Specialty Coffee Association

A Specialty Coffee Association Report

Understanding and Evolving the SCA Coffee Value Assessment System:

Results of the 2020-2021 Cupping Protocol User Perception Study and Proposed Evolution

August 2022

Research Team:
Roukiat Delrue, Consultant
Dr. Mario Roberto Fernández Alduenda, Technical Officer, SCA
Katie von der Lieth, Research Programs Manager, SCA (2021)

Publication Team:
Jenn Rugolo, Editor
Dr. Mario Roberto Fernández Alduenda, Reviewer
Peter Giuliano, Reviewer
Kim Elena Ionescu, Reviewer

Table of Contents

About the SCA Coffee Value Assessment System Evolution Project

Since it was created in 2004, the existing SCA cupping system has become a globally recognized industry standard used by many stakeholders across the globe to assess coffee quality and to assist in the determination of a coffee's value in contracts. Meanwhile, our understanding of sensory science and coffee's sensorial properties, as well as the specialty coffee industry and its global context, have advanced significantly.

Despite its ubiquity, the protocol has not been thoroughly investigated or substantially updated by its creators since its foundation. In line with efforts to create more equitable and distributive models of governance (decision-making power) in the specialty coffee market, one workstream of the SCA's sustainable coffee agenda is to rebalance the power of buyers relative to growers/processers in assessments of coffee quality by embedding mechanisms into coffee quality evaluations that resist the hegemony of buyers. Additionally, the SCA has been working to produce a collective mindset shift within the specialty coffee sector by developing, publishing, and amplifying an interpretation of "specialty coffee" that includes not only taste quality, but other attributes that differentiate specialty coffees, including sustainable growing and sourcing practices.

As the existing coffee value assessment system, including the protocol, impacts and is impacted by both work streams, the first step of the review and evolution process was to commission a project in 2020 to better understand how members of the coffee industry currently use the protocol and its perceived strengths and weaknesses in assessing coffee's value and sensory attributes. Another step, also completed in 2020, was to perform a review of the existing literature at the intersection of coffee and sensory science (since published as *The Coffee Sensory and Cupping Handbook* in 2021[1]) and to proffer an updated, descriptive definition of specialty coffee[2] upon which a new framework for a coffee value Assessment system could be built.

[1] Dr. Mario R. Fernández-Alduenda and Peter Giuliano. *Coffee Sensory and Cupping Handbook*, Edition No. 1 (2021).
[2] Specialty Coffee Association. *Towards a Definition of Specialty Coffee: Building an Understanding Based on Attributes* (2021). https://sca.coffee/attributes-whitepaper

Executive Summary

Conducted concurrently with a literature review, the SCA commissioned a research study in 2020 to better understand industry perception of the SCA cupping system. Across a wide-reaching online survey (1575 respondents) and a selection of semi-structured interviews (47), researchers asked: How is the SCA Cupping Protocol being used, and what are its perceived strengths and weaknesses in assessing coffees' value and sensory attributes?

Findings indicate that, although factors outside cupping results influence a coffee's value, the score and descriptor results of the existing SCA Cupping Protocol are frequently used in contracts and price-setting negotiations in coffee transactions. Some technical aspects of the protocol are seen as problematic, including sweetness, reference words, overall, clean cup, body, and complexity. Finally, although there was strong variance in the reported level of its objectivity, the SCA cupping system was rated as a strong tool in creating a common language for the coffee community (except for consumers).

Definitions

By SCA cupping system, this report means the set of standards, protocols, and tools used to assess the quality of arabica coffee, developed primarily by the Specialty Coffee Association of America between 1997-2016. The original purpose of the system was to define "specialty grade coffee" in order to distinguish it from "commercial" coffee, so the foundation of the system was the standard defining specialty grade green coffee and its specifications. The two most important standards in the existing system are the *SCA Green Grading Handbook*, which describes the method to grade coffee in green bean form, and the SCA Cupping Protocol, which describes the procedure and criteria to prepare the coffee sample and assess its quality through cupping (tasting), resulting in a "cupping score."

Although this project focuses on the two elements of the system dealing most closely with cupping (the SCA Cupping Protocol and the SCA Cupping Form), the following definitions were provided to survey respondents and are subsequently used for this report:

SCA Cupping Protocol: The method to cup a coffee, including both the steps to carry out the cupping and the criteria to score it.
SCA cupping results: The outcoming information from the cupping, including the final score, each attribute's score, attributes' intensities and freely elicited descriptors and notes.
SCA Cupping Form: The paper form where the cupping results are recorded.

Project Description and Methodology

Data was collected via an online survey and in-person semi-structured interviews. The purpose of the surveys was to gain a broad understanding of how the SCA Cupping Form and Protocol are currently used and perceived by the larger coffee community. The semi-structured interviews provide more qualitative data and allow for more in-depth and nuanced data collection.

First, the survey questions were drafted and revised. The survey asked respondents about the frequency and context of use of the cupping protocol as well as the detailed technical elements which compose the cupping protocol. The survey questions were reviewed by both the consultant and key members of the SCA's leadership team and, after consensus on content, the survey was translated to Spanish, Korean, Traditional Chinese, and Simplified Chinese. The survey was published in English first, about a week later in Spanish, and a few days after in the Korean, Traditional Chinese, and Simplified Chinese.

In parallel to the survey, the SCA and the consultant determined the interview methodology and content of the research, including who would be interviewed and which consistent prompts would be used (this was to ensure the information was, to some degree, comparable between interviews and easier to interpret). Interviews ranged in time from 30 minutes to 1.5 hours depending on the interviewee, and were conducted over a span of two months from December 2020 to January 2021.

Surveys

Timeline

The Spanish and English versions of the survey were open for approximately two months between November 14, 2020 to January 21, 2021. The surveys in Korean and Simplified and Traditional Chinese were open for shorter timeframe from December 14, 2020 to January 21, 2021. Across all languages, a substantial majority of responses were collected in the first two weeks after each survey's release, and for this reason we do not believe that the shorter time frame for the surveys in Asian languages skew the results relative to the English and Spanish surveys.

Respondent Demographics

The project goal was to obtain 500 "useable" responses (i.e., most if not all questions answered, as some respondents exited the survey after the first few questions); we obtained more than twice that amount.

The number of respondents for each survey language were:

557- English
270 - Spanish
137 - Traditional Chinese
61 - Korean
18 - Simplified Chinese

In terms of geographic dispersion, respondents were located in 86 countries across the globe. The United States and Taiwan were the two most represented countries in the survey, accounting for 16% and 14% of total respondents, respectively. Mexico and Korea were the third and fourth most represented countries, each accounting for 7%.

Survey respondents were unknowingly filtered into two categories, depending on how often they used the SCA Cupping Protocol. Anyone responding that they used the protocol less than once per month was taken to a more simplified version of the survey; respondents answering more frequently than once per month did the full survey. This filtering of respondents had two purposes: the first was to ensure only those that cupped using the protocol on a regular basis responded to the more technical questions and second was to ensure respondents didn't feel "put on the spot" with something not as familiar to them.

Interviews

Selection of Interviewees

Lead SCA staff members and the consultant developed a matrix in order to ensure that at least three key components were considered while determining people to interview, so as to guarantee a diverse representation of stakeholders:

- **Are they a "super-user"?** "Super-users" are those extremely experienced with the cupping protocol and form, often early developers or adopters. Some of these were also SCA board members.
- **Where are they geographically located?** We wanted to ensure most regions of the world were represented and in approximately equal amounts, and split these as North America; Latin America; Europe; Asia and Middle East; Oceania and Africa.
- **What role do they play in coffee's complex value chain?** We wanted to ensure all actors of the coffee system were represented and we used the following categories: retail or barista; roaster; trader, importer, or exporter; producer or processor; and educator, non-governmental organization, or association.

Our goal was to conduct 50 interviews. To provide flexibility, the matrix created was for a little over 75 people. All geographic areas and actors in the value chain were successfully represented at a total of 47 interviews. Additionally, to those 47 executed interviews, around 15 other invitations were sent repeatedly with no response. It is not believed this impacted the outcome in any way as the trends and repeated topics are visible and clear. Most interviewees were avid cupping form users with a few exceptions which were also valuable in terms of perspective.

Timeframe

Interviews were conducted from December 1, 2020 to January 22, 2021. The average duration of each interview was around 45 minutes.

Interviewee Demographics

All value chain roles as previously defined were successfully represented in the interviews, including even some additional roles such as authors or developers of diverse cupping methodologies. The interviewees were from approximately 30 countries, including but not limited to: Burundi, Kenya, Guatemala, Brazil, Colombia, Taiwan, Singapore, USA, Indonesia, Sweden, Ireland, Poland, Russia, and Greece, among others.

Data Analysis

For the purposes of data analysis, the responses to the surveys were split into two parts: the "Common Questions" part (questions 3-15, after country and role in coffee), with the questions made to all respondents about the value, usefulness and objectivity of the SCA Cupping Protocol, and the "Protocol Questions" (questions 16-28), for which non-SCA cuppers (respondents reporting they use the SCA Protocol less than once a month) were exempt, asking about details of the SCA Cupping Protocol, its strengths and weaknesses.

A table was produced for the Common Questions, with respondents as rows and response variables as columns, integrating the responses from all languages into a single table. Responses from different languages were coded using a simple number code, where higher numbers imply a more positive perception of the SCA Cupping Protocol. Descriptive statistics for each variable were done. A table of correlations between all variables was produced. An analysis of variance (ANOVA) was done, to determine the effect of language, country, role in coffee and cupping frequency (both with any method and SCA) on the rest of the Common Questions, and a principal component analysis (PCA) was also done, using the variable where such significant interaction with demographics occurred, to explore the main trends between demographics and attitude towards the SCA Cupping Protocol.

The Protocol Questions (16-28) were analyzed using the question-by-question (univariate) approach used by SCA in the past for other surveys, as this seemed more fitting for the type of data collected on them.

Findings and Implications

Although we cannot speak to causation, there was a positive correlation between both the frequency the system's use and an actor's proximity to coffee production activities in relation to more positive sentiment about the cupping system. Respondents were also generally positive about the cupping system as a tool for price discovery as well as a widespread common language within the industry. In regards to the question of subjectivity and objectivity, respondents across the survey and interviews were split—but in light of sensory science advancements, this is an understandable result (the existing protocol and form are a mix of subjective and objective sensory tests). Most respondents remarked that the existing system only works as a common language if users are well-trained.

Across the form's specific sections and functionality, questions were raised about reference words, the placement of some boxes ("Uniformity," "Clean Cup," "Sweetness," and "Defects"), and the name of some categories ("Overall" and "Body"). Many respondents also noted the form could be more ergonomic, or even digitized. Beyond the form, there was also a call to adapt the protocol itself as both the underlying standards and roasting technology have changed.

Notably, many respondents—while identifying key areas or improvement to improve the systems facility—also raised concerns that any small change could affect the livelihood of a producer or group and requested a considered approach to its evolution.

Interactions Between Respondents and Responses

The clearest interaction between a respondent's characteristics and their attitude towards the SCA Cupping Protocol relates to respondents' frequency of protocol use. There is a very clear correlation between how frequently a respondent uses the SCA Cupping Protocol and their general appreciation for it in terms of usefulness, strength, and objectivity. There are two opposite ways to explain this situation: (a) respondents find the SCA Cupping Protocol is useful, and thus they use it frequently, or (b) frequent users learn to appreciate the SCA Cupping Protocol's strengths as they practice more.

Both situations could be happening in real life, depending on each respondent. However, cupping frequency is not always a person's choice: some must cup frequently because of their position in coffee, and the choice of cupping method is usually a company policy, rather than an individual choice. For this reason, explanation (b) is perhaps more likely: people who need to use the SCA Cupping Protocol frequently for whatever reason, spend more time using it and find more reasons to appreciate the system and its impact. This can be likened to the effect termed "anchoring bias" or "cognitive bias."

The relationship between a respondent's role in coffee and their responses, as identified in one of the produced principal component analysis plots, is very revealing. The four roles with the most positive attitude towards the SCA Cupping Protocol were cooperatives, processors, producers, and exporters, in that order. The four roles with the most negative attitude were consultant, importer, other, and roaster. This offers interesting perspective on a common sentiment (and critique) of the protocol, that producers and cooperatives benefit least from the use of the system. Interestingly, this sentiment is not stated by producers, cooperatives, processors, and even exporters—who were the most appreciative of the system in the results, and who stated they felt empowered by the protocol—but by others in different roles. At the same time, roles closer to the consumer perceived that producers and cooperatives were not benefiting from the protocol. Perhaps roasters and importers know something producers don't know, and they realize that producers are not getting as much from their coffee as they should, based on cupping scores. An alternative, and perhaps most likely, explanation is that of selection bias among the producer respondents, who are likely to be specialty coffee producers, well-connected to the internet, and in a position to be empowered by their use of the protocol and form.

It is also worth considering why consultants, importers, and roasters were the roles least enthusiastic about the cupping system. Some have developed their own systems and are thus critical about other methods or feel the SCA protocol is cumbersome or complicated. Some highlighted that the protocol is not designed for many of the applications they need, like development and quality control of roasted coffee products. Perhaps another source of this frustration is the system's inability to effectively communicate coffee quality to consumers—the survey results indicate it's a much better communication tool within the industry, which might explain why the "first links of the chain" feel more positively about it (their customer is another coffee person!). Roles closest to the consumer may struggle with using a tool that doesn't effectively communicate a coffee's attributes to their own customer/consumers.

Regarding the relationship between a respondent's language and their response, our results are not very clear. The more positive perception among Spanish language respondents, compared to other languages, may be explained from the demographics of the group, with a much larger percentage of producers, processors, cooperatives, and exporters compared to the other languages. It could also be partly due to a cultural tendency among Spanish speakers to try to please. On the other hand, the Korean language responses, though with a large dispersion, reflect in general a more negative attitude towards the cupping system than the other languages. We do not have an explanation for this trend.

Price, Contract, and Transactions

The results of the protocol's use were generally perceived to be a useful tool for price discovery. Through the stakeholder interviews, we found that many contracts buying and selling coffee included (either implicitly or explicitly) a coffee's cupping score. Spanish-speaking respondents, who represent the majority of producing-side businesses captured in this survey, were especially positive about the usefulness in price discovery: they often drew a strong relationship between higher cup score and higher price and also reported that knowing a coffee's score served as a reference for setting a price. That said, respondents closer to consuming businesses drew a less linear relationship between score and price, noting other that there are factors other than influence price. Finally, and somewhat paradoxically, price discovery was selected by survey respondents as the second-to-least used application of the system.

A Widespread Common Language

Both the survey and the interview respondents noted the unique value of the SCA's cupping system, particularly its widespread adoption by the industry along the chain and across the globe to create what was commonly described as "the common language of specialty coffee." Survey opinions about this value were the least dispersed piece of data (81% of respondents rated the SCA Cupping Protocol as a "strong" or "very strong" tool to create a common language), while interviewees pointed out that this common language created by the protocol is in fact one of the main assets of the system, as it greatly helps aligning the expectations between the buyer and the seller.

However, both the survey results and the interviews show that, though a great communication tool within the industry, the system is not an effective way to communicate a coffee's attributes to consumers: there is no standard way to "transfer" descriptors from a cupping and convey them onwards. Also, with most coffees scoring between 82 and 88, the cupping score—supposedly an intuitive 100-point scale, designed to be consumer-friendly—is in fact confusing to consumers; besides, many final consumer-facing products have lost any direct link to the flavor that originated the cupping score for the green coffee, as roasted and brewed products oftentimes have little resemblance with the coffee used for cupping.

Objectivity, Subjectivity, and Intersubjectivity

The topic of subjectivity and objectivity was polarizing among respondents, who were asked to rate the objectivity of the cupping system on a scale of 0-100, where 0 was "not at all objective" and 100 was "completely objective." Analysis identified four distinct groups of respondents: 58% of respondents rated the system's objectivity at 50 or above, and this group was centered around a score of 75; 42% of respondents rated the system's objectivity at 49 or below, with this group centered around a score of 45. However, 2% of respondents rated the system as "not at all objective" (a score of 0), and 5% as "completely objective" (a score of 100). Using multivariate analysis, it seems that Spanish language respondents tend to perceive the SCA Cupping Protocol as more objective than respondents of other languages, and respondents who never use the SCA Protocol tend to rate its objectivity as lower than rated by respondents who use it.

The question of objectivity was further explored in question #20 of the survey, where respondents were asked to rate the level of objectivity of each of the SCA cupping categories and form sections. The perception of "higher objectivity" among Spanish-language respondents was also clear here across all categories. The categories rated as least objective in both languages were "Overall" and "Sweetness," while "Uniformity," "Acidity," and "Defects" were the categories considered most objective in both languages. Some of the interviewees explained these perceptions: "Overall" is sometimes referred to as "cupper's points," and is generally viewed as the one category where cuppers are allowed to express their opinions, likes, and dislikes. "Sweetness" is usually a 10-point "giveaway" for most coffees and does not reflect the level of a coffee's perceived sweetness. "Uniformity," on the other hand, does not require a cuppers judgement but acts as a sort of discriminative sensory test. The category "Defects"—particularly for the most widely known defects such as phenol, ferment, onion, and potato—do not require cuppers' judgement: if the off-flavor is present, a defect should be marked. These nuances explain why the majority of respondents believe the protocol lies around halfway between "not at all objective" and "completely objective."

The prevalent opinion across interview respondents is that the protocol is objective when cuppers are well-trained. This is closely related to the concept of calibration: no matter their personal opinions as coffee persons, cuppers are able to "calibrate" against a group of peers or against "their market," meaning they are able to pick up the scoring criteria from the group and adjust their results accordingly. One or two of the interviewees call this process "intersubjectivity"—cuppers' results do not reflect an objective reality nor each cupper's individual hedonic reactions to a coffee. Instead, cuppers strive to follow "cupping criteria" which define an attribute's desirability or undesirability for an abstract collectivity sometimes referred to as "the market" or "the industry," though sometimes it can be recognized as "the buyer," "the senior cupper," or even the group of peer cuppers. A few interviewees even said the SCA Cupping Protocol is completely subjective, though they also said that does not make it less valuable, as it makes it possible to align with the buyer's likes and dislikes, which also brings us back to the idea of intersubjectivity.

A method's level of subjectivity or objectivity is not subject to democratic vote—it cannot be stated that a majority of people perceive the SCA Cupping Protocol to be objective, therefore it must be objective. Rather, this reveals the industry's

perception, though the "real" level of the method's objectivity is not easy to assess. Philosophy and science have discussed objectivity and subjectivity for millennia, and both terms are fraught with connotations ("I think you are not being objective," is another—perhaps polite—way of saying "I think you are wrong"), although we have historically been led by physics in our understanding more recently. For sensory science, the concepts of objectivity and subjectivity have dramatically changed as the discipline has evolved. In the mid-twentieth century, any kind of sensory assessment was considered subjective, compared to chemical tests, which were considered objective.[3] Today, the generalized view in sensory science is that an "analytical" quality (such as taste intensity or body level) is considered objective, even if not all observers perceive it, whereas a value judgement (such as cupping score, grade, preference, liking or acceptability) is considered subjective.[4] The SCA Cupping Protocol would therefore be a mixture of objective and subjective assessments (hence the respondents' perception of the protocol being halfway between both extremes), where the intensity ratings, descriptive notes, "Uniformity," and "Defects" are examples of the more "objective" parts; "Overall" represents the most subjective part (the individual cupper's hedonic response); and the rest of the scores ("Flavor/Aroma," "Fragrance," etc.) represent an effort to interpret someone else's hedonic response, ideally "the market's."

Training or Experience Required

Many respondents remarked that the existing system only functioned as a common language if users were trained, experienced, or calibrated to the protocol and form. Most times, this was stated without judgement attached, but rather simply as fact. All respondents understood effective use of the system requires training, experience, and calibration.

Reference Words

Responders across the survey and the interview often questioned the reference words listed on the form. Some questioned if our understanding of these reference points have changed, or if the words have been insufficient all along. Others suggested that, as most coffees evaluated with the form are specialty coffees, most are arguably "good" if compared to "gas station or airline coffee."

As such a widespread tool, so reliant on training and education, providing reference points is extremely important for users to learn how to use it properly. However, we note that the current reference words are not assisting users: several interviewees mentioned that if all coffee is scored at "very good," it does not even cross the existing threshold of a specialty coffee score! The reference words are misaligned with the scores.

[3] Sjöstrom, "Correlation of Objective-Subjective Methods as Applied in the Food Field."
[4] Carpenter, Lyon, and Hasdell, *Guidelines for Sensory Analysis in Food Product Development and Quality Control.*

Uniformity, Clean Cup, Sweetness, and Defects

Although the boxes for "Uniformity," "Clean Cup," "Sweetness," and "Defects" sit towards the end (or right) of the cupping form, several respondents suggested this section should be moved to the front (or left) of the form, as it is often the first evaluation marked.

Additionally, several respondents highlighted the need for an expanded section to discuss evolving issues that aren't captured by the existing categories. One example given was the evaluation of aged coffee—the perception of an (unintentionally) aged coffee is not technically "defective," so how should a user incorporate this into their evaluation? (It's important to note that "perception" is a key word, as these flavors could potentially be caused by more than just aging.) The same problem exists with other contaminations because the definition of both "clean cup" and "defects" have taken different interpretations through time. Respondents indicated that these kinds of issues are often clear, have names, and are very limited in number, so they could all be listed on the protocol in the same way green defects are. All "other" issues need an indication of how to be reflected on the form.

Sweetness

Of all the attributes evaluated by the system, the category of "sweetness" was the most referenced as the one needing adjustment. Many suggested an "obvious" solution, to replace sweetness evaluation as a presence (descriptive) to a quality assessment (quality scale), with or without an intensity scale included as reference of the evaluator.

However, some of the respondents were completely aware that this small change would instantly affect the final score—which can, in turn, affect the price a producer receives for their coffee. Additionally, from our concurrent literature review into coffee sweetness and sensory evaluation, we know that just like the other attributes captured in the protocol, this category doesn't offer an "objective measurement of sweetness," but rather, a perception of sweetness which can be affected by many things, like the type of aromas perceived.

Any evolution of the cupping system will need to rethink the presence and evaluation of sweetness, bearing in mind that any change to this important yet complicated section stands to impact the livelihoods of coffee producers. It is clearly an important category to evolve, but incredibly complicated to change.

From "Overall" to "Complexity"

As mentioned previously, the "Overall" category is perceived as the least objective, with some respondents remarking they felt it was almost an afterthought to just make the total get to 100 or as a replication of "Balance." The survey offered "complexity" as an alternative to "Overall," which yielded mixed results in the interviews. Those who had already responded to the survey often suggested "complexity" as a replacement for "overall," while those who hadn't yet participated in the survey did not.

There was no consensus in how to approach the evolution of this category. While some experienced cuppers have decided to not consider it "cupper's points" as the protocol indicates, everyone seems to have determined their own definition for it, which for the sake of common language, needs attention.

From "Body" to "Mouthfeel"

The word "body" as a descriptor of categorical evaluation causes issues for some users, likely in some cultural contexts. Most interviewees from Europe or Asia held strong opinions that a more accurate word for this evaluation is "mouthfeel."

If the definition and protocol is improved, mouthfeel could include "astringency" which would in turn, provide an approach for respondents who requested that this attribute be considered (however this needs to be analyzed further).

Other alternatives such as viscosity, tactile evaluation, or others could be applicable as well.

Moving into the Future: Ergonomics, Complexity, Digitization

Many respondents noted the design of the form—particularly the small space available for notes, small boxes for numbers, and small lines for scales—could be evolved to create something more ergonomic. Some, although infrequently, noted newer forms have been designed in a vertical format. For some, the digitalization of the form was deemed very important, but it was less so for others, particularly those noting that accessibility to everyone (particularly producers) might be slightly more challenging than the paper and pencil form currently available.

Finally, while some indicated the form is very thorough and with that, complex, others said it would actually benefit from including some flavor modulation components. In the survey, many mentioned the notion of "complexity" of the coffee being added as an element, specially to replace "Overall." For others, the type of interactive scales as seen in the robusta form ("Bitter-Sweet" ratio and "Salt-Acid" ratio) could be incorporated. Of these, some suggested the same exact type of scale, while others suggested a type of interactive scale linking the fermentation balance to the taste balance.

Not Just the Form: Evolving the Protocol

Respondents across the survey and in interviews identified that, although the industry has learned to use the existing form "as is," the protocol itself also needs adapting. Beginning with sample identification and initial assessment, respondents identified that green grading, quakers, water activity, and the inclusion of a roast protocol are required. Not only have the standards underlying these elements of the protocol changed (sometimes without explanation), but the existing standards don't apply to new techniques and technologies (i.e., sample roasts conducted in an air-roaster).

This desire to evolve the protocol also extended to the cupping procedure itself, particularly the number of cups used. Some interviewees, most linked to production, noted that five cups per sample is too many, particularly when evaluating quality several times. Conversely, some traders mentioned needing to cup up to ten cups per sample. Other questions included: Is steeping time at the suggested range ok? How well is it actually enforced, or how different is steeping time from Sample 1 to Sample 6? Many also expressed concerns around the potential of the protocol to spread infectious disease (understandable, given the timing of the project during the COVID-19 pandemic), as many said they struggled with the temporarily suggested method.

Caution and Impact of Changes

Although many users felt several changes could improve the facility of use for them, many specifically emphasized their awareness that any small change could affect the score, which could in turn affect a price and the livelihood of a producer or group. These respondents also emphasized that we should be very considered in our approach to change for this reason.

Furthermore, respondents noted how widespread the protocol and form currently are (a very simple example: the number of people willing to participate in this research project!). Everyone manages change differently; the success of any transition to an evolved system will be dependent on the SCA's ability to provide adequate supporting tools and training.

A Vision for a New SCA Coffee Value Assessment System

Building on the findings of the literature review and perception study, the research team has proposed a vision to evolve the existing cupping system into its next phase as a coffee value assessment system. The new system will be founded upon the newly-released attributes-based definition of specialty coffee and advances in sensory testing best practice (including the separation of "descriptive" and "affective" testing within the updated protocol and form). It will also incorporate two new sections, one capturing coffee's extrinsic attributes and another offering tool for value discovery, as mechanisms to resist the hegemony of buyers.

The research team recognizes that not only will any evolutionary changes take time to move throughout coffee's complex value-generating system, but that it also stands to impact producer livelihoods. With this in mind, we propose a phased approach to executing the vision.

A New Foundation

The previous cycle of the SCA cupping system was based on the existing definition of specialty coffee at the time of creation, namely the "Specialty Grade" standard for washed green arabica coffee. Although the SCA Cupping Protocol and various other standards and tools remain in effect, the Specialty Grade specifications are no longer easily found on the SCA's website nor in its publications, save for the *Arabica Green Coffee Defect Guide*. This has made the current standards and tools seem "orphan" and the whole system somewhat incongruous. If a new, reengineered version of an SCA Coffee Value Assessment System is to be developed, for the sake of congruency, it needs to be built from a new definition of specialty coffee.

If we want the SCA's cupping system to help actors discover a coffee's quality, attributes, or value, we first needed to begin by defining how these concepts are understood in today's specialty coffee industry. For some time now, this new definition of specialty coffee has slowly been evolving from internal discussions among SCA staff, from some lectures presented at SCA events and from external research, showing how a coffee's value is discovered. This definition has since been refined and proposed in a 2021 SCA white paper, *Towards a Definition of Specialty Coffee: Building an Understanding Based on Attributes*.[5]

Any product, including coffee, has different kinds of attributes, which may be valuable for the potential user or buyer. Some of these attributes are *intrinsic* to the product—in the case of coffee, examples of intrinsic attributes could include its chemical composition, physical properties and appearance, and its sensory properties, namely flavor. Some attributes are *extrinsic*, meaning we cannot discover them from the product itself, but need additional information. Examples of extrinsic attributes in the case of coffee include traceability and sustainability information about the coffee lot: where it was produced, by whom and how; what variety it is; how it was processed; what is its relationship with the environment, with the humans along the chain and with the economics along the chain, and many others.

A potential buyer could value some of these attributes more than others. For some, perhaps a coffee is most valuable when it shows unique flavor or specific flavor characteristics; for others, perhaps a coffee is most valuable when we can be sure the farmer was paid a fair price, when carbon was captured during its production, or when it was produced in a bird-friendly environment. Thus, the same coffee could be highly valuable for one person and have very little value for another person.

Therefore, the SCA has chosen to define specialty coffee as "a coffee or coffee experience recognized for its distinctive attributes, and because of these attributes, has significant extra value in the marketplace." Conversely, coffee which does not present these valuable attributes—or for which attributes are not discovered along the chain—remains undifferentiated and will probably end up as commodity coffee.

[5] https://sca.coffee/attributes-whitepaper

A transparent approach to discover a coffee's value for a *specific* buyer would be to compare their needed or sought-after attributes in a coffee versus the attributes shown by the coffee in question. In this way, for example, if a buyer needs a fair-trade coffee with a distinct chocolatey flavor, it will become clear to the seller why they reject or offer a low price for a coffee which, though organic and high-scoring, does not exhibit any of the sought-after attributes.

Conversely, if a coffee exhibits rare attributes which are sought-after by the buyer, the seller should expect a fair premium.

The Concept

We propose a revision of the SCA Coffee Value Assessment system, building on this new foundation of attribute-based value discovery in specialty coffee, with three objectives: enhanced usefulness for the community, congruency with sensory science, and transparency about how a coffee's attribute compare with the buyer's needs. For these purposes, an assessment of coffee in four separate "compartments" is proposed. They are called "compartments" because each of the four evaluation steps is independent and should be completed separately to avoid bias when assessing the intrinsic attributes: tasters and cuppers can be easily biased if they are given access to extrinsic information about the coffee.

Descriptive Compartment

This is the assessment where "objective," intrinsic attributes are assessed and recorded. The core of this assessment is a flavor descriptive analysis of the coffee—an analysis that can be done to coffee in different points along the chain (parchment, green, roasted, brewed...) and can therefore be used as a common language about a coffee's flavor attributes to communicate along the chain and even to consumers. Though the *World Coffee Research (WCR) Sensory Lexicon*, the SCA/UC Davis/WCR Flavor Wheel, and other tools would be used as references, this would not be classic descriptive analysis: the analysis could be broken by tasting step (fragrance, aroma, flavor, aftertaste, mouthfeel, etc.) and each of these steps would call for a description. Intensities could be grouped by step, as they are now (fragrance intensity, aroma intensity) or broken down by attribute (e.g., sour and salty tastes as part of the flavor "step"). Since it is not possible to rate the intensity of all WCR Sensory Lexicon attributes, the main categories of the SCA's Flavor Wheel (*Floral*, *Fruity*, *Sour/Fermented*, *Green/Vegetative*, *Other*, *Roasted*, *Spices*, *Nutty/Cocoa* and *Sweet*) could be rated by intensity or just marked when present, as in a check-all-that-apply (CATA) test. Key information could be extracted from this assessment compartment to communicate to consumers or to inform specific product- development applications, such as the roasting profile development.

Cup Quality Compartment

Cup quality assessment is the part most resembling our current cupping method, and some essential features should be retained: a final score will be produced by adding up the scores of the different cupping "categories" or steps, which are rated separately, though this does not mean we will keep the same rating scales or even the same categories we currently have. This assessment must be done without disclosing any information from the other assessments to the cupper. The current protocol's shortcomings identified during the survey will be addressed here. This means that a "Sweetness Perception" and a "Complexity" category could be included as part of this assessment.

Extrinsic Attributes Compartment

In this compartment, a coffee's extrinsic attributes are recorded. These may include all types of information about a coffee's history, including relevant traceability, origin, processing, and sustainability information. A standard form for capturing this information could be promoted, so that relevant information about a coffee which is currently overlooked might "accompany" a coffee lot in the

future. Just the amount of information available about a coffee lot could be indicative of its value from the extrinsic attributes point of view: a lot for which all types of information are known would immediately become more valuable for certain buyers than an anonymous lot with little more than a country of origin and a lot number.

Value Discovery Tool

This compartment is not really an assessment but a tool that aids in a coffee's value discovery through comparing the attributes that are valuable or sought-after by the buyer versus all the attributes found in the coffee in question, brought here from the other compartments.

Continuing Evolution: Next Steps

The vision outlined in the prior section might take a minimum of two or three years to be accomplished; it requires working with partners and executing multiple projects to research and design iterations of the different components of the SCA's cupping system. In the meantime, the research team has identified a number of short-term updates to the existing system, as a phased approached to the changes or as a compliment to the outlined vision. While none of the suggestions given by the interviewees are dismissed, given the collective experience of the authors, there is a deeper understanding of the implications and impact of each category.

These potential updates have been grouped by their level of impact, with "mild" or short-term updates, which do not impact scoring, and "medium" or mid-term updates, which might affect scoring. These are NOT mutually exclusive but rather complementary to each other through time. Both categories below imply re-writing the protocol as well as a thorough pilot program of alpha and beta testing prior to official change implementation.

"Mild" or Short-Term Changes (Does Not Affect Scoring)

These changes are currently in an alpha-testing pilot project at the time of this report's publication (July 2022) and will shortly be released as an early-adopter beta testing program in the coming year (by 2023):

- "Face lift" and better overall form design
- Move UCS to the front of the form
- Change "Body" to "Mouthfeel," which includes "Astringency" as a part of the definition
- Update reference words
- Re-write the SCA Cupping Protocol, tailoring for clarification, industry standards, better definitions, etc.
- Offer lectures and workshops on these interim changes and the new proposed protocol

"Medium" or Mid-Term Changes (Affects Scoring)

These changes require significant further research and testing as they are likely to impact scoring, some of which is (at the time of writing) already under way:

- Develop a "perception of sweetness in coffee" research project to determine course of action for the "sweetness" attribute
- UCS + Defect: only one defect deduction (no T / F distinction)
- More vertical scales to expand the descriptive element of the cupping form

Conclusion

Everyone manages change differently, and we know we need to tread extra carefully with something as far-reaching and globally understood as cupping. While the existing SCA Coffee Value Assessment system is familiar to many—and its evolution presents an incredible opportunity to integrate what we now know about how, when, where, and by whom coffee is valued—it has also become increasingly enmeshed in our ways of buying and selling coffee, featuring regularly in contracts. In late 2020, as we embarked on our first step of the evolution process with a wide-ranging survey of 1700 coffee professionals across the value chain to understand how the form is currently used and perceived by the larger coffee community, we found the perception of the form and protocol as tools were generally positive, particularly with producers and cooperatives. We also captured a "wish list" of potential improvements, which we then categorized based on their radicality (i.e., how significant would this change feel?) and their potential to impact scoring. Any evolution will need to be consistent with the SCA's sustainability agenda, to better understand how coffee is valued so that actors in the coffee chain can have more equitable access to the value that specialty coffee creates. In addition, as cupping is a sensorial activity—and coffee's sensory attributes are an important factor in determining a coffee's value—any evolution will also need to factor in the advancements in sensory science to create a robust tool.

Just as the original cupping protocol was based on a definition of specialty coffee, specifically the specialty grade standard for washed green arabica coffee, the evolved cupping system will need to be based on a definition, too. A new, attributes-based definition of specialty coffee builds a framework that respects diverse consumer preferences while simultaneously strengthening producers' understanding of how to communicate and increase the value of the coffees they produce.

An evolutionary concept—to "compartmentalize" four evaluation steps—was internally proposed in early 2021 to avoid bias when assessing a coffee's intrinsic attributes and to maximize cupper focus and effectiveness in assessing descriptive attributes and affective liking, with initial prototypes since developed for testing. At the time of writing, in addition to initial alpha-tests run in collaboration with World Coffee Research to understand how cuppers actually use the current protocol when buying and scoring coffees in the real world, the SCA has since completed its first round of initial beta-tests of the descriptive and affective cupping sections with an initial group of coffee producers and exporters. In the coming months, the SCA will a call for volunteers to join an "early adopter" group for further testing and feedback as we refine the design and presentation of the form and protocol documents.

Oak Lodge Farm, Leighams Road,
Bicknacre, Chelmsford,
Essex, CM3 4HF
United Kingdom

505 Technology Drive, Suite 340
Irvine, California, 92618
United States

sca.coffee