pursue a graduate degree in Food Science under the guidance of Dr. Ingolf Grün in the Milton E. Bailey Flavor Chemistry Laboratory. In 2016, after two years of preparatory course work in the sciences, he officially began study and research as a doctoral student, and subsequently successfully wrote grant proposals that obtained funding for his project, including one from the Professional Manufacturing Confectioners Association (PMCA), allowing him to travel to Penn State's storied Sensory Evaluation Center for a large portion of his research. In the fall semester of 2020, Alan earned the degree of Doctor of Philosophy in Food Science, with a non-designated minor in Chemistry, graduating with a 4.0 grade-point average, prior to embarking upon exciting new adventures in the fascinating world of food and food science.