White paper

Confidence you can defend:

How p-value analysis supports field chemical identification

In field chemical identification, confidence must be earned. Operators are often asked to make rapid decisions based on unknown powders, liquids, residues, mixtures, or suspected narcotics in uncontrolled environments. A result that appears simple on screen may still depend on complex analytical assumptions beneath the surface.

That is why the method used to generate a chemical identification matters. Many handheld Raman and FTIR instruments rely on spectral similarity scoring to compare an unknown sample against a reference library. Thermo Scientific™ handheld analyzers use a statistical p-value approach designed to go further: The method evaluates whether the differences between the sample and reference are statistically meaningful or consistent with normal measurement variability.

This application note explains how p-value analysis helps reduce ambiguity, improve confidence, and support more defensible decisions in real-world safety and security applications.

Hit quality index (HQI)

Spectroscopic chemical identification relies on reference libraries that compare an unknown spectrum with Raman or FTIR reference spectra stored on the device. Much like fingerprint identification, the system looks for patterns that align between the unknown sample and known references. The closer the patterns match, the higher the similarity score. This similarity score—called a hit quality index (HQI) —is calculated for each reference spectrum in the library. The top HQI scores are displayed for chemicals that have the greatest spectral similarities with the unknown sample.

The similarity score is a measure of the overall resemblance between spectra i.e. how closely their spectral features align. A value of 1 indicates a perfect positive correlation between the spectra, while a value near 0 indicates little to no relationship in their spectral patterns. In many algorithms, thresholds such as 0.95 are commonly used as practical cutoffs for declaring a match. However, this value is a peak shape similarity metric, not a match probability. A score of 0.95 does not mean there is a 95% chance the materials are the same, nor does it represent 95% statistical confidence of a match. A significant HQI limitation is that the similarity score does not account for noise in the measured sample. In real-world measurements, an unknown sample exhibits more noise than a less noisy reference spectrum in the library. As a result, the score calculated appears more reliable than it is.

It should be noted that the HQI score is heavily affected by large peaks within a given spectrum. Large peaks, such as those from active cutting agents can strongly influence the HQI score for both Raman and FTIR techniques. Smaller peaks, particularly from low concentrations of highly potent designer narcotics, can be missed by this method if large peaks are also present in the spectrum.

Because HQI focuses on overall spectral similarity, it may not adequately distinguish subtle but meaningful differences between a chemical mixture spectrum and a single-component reference. Accurate mixture identification is especially important in narcotics investigations, where street samples and altered brand-name pills often contain multiple active ingredients, cutting agents, or contaminants.

Statistical evaluation: p-value

Given the accelerated pace of evolving illicit drug formulations, law enforcement agencies must be able to reliably identify potent substances within complex mixtures. Conventional similarity-based techniques, such as HQI, were designed for controlled laboratory environments. By contrast, impure samples and variable environmental conditions encountered during field testing necessitate the use of more robust spectral analysis methods.

To address this need, Thermo Scientific™ Handheld Raman and FTIR analyzers employ a statistical p-value spectral analysis. Rather than simply asking, “How similar do these two spectra look?”, this approach evaluates a more meaningful question: “Given normal measurement variability, what is the probability that the sample spectrum differs from the selected library reference?”

By employing statistical analysis rather than relying solely on spectral similarity, the p-value algorithm assesses whether observed differences align with expected measurement variations or signify a genuine chemical mismatch. Results are disclosed only when they meet predetermined similarity thresholds, thereby reducing the likelihood of false positives and supporting more confident decision making in operational settings. The acceptance criteria for p-value “pass,” “similar,” or “fail” ratings is color-coded on the device screen: Green denotes a positive chemical identification, orange denotes chemicals that have similar spectral features, and red denotes no chemical identification, respectively. Alternatively, dangerous chemicals can be “screened” to ensure relevant information is displayed to the user.

This statistical approach provides clearer and more reliable identification, particularly when analyzing chemical mixtures such as common pharmaceuticals or potent narcotics containing cutting agents. The p-value method is thus well suited for real-world safety and security applications that may include identifying chemicals from a broad range of sources including clan labs, street narcotics, and military CBRNE missions.

To illustrate the superior performance of Thermo Fisher’s p-value identification, the Raman mode of the Thermo Scientific™ Gemini™ Chemical Identification Analyzer was used to analyze a set of chemicals that produce similar spectra. Ammonium nitrate, an ingredient used in homemade explosives, is a substance of concern in law enforcement and public safety, making correct identification and differentiation from similar chemicals critical. Spectral data were compared from the two main identification methods discussed above: The traditional Hit Quality Index (HQI) and the statistical p-value method. Both methods correctly identified ammonium nitrate. However, HQI provided multiple results which could include additional false-positive chemical identifications, thereby potentially delaying or complicating operator response. By comparison, the p-value method provided one clear result which the operator can confidently use for decision making (see Table 1).

Table 1 shows the p-value result using the Raman mode of the Gemini analyzer, along with the derived HQI values for the identification of ammonium nitrate. The p-value identifies only ammonium nitrate. Since HQI assesses spectral similarity on an overall basis, it is possible for chemically unrelated materials to appear as strong matches. In this case, HQI yields minimal differentiation between ammonium nitrate, 2-chlorobenzoyl chloride, and nicotinamide. Although the latter two substances

(a corrosive compound and vitamin B3, respectively) are chemically distinct from ammonium nitrate, their HQI scores are very close to the true positive HQI value of 96.2, requiring further user interpretation of the results.

Table 1. Gemini analyzer Raman p-value results and derived HQI data for the identification of ammonium nitrate

Identified chemicals	P-value result	HQI result
Ammonium nitrate	Pass	96.2
2-Chlorobenzoyl chloride	No match	94.8
Nicotinamide	No match	94.1

Though the Raman spectra of the 3 chemicals discussed above share some similarities, they also contain meaningful differences that are not reflected in the HQI score. By contrast, the p-value result unambiguously identifies the unknown sample as only ammonium nitrate.

Figures 1a and 1b show the displayed spectra of these three chemicals from the Gemini device. Note the visual spectral differences between the ammonium nitrate spectrum overlaid with 2-chlorobenzoyl chloride spectrum (1a), and those between the ammonium nitrate spectrum overlaid with nicotinamide (1b).

Displays for chemicals tested using Gemini analyzer Raman mode:

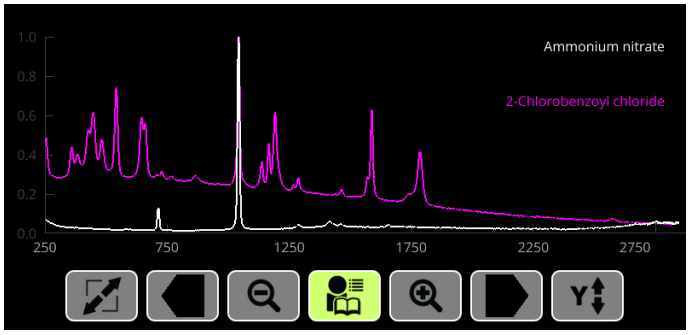

Figure 1a. Ammonium nitrate (white) overlaid with 2-chlorobenzoyl chloride (purple).

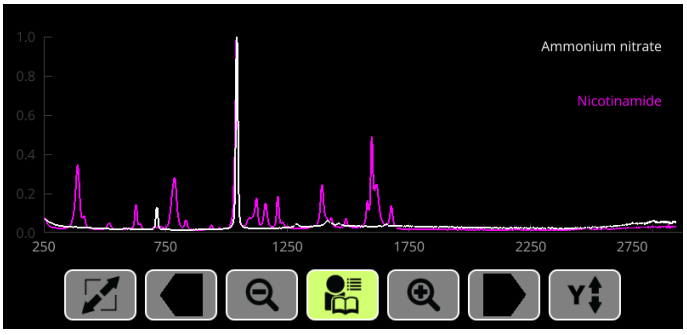

Figure 1b. Ammonium nitrate (white) overlaid with nicotinamide (purple).

The p-value algorithm also provides fast, reliable chemical identification of mixtures. The tactic of leveraging the p-value approach to automatically evaluate potential mixture components is based on the evaluation of the full spectral profile. This method extends statistical analysis from the single-component identification algorithm to accurately identify the best match of up to 4 mixture components. The p-value algorithm helps deliver results with both speed and accuracy, proven through extensive field testing.[1,2,3]

For an illustrated example of the p-value performance for a sample mixture analysis, consider an 80/20 mixture of caffeine and heroin, (weight%, respectively). Caffeine produces several strong peaks (Figure 2a, yellow spectrum) that can dominate similarity scoring such as HQI, while heroin contributes additional spectral features that are small by comparison (Figure 2b, yellow spectrum).

Displays for spectra of pure chemicals and mixtures tested using Gemini analyzer Raman mode:

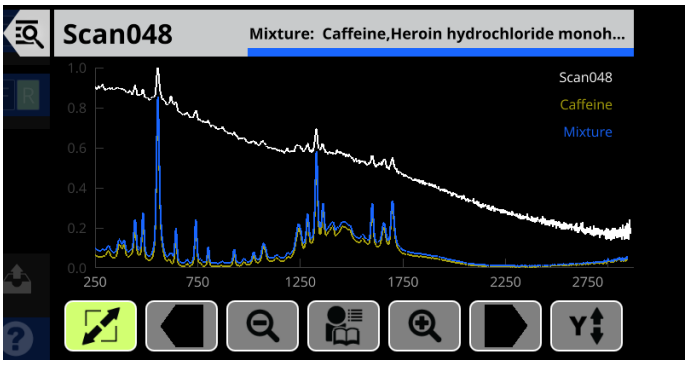

Figure 2a. White spectrum is the scan of 80/20 caffeine/heroin sample. Yellow spectrum is caffeine; blue spectrum is the algorithm calculated spectrum generated from the library of both caffeine and heroin.

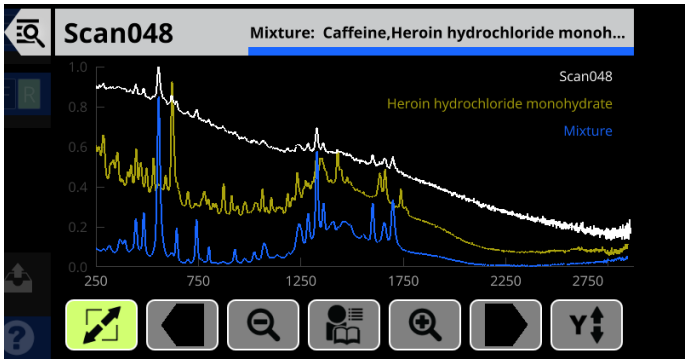

Figure 2b. White spectrum is the scan of 80/20 caffeine/heroin sample. Yellow spectrum is Heroin HCl•H_2O; blue spectrum is the algorithm calculated spectrum generated from the library spectra of both caffeine and heroin.

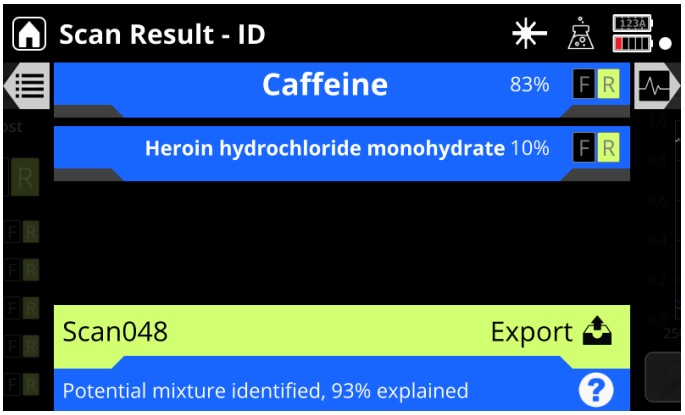

Figure 2c. Gemini analyzer results screen for the analysis of 80/20 caffeine/heroin. The algorithm has correctly identified the mixture of the sample and indicated that the mixture spectrum is generated by a high contribution from caffeine (83%) and a smaller contribution from Heroin HCl•H_2O (10%).

In this mixture example, the p-value algorithm of the Gemini device has correctly identified that the sample spectrum is a combination of caffeine and heroin (Figure 2c). The algorithm has also indicated the percentage of statistical certainty contributing to the sample spectrum from both components. This is a particularly challenging case because caffeine has many similar peaks to heroin, and heroin is present in a significantly lower concentration.

By comparison, HQI reports caffeine as the primary result, while failing to clearly identify the presence of heroin, the hazardous component. The p-value and HQI data for 80/20 caffeine/heroin mixture are summarized below in Table 2.

Table 2. HQI and p-value results for 80/20 caffeine/heroin (wt%) mixture analysis

Result	P-value result	HQI result
Heroin	PASS – heroin detected	0 – no chemical ID
Caffeine	PASS – caffeine detected	91.2 – only caffeine suggested
Mixture	PASS – both heroin and caffeine detected	0 – not identified

The p-value approach evaluates the full spectral profile and can identify statistically significant contributions from multiple components within a mixture. In this case, the individual components may not meet the required reporting threshold on their own, but when evaluated together the combined mixture spectrum is sufficient to explain the sample spectrum. This allows the correct combination of multiple components to be recognized, rather than reporting only the “dominant” material. This capability is particularly important for field users, as modern drug formulations often contain low concentrations of active narcotics mixed with cutting agents or other narcotics.

Thermo Fisher’s p-value approach provides a simple, clear and reliable basis for action. By distinguishing true matches from statistically inconsistent results, it reduces the likelihood of false positive alarms, strengthens confidence in on-scene decisions, and supports more effective public safety operations.

Conclusions

Thermo Scientific™ handheld analyzers with p-value algorithms provide field users reliable chemical identification for both pure chemicals and complex mixtures found in many street narcotics. This enables officers and safety personnel to work with greater confidence, appropriately escalate credible threats, and make faster, better-informed decisions in dynamic environments.

References

1. Kranenburg, Ruben F., et al. “Performance evaluation of handheld Raman spectroscopy for cocaine detection in forensic case samples.” Drug testing and analysis 13.5 (2021): 1054-1067.
2. Harruff, Richard C., William L. Barbour, and Nicole A. Yarid. “Evaluation of in-house drug evidence testing by King County Medical Examiner’s Office in “real-time” fatal overdose surveillance.” Journal of Forensic Sciences 69.4 (2024): 1350-1363.
3. Grover, Casey, Justine How, and Reb Close. “Fentanyl, etizolam, and beyond: A feasibility study of a community partnership using handheld Raman spectrometry to identify substances in the local illicit drug supply.” Journal of Public Health Research 12.2 (2023): 22799036231166313.

Learn more at **thermofisher.com/threatdetection**