Final Analysis Report - ISO/IEC 19795-2

Biometric Performance Testing

Prepared for

BRYK Group

By

BixeLab

3/16 Bentham Street, Yarralumla,
ACT 2600, Australia

BRYKGROUP
BRYK.ID

Technology Evaluation for:

BRYK.ID Extension v9.6.30 (Android v9.6.20, iOS v9.6.21) – System Under Test with three APIs integrated and working together:

- ***first: match-3d-2d-face-portrait***
- ***second: match-3d-2d-3rdparty-idphoto***
- ***last: match-3d-2d-3rdparty-idphoto-low-quality***

Attention: Matthew Smith
matthew.smith@brykgroup.com

Document Code	***23-018-TR-13***
Submit Date	***20 June 2023***
Submitted by	***BixeLab Pty Ltd***
Contact	***Ms. Somya Singh***
	M: 0412802334
	E: s.singh@bixelab.com

NVLAP®
TESTING
NVLAP LAB CODE 600301-0

Version History			
Version Number	**Description of change**	**Author**	**Date**
V1.0 (Interim)	Initial report release	S.Singh	19th May 2023
V1.1 (Interim)	Fixed graph for Sex distribution on page 17	S.Singh	22nd May 2023
V2.0	Final analysis report (draft)	S.Singh	16th June 2023
V2.1	Final analysis report	S.Singh	20th June 2023

Executive Summary

This report contains the results and findings of BixeLab's evaluation of BRYK's **BRYK.ID Extension version 9.6.30 – the system under test (SUT).** The primary goal was to establish if the biometric matching accuracy for the SUT is performing within TDIF identity binding role expectations for secure live deployment.

The BRYK.ID Extension version 9.6.30 was tested with the help of a test corpus collected using the BRYK.ID Extension (Early Access) app. The test crew comprised of 1,002 test subjects, who used the front-facing camera of their iOS or Android phones to submit a 3D reference (template), the back-facing camera to submit a high-quality photo of their photo on an identity document and an 2D image extracted by scanning the NFC chip on Australian ePassports[1] (see Test Constraints).

The test effort was conducted in full compliance with the protocol approved by Australia's NHMRC's Human Research Ethics Committee – Australian Institute of Health and Welfare (AIHW).

Testing included a full cross comparison test to verify a False Match Rate (FMR) of 0.01% or lower and False Non-Match Rate (FNMR) of 3% or lower. The principle operating point corresponding with the False Match Rate was "match level 5" and was tested so that there was at least 90% confidence interval that the FMR and FNMR were equal to or less than the required values.

To validate the FMR requirement of 0.01% or lower and FNMR or 3% or lower BRYK.ID Extension version 9.6.30 was tested to the applicable specifications of the ISO/IEC 19795-2 standard.

This report is intended for the key personnel in the Service New South Wales (SNSW) and BRYK Group and the results reported are applicable only to the testing undertaken and solution version tested.

Key Results

Verification (1:1) biometric matching performance for the SUT, with three APIs integrated and working together at threshold of match level 5 and 6, was measured using a demographically diverse analysis set with 1,005,007 (one million five thousand seven) comparison results – 1,002 mated comparison results, 1,004,005 non-mated comparison results.

The following metrics were measured as per TDIF guidance, with 90% confidence interval for each matcher within the SUT at the reported matching thresholds of 5 and 6.

Table 1 Biometric Matching Performance Summary: Using an analysis set with 1002 users.

Error Rate	Threshold	TDIF Biometric Requirement	match-3d-2d-face-portrait	match-3d-2d-3rdparty-idphoto	match-3d-2d-3rdparty-idphoto-low-quality
False Match Rate (FMR)	5	0.01%	0.041%	0.067%	0.067%

[1] For 3.3% users, a good quality image from the passport was used rather than NFC image.

Error Rate	Threshold	TDIF Biometric Requirement	match-3d-2d-face-portrait	match-3d-2d-3rdparty-idphoto	match-3d-2d-3rdparty-idphoto-low-quality
False Non-Match Rate (FNMR)		3%	1.71%	2.41%	2.11%
False Match Rate (FMR)	6	0.01%	0.005%	0.011%	-
False Non-Match Rate (FNMR)		3%	2.21%	5.62%	-

Refer to section *Performance Results* for a detailed overview of biometric performance measured.

To evaluate the effect of a small cohort of users who had lower quality (i.e., low lighting, bad pose, glasses with reflections) the performance of the three matching APIs was also measured on a reduced set of 959 users. This resulted in no significant change to the False Match Rate and approximately a 1% reduction in the False Non-Match Rate (see section 7.0).

Key Findings:

- At a matching threshold of 5 none of the engines tested meet the TDIF requirements for matching accuracy.
- At a matching threshold of 6, the FMR and FNMR for the best performing engine (match-3d-2d-face-portrait) are within the TDIF requirements for matching accuracy (see section 6.3 and 6.4)
- With the quality restricted set, the match-3d-2d-3rdparty-idphoto API was close to TDIF minimum accuracy requirements.
- No obvious correlation of demographic or environmental factors were discovered in leading to the high FMRs.

Table of Contents

1.0 Introduction

1.1. Background

BixeLab is a biometric testing laboratory that has obtained accreditation from the National Institute of Standards and Technology (NIST) under the National Voluntary Laboratory Accreditation Program (NVLAP Testing Lab Code 600301-0). This accreditation confirms that BixeLab meets the requirements outlined in the NIST Handbook 150-25, which align with ISO/IEC 17025:2017, the General Requirements for Competence of Testing and Calibration Laboratories.

BixeLab has been engaged by BRYK Pty Ltd to conduct additional biometric matching accuracy testing of the complete solution, using all three applicable matching APIs and a larger test crew size.

This document outlines the final technology test performance relating to the evaluation of the BRYK.ID Extension v9.6.30. More specifically, this report outlines the matching results of the system for the test corpus with 1002 unique, demographically diverse and consenting individuals. The intended readers of this test plan are key personnel from BRYK.ID and Service New South Wales (SNSW).

1.2. Objectives

The evaluation involved the following key components:

- **Biometric Matching** – *Is the person in the acquired image the same as the claimed identity document?*

This included testing the BRYK.ID Extension v9.6.30 for 3D to 2D biometric matching accuracy, which is listed as the System Under Test (SUT). This SUT is comprised of three APIs that work together to verify each identity transaction, including "3D:2D 3rd Party ID Photo Low Quality", "3D:2D 3rd Party ID Photo", and "3D:2D Face Portrait Matches".

This report focuses on the individual performance for each API in addition to the combined biometric performance.

The evaluation intended to result in two key outcomes:

1. 1:1 biometric matching performance report of the integrated solution, which includes a combination of three API engines, against the applicable TDIF biometric requirements, using a sample size of 1002 unique and consenting live subjects.
2. Provision of ISO 19795-2 compliant biometric performance test report.

1.3. Scope

BixeLab conducted a technology evaluation of BRYK.ID Extension v9.6.30, which included the following components:

1. Testing the biometric matching accuracy of the SUT with biometric matching data acquired from real human test subjects.
2. Reporting on the following performance metrics: False Reject Rate (FRR), False Accept Rate (FAR), and Failure to Acquire Rate (FTA) if applicable.
3. Reporting Bona fide Presentation Classification Error Rate (BPCER) for the SUT.

The evaluation does not include the following components:

1. Any components that are not explicitly stated in the evaluation scope.

1.4. Test Constraints

- The test subjects accessed the test application (SUT) remotely using their own mobile devices that complied to the pre-specified constraints for the acceptable software and hardware. Test running and data collection processes comprised largely of provision of written and video instructions related to test completion, consent, application installation, FAQs etc.
- The test environment was uncontrolled. This means that the test subjects completed the test transactions outside of the controlled laboratory conditions. Instructions were provided to ensure that the subjects were present in indoors, in a well-lit presentation medium. Where this provided a realistic overview of the performance in real-world, the varying influencing factors are not traceable.
- BixeLab sourced test subjects through a third-party. Demographic breakdowns were recorded but they were largely uncontrolled. All deviations from the planned demographics are declared in this report.
- BixeLab has tested what it believes to be a representative sample of a real-world system. Testing results are specific to the testing undertaken as in the Test Plan. Any deviations from the planned testing methodology are clearly identified in this document.
- Technology evaluation of the SUT was undertaken through a single integrated API issued by the client. The test execution process involved submission of a 2D image (enrolment image to compare against a 3D reference database made up of 3D FaceMaps in mated and non-mated comparison trials. The goal was to receive a comparison outcome and an associated match level where applicable. Where errors were outputted by the SUT, these were not specific enough for the evaluator to identify whether these were owing to a failure to enrol or failure to acquire, hence any comparison trials for which the SUT outputted an error these were measured as FTAs.
- The outcomes of the test report was specific to the performance of the SUT as measured in this specific evaluation. BixeLab is not liable for any claims made for the SUT that fall outside the scope of this evaluation.
- On April 26th, the client was informed that test subjects were having difficulty with the NFC scan process due to unclear on-screen feedback between taking a photo of the ID document and the NFC chip reading process step. The client then added on-screen user instructions for the NFC flow to the Android application version. It was confirmed that no other changes were made to the system under test (SUT). This was verified by checking the application version at the bottom of the screen, which remained the same. BixeLab trusts the SUT provider and has tested what it believes to be the agreed version of the SUT. The SUT

version has remained consistent throughout the evaluation process, as confirmed by the client.

- Analysis of non-mated comparison trials revealed instances of unusual outcomes related to false acceptance. Upon manual review of these high scoring non-mated comparison trials, we are investigating potential causes for this based on test metadata.
- A subset of users encountered difficulties in utilizing the NFC scanning feature within the application, however, they were still able to successfully complete the selfie and Document Scanning procedures. In order to accommodate these users, BixeLab contacted them individually to request high-quality 2D passport photos. Out of the total of 1,002 distinct individuals, 3.3% had submitted their 2D images separately. Whenever necessary, these images were enhanced to maximize their resemblance to images acquired through NFC scanning.
- Following a comprehensive manual review of the false matches and false no-matches, a certain percentage of users were excluded from the analysis set owing to their non-compliance with the SUT 3D selfie requirements. Non-compliant instances encompassed users who captured selfies wearing glasses that exhibited visible reflections, users who submitted 3D selfies taken in inadequately lit environments, users who provided selfies with unfavourable pose angles, or users whose 2D images did not meet the relevant quality standards for evaluation. The reduced analysis set was analysed to derive performance metrics relating to the quality restricted set.

2.0 Scenario Description

2.1. System Information

This section provides information related to the System Under Test (SUT) - listed as BRYK.ID Extension v9.6.30 for biometric matching accuracy (Technology) evaluation. The specifications related to the SUT are listed in Table 2 below.

Table 2 System Information

System Specifications BRYK.ID Extension v9.6.30		
BRYK.ID Extension v9.6.30 with three APIs integrated and working together: • highMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-face-portrait • moderateMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-3rdparty-idphoto • lowMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-3rdparty-idphoto-low-quality	System Version	*Facetec Server SDK 9.6.30* *Android: v9.6.20* *iOS: v9.6.21*
	Modality	*Face*

2.2. Concept of Operations

This section describes the characteristics of the SUT from the viewpoint of an individual who will use that system.

The use case for the SUT involves 3D mapping a selfie video of a person's face in real time to produce a 2D reference (template) which is then compared to a passport style image acquired through a document scan. The video is captured by the SUT when each user completes the liveness challenge of moving closer to the device. This video is then processed to produce the 2D reference (template) which is saved to BRYK's server. Each user also scans an ID document through NFC to complete the application process flow. The end-to-end SUT process flow is illustrated in Figure 1 below.

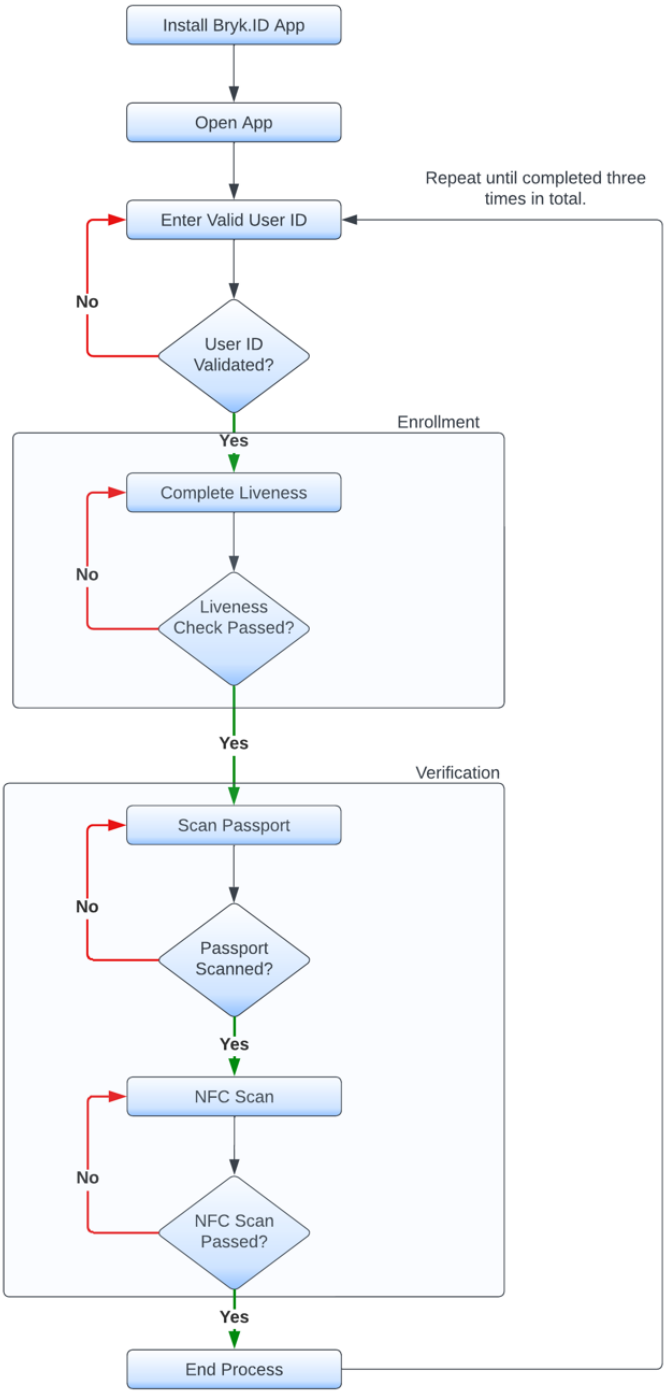


Figure 1 Overview of the application process flow for system under test (SUT)

Figure 2 provides an overview of the testing process flow based on data collected using the SUT. This process flow was used to undertake an offline evaluation to measure the FRR, FAR, FTA and BPCER associated with all three underlying biometric matchers.

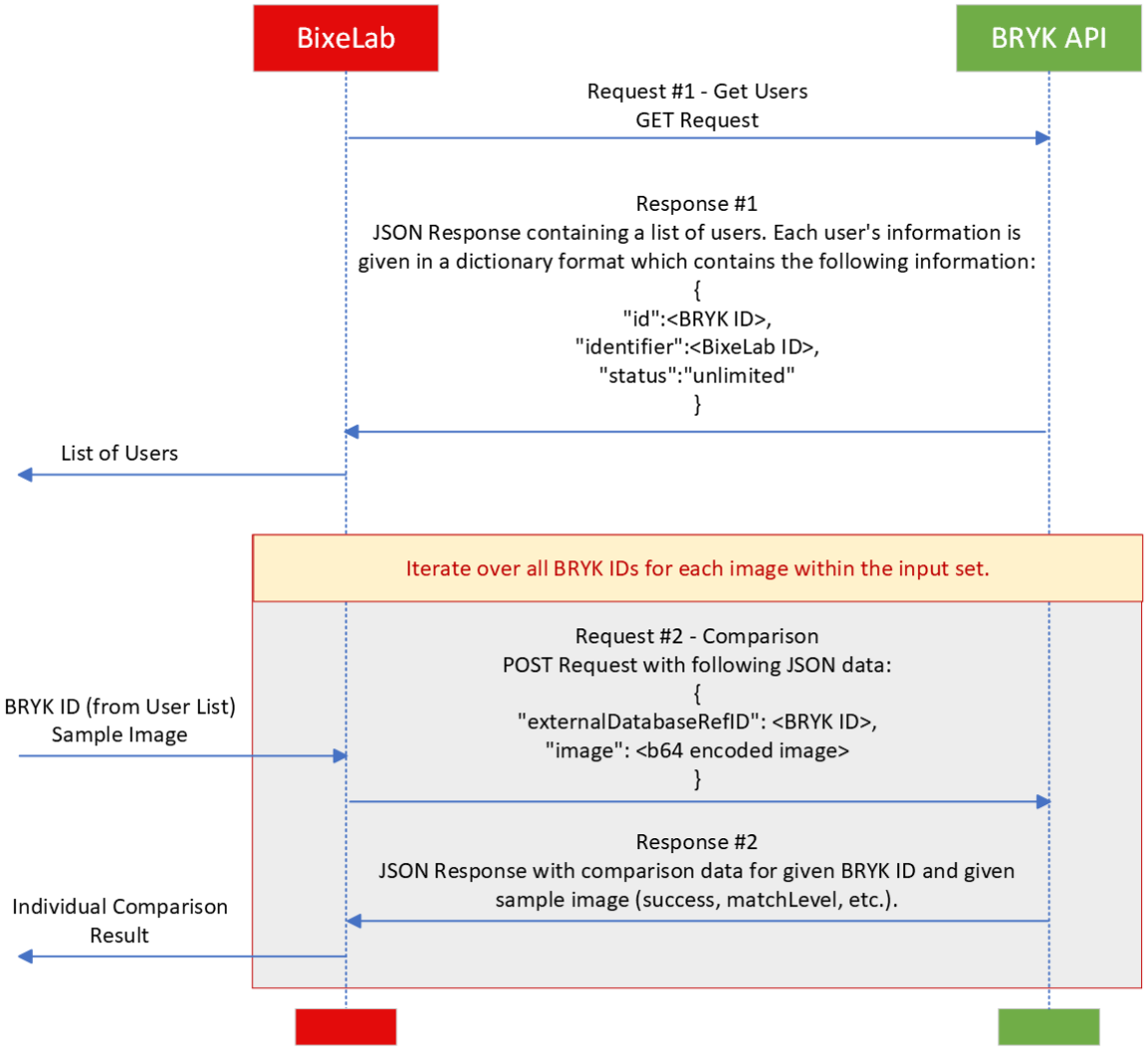


Figure 2 Process flow for system under test (technology evaluation)

Note that it is the responsibility of the owner of the technology provided for testing to ensure that the settings align with the objective of the evaluation and no configurations associated with the SUT are changed during the evaluation. Before the testing process is initiated, BixeLab ensured that the software versions are installed and configured appropriately and verify that the system is operating correctly.

A summary of pre-testing activities undertaken to setup the SUT for the evaluation is outlined in this report.

2.3. Test Environment

The test apps were accessed remotely by all test subjects with concise written and video instructions. The test subjects were encouraged to test the app in good quality lighting; however, the physical test environment was largely uncontrolled. Since the app will be used similarly in a real-world environment, the uncontrolled test environment was justified.

Additionally, the app was available on both iOS and Android. Out of the 1002 users, 56.3% users used the app on iOS and 43.7% used the app on Android. Since the app will be similarly used in a real-world environment, the uncontrolled software and hardware test environment was justified.

2.4. Configuration Audit Log

The nature of the testing meant that it was not possible to do a software and hardware audit log. An alternate approach was taken instead. An audit of the system configuration of the SUT was conducted during pretesting. The information recorded is outlined in a log that includes hardware and software specifications, OS versions, serial number, and build at the outset of testing. During testing, the app version, and the dashboard was audited daily.

During testing, BixeLab requested BRYK to make changes to the user interface for the android app to improve usability. This was noted in the audit log.

This log is to verify that client software and hardware configurations related to testing do not unexpectedly change through the execution. The aim is to mitigate system software and hardware related variations and to perform essential laboratory protocol for integrity. The table in *Appendix 4: Configuration Log* provides details of the software and hardware specifications with no identified anomalies.

2.5. Output Measures

The output measures define the outcomes that were measured during the evaluation.

The System Under Test (SUT) – listed as BRYK.ID Extension v9.6.30 produces the output measures listed below. Outputs that were considered for analysis have been identified in bold formatting.

Table 3 4System Under Test

System Under Test component	Interpreted components	Output Measures (Dashboard)
Biometric Matching	**3D Enrolments**	**Date/Time** Geolocation **3D Liveness** **2D Reference** Estimated Age **SDK Ver. (9.4.16 for android/9.4.22 for iOS) including Identity of the Tester**

System Under Test component	Interpreted components	Output Measures (Dashboard)
		Device Model Device platform Fraud Flags On Fraud List Details
	NFC extracted 2D verification images	**Date/Time** Geolocation **3D Liveness** **SDK Ver. (9.4.16 for android/9.4.22 for iOS) including Identity of the Tester** Device Model Device platform Face Scan Audit Trail Doc. Scan Face Crop **NFC Image** **Doc. Front Scan & Crop** Doc. Back Scan & Crop NFC Data & NFC Cert. Match Level (ID Photo) Match Level (NFC) Doc. Type
	3D to 2D matching	**Success (true/false)** **Matched (true/false)** **Match level (>=5 corresponds to a match, <5 corresponds to a no match)**

2.5.1. Biometric Matching Subcomponents

The SUT has a minimum Match Level = 5 in all 3 endpoints. Match levels greater than or equal to 5 result in a Match and match levels lower than 5 result in a non-match. The table below provides the expected output performance for biometric matching subcomponents:

Table 4 5Expected Outputs for Biometric Matching Subcomponents

System Under Test component	SUT Subcomponent	Ground Truth	Output Measures (Dashboard)	
Biometric Matching	All three APIs	Genuine (Mated)	Success	True
		Impostor (Non-mated)		False
			Message	-
		(FTA)		“There was unexpected error,

System Under Test component	SUT Subcomponent	Ground Truth	Output Measures (Dashboard)	
				please contact BRYK.ID”
	highMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-face-portrait	(FTA)	Match Level	0
		Impostor (Non-Mated)		1
		Impostor (Non-Mated)		2
		Impostor (Non-Mated)		3
		Impostor (Non-Mated)		4
		Genuine (Mated)		5
		Genuine (Mated)		6
		Genuine (Mated)		7
		Genuine (Mated)		8
		Genuine (Mated)		9
	moderateMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-3rdparty-idphoto	(FTA)	Match Level	0
		Impostor (Non-Mated)		1
		Impostor (Non-Mated)		2
		Impostor (Non-Mated)		3
		Impostor (Non-Mated)		4
		Genuine (Mated)		5
		Genuine (Mated)		6
		Genuine (Mated)		7
	lowMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-3rdparty-idphoto-low-quality	(FTA)	Match Level	0
		Impostor (Non-Mated)		1
		Impostor (Non-Mated)		2
		Impostor (Non-Mated)		3
		Impostor (Non-Mated)		4
		Genuine (Mated)		5

3.0 Dataset Description

3.1. Dataset Profile

This section aims to specify the high-level attributes of the test data associated with this evaluation based on the requirements established in the test plan provided previously. These attributes include but are not limited to age, gender, ethnicity, and other factors under which the corpus was collected.

For the final analysis, a test crew of 1002 unique and conformant test subjects was identified following data collection for running matching accuracy tests offline. All subjects were required to read and sign consent to participate in this testing prior to commencing testing activities.

The table below provides an overview of the demographic attributes of the test crew.

Table 5 Demographic Inclusions for Test Crew

Subject Metadata / Input Factors	Metadata breakdown	ISO 19795-5 Recommended Distributions	BixeLab Achieved Test Crew Inclusion
Age Range	**< 18 y/o**	0%	0%
	18 – 30 y/o	25-40%	20.4%
	31 – 50 y/o	25-40%	64.7%
	51 – 70 y/o	25-40%	14.4%
	> 70 y/o	0%	0.5%
Gender	**Male**	40-60%	45%
	Female	40-60%	54.5%
	Non-Binary	-	0.3%
	Gender Fluid	-	0.2%
Ethnic Origin[2]	**Caucasian**	-	59.4%
	Southern and Central Asian	-	16.5%
	Undisclosed	-	6.9%
	Oceanian (AU First Nations, Maori, etc)	-	5.4%
	Northeast Asian	-	5.3%
	North African and Middle Eastern	-	1.8%
	Southeast Asian	-	1.5%
	Mixed Asian	-	0.8%
	South American	-	0.8%

[2] Adapted from Australian Standard Classification of Cultural and Ethnic Groups (2019) available at https://www.abs.gov.au/ausstats/abs@.nsf/mf/1249.0

Subject Metadata / Input Factors	Metadata breakdown	ISO 19795-5 Recommended Distributions	BixeLab Achieved Test Crew Inclusion
	Sub-Saharan African	-	0.5%
	Latino	-	0.4%
	Polynesian	-	0.3%
	Mixed Race (undisclosed)	-	0.2%
	Asian	-	0.2%

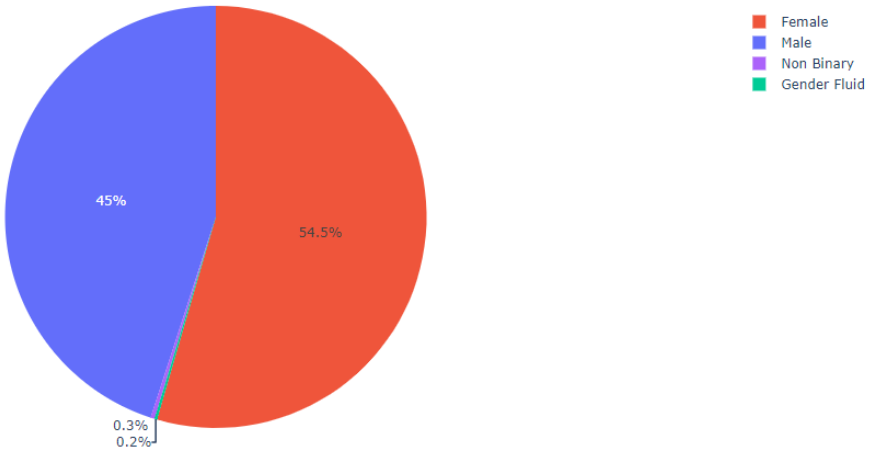


Figure 3 Gender Distribution

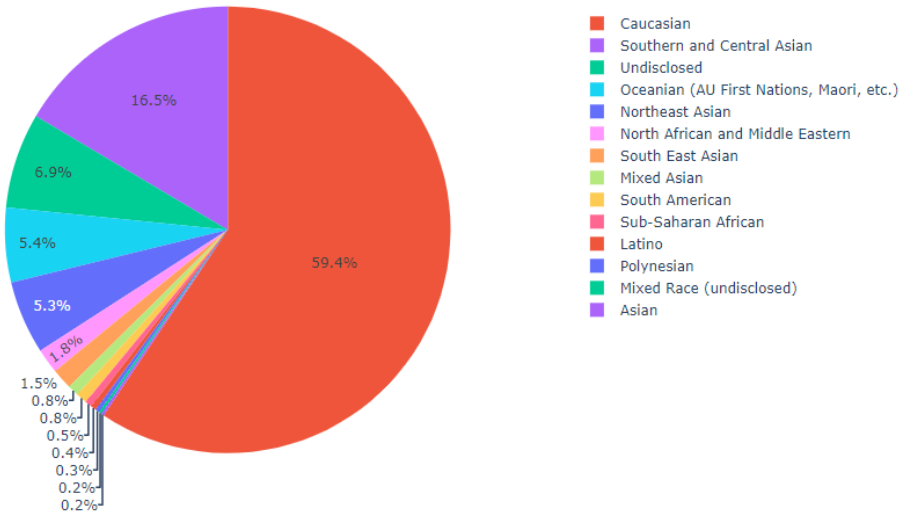


Figure 4 Ethnic group Distribution

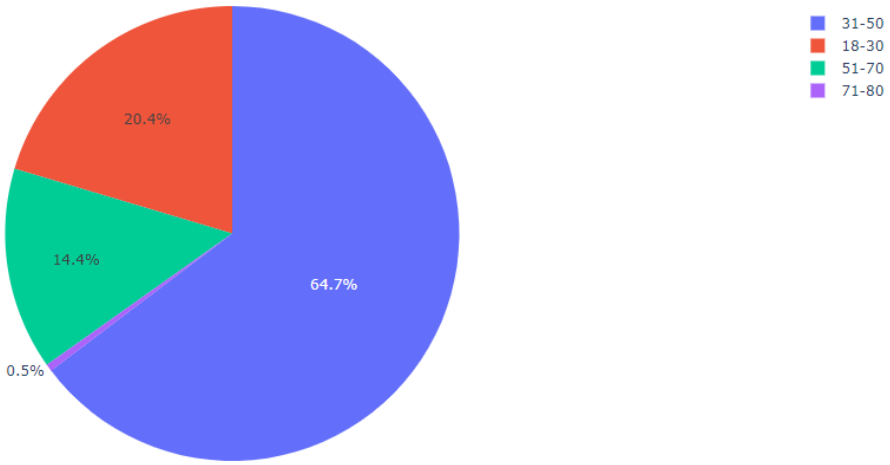


Figure 5 Age group Distribution

A brief overview of the databases used for biometric mated and non-mated comparison trials is provided below:

- Biometric reference database comprised of 2D source images (face image that is digitally stored on a passport) extracted with the help of NFC chip reading at the point of capture during test completion and stored within the test environment.
- 3D Facemaps for each test subject were created in the SUT database and these formed the biometric references to be used as the subject of biometric comparison later. Due to design constraints in the SUT, the 3D Facemaps were not directly reviewable. However, audit images (*auditTrailBase64*) were made available on the dashboard, which provided indications as to whether a given 3D FaceMap would be valid or not for the purpose of running a full crossmatch at a later stage.

4.0 Data Collection

This section provides a detailed overview of the data collection process. BixeLab leveraged a trusted and reliable test crew recruitment agency to recruit consenting test subjects. The data collection process was divided into the following key phases:

1. **Pre-screening**
2. **Application installation & Test completion**
3. **Post test completion processes**

4.1. Pre-screening

Owing to the remote nature of this evaluation, it was essential to ensure that each participant had access to suitable hardware, software, biometrically readable passport and reliable internet connection. Within this context, each participant was requested to ensure that they met the following criteria as part of pre-screening process:

Table 6 6 Pre-Screening Process

Checklist Item	Description
1	Participant has an Australian Biometric Passport (Australian Passport with an NFC-readable chip) The passport must: • Be an Australian Passport issued after 2014. • Bear the biometric logo as highlighted in red in the image below. 
2	Participant has access to a phone that meets the following requirements: 1. **For iPhones**, the model must be an iPhone that came out after the iPhone X (this includes iPhone XR, XS, 11 etc.). The version of iOS must also be 15.6 or above 2. **For Androids**, the version of Android must be 12 or above

Checklist Item	Description
3	Participant must perform the steps below to confirm the requirements mentioned in points 1 and 2: Check NFC functionality: • To check NFC functionality, install the following ReadID Me app, from the links provided. Android: https://play.google.com/store/apps/details?id=nl.innovalor.nfciddocshowcase&hl=en&gl=US iOS: https://apps.apple.com/us/app/readid-me/id1463949991 • Use the app to scan your Biometric Passports. • The scan will be completed only if your phone has NFC ability AND your passport has a Biometric Chip (NFC chip). If ReadID Me doesn't show up on the App store/Play store, this means your device does not have NFC and therefore you will be unable to participate in this testing.
4	Access to a reliable internet connection.
5	Any environment with good, even, and ambient lighting.

4.2. Application Installation & Test Completion

Refer to *Appendix 5: Test Application Installation Instructions* and ***Error! Reference source not found.*** for more information related to instructions provided to the test subjects. These sections also provide an overview of the test running process.

4.3. Post test completion processes

After completing the test, the test operators carried out post-test processes, which included auditing the dashboard to verify that all users had completed the test accurately. In cases where users were unable to complete the test due to issues with the user interface, difficulty in interpreting instructions, or any other related challenges, they were prompted to retry. If the users were still unsuccessful after attempting to retry, they were then requested to withdraw from testing. All biometric data related to the tests completed, was audited, validated, and logged for later use in analysis.

5.0 Test Execution

5.1. Pretesting

This section provides a summary of the pretesting processes undertaken within the controlled lab facilities prior to test commencement. The main goal was to ensure that the SUT application was operating as expected on both types of mobile operating systems (iOS and Android) in addition to ensuring that the test running procedures planned for data collection and remote testing were fit for purpose.

Key observations related to the pretesting process are noted below:

- Both versions were put through different tests to simulate the environmental conditions that remote users may be in. One was of extreme lighting to simulate a user being in a brightly lit room and the other was of dim lighting to simulate a user being in a darker room. Both versions of the app were unbothered by the different lighting conditions and were able to pick up and pass the liveness checks.
- Performance of the app regarding the images taken was consistent and accurate.
- The flow of the app was easy to follow.
- iOS application had a better user interface in comparison to Android assisting the users on each step with the help of on-screen instructions.
- The processing/loading screens seemed to take longer for iOS as compared to Android.

5.2. Testing procedures

This section provides an overview of the testing process to receive the biometric performance statistics for analysis. This involved running a full crossmatch of 1002 3D Facemap templates (references) against a set of 1002 2D image samples.

5.2.1. Technology evaluation workflow

At the outset of each evaluation, the SUT's hardware and/or software specifications are logged as part of BixeLab's audit tracking requirements.

The BRYK app was used remotely by test subjects to perform the first 6 steps of the workflow.

1. User entered the provided username.
2. User completed liveness check, then 3D Facemap is saved in the database.
3. User completed ID Scan (AU Passport from 2014: P and R Series)
4. The front-page ID scan is run against the endpoint:match3d-2d-id-scan. A successful match, corresponding to a Match level >=5 let the user proceed to the next step.
5. User completed NFC Scan.
6. The SUT saves the Extracted NFC Digital Photo in the database.

Once the users had submitted their enrolment and verification images remotely, the rest of the evaluation was carried out on site at BixeLab in offline testing. A brief overview of the offline testing undertaken to complete a technology style evaluation of the SUT with all three matchers is provided below:

1. Download the 2D sample images from BRYK using the GET USERS endpoint (https://api.brykid.brykgroup.ai/users/find/{user_id}') and collate them into a dataset.
2. Then the technology evaluation script is run using the Comparison endpoint (https://api.brykid.brykgroup.ai/match-3d-2d).
   The script does the following:
3. Read in all of the 2D sample images and convert them to a base64 format to be readily available when sending the data over BRYK's web API.
4. Request the equivalent BRYK ID for each sample.
5. Then, iterate over each 2D image and BRYK ID, calling BRYK's comparison endpoint POST '/match-3d-2d' until all comparison options are exhausted.
6. The SUT runs the extracted NFC Digital Photo against the 3 endpoints: 3D:2D Face Portrait | 3D:2D 3rd Party ID Photo | 3D:2D 3rd Party ID Photo Low Quality outputting matching scores for each endpoint.
7. After completion of all comparisons, these results are then exported to Excel document.
8. The match scores are used to develop performance statistics for analysis.

6.0 Performance Results

This section provides the performance results associated with the SUT for technology evaluation. The reported results are in line with the applicable specifications of ISO/IEC 19795-2 (clause 6.3).

6.1. Failure-to-enrol rate (FTER)

The failure-to-enrol rate represents the proportion of enrolment transactions where the system fails to create and store a biometric reference in accordance with the enrolment policy. This rate includes three scenarios:

- Individuals who are unable to present the required biometric characteristic.
- Individuals who cannot produce a sample of sufficient quality during enrolment.
- Individuals who are unable to reliably produce a match decision with their newly created reference when attempting to confirm the enrolment's usability.

In the technology evaluation conducted, an analysis was performed based on a previously collected corpus, eliminating issues in obtaining a biometric sample. However, enrolment failures can still occur, such as when the biometric sample quality is too low for feature extraction. During the evaluation of 1002 unique and consenting individuals, no Failures to Enrol were observed.

6.2. Failure-to-acquire rate (FTAR)

The proportion of verification attempts for which the system fails to capture or locate an image or signal of sufficient quality corresponds to a Failure to Acquire rate. The failure-to-acquire rate shall be estimated as the proportion of test attempts (for mated comparison trials) that were not able to be completed due to failures at presentation (no image captured), segmentation, feature extraction or quality control.

BixeLab used a previously collected test corpus ruling out capture failures. Further, feature extraction was successful for all samples. Hence, no failures to acquire were observed. This led to a 0% failure-to-acquire rate.

6.3. False Match Rate

The false match rate is the proportion of a specified set of completed non-mated comparison trials that result in a comparison decision of "TRUE".

The table below provides the measured False Match Rates for each individual matcher within the SUT.

Table 7 7 False Match Rates for each matcher at threshold 5 and 6

Error Rate	False Match Rate (FMR) @ threshold of 5	False Match Rate (FMR) @ threshold of 6
match-3d-2d-face-portrait	0.041%	0.005%
match-3d-2d-3rdparty-idphoto	0.067%	0.011%
match-3d-2d-3rdparty-idphoto-low-quality	0.067%	-

6.4. False Non-Match Rate

ISO/IEC 19795-1 defines FNMR as the proportion of completed mated comparison trials that result in a comparison decision of “FALSE”.

The table below provides the measured False Non-Match Rates for each individual matcher within the SUT.

Table 88 9 False Non-Match Rates for each matcher at threshold 5 and 6

Error Rate	False Non-Match Rate (FNMR) @ threshold of 5	False Non-Match Rate (FNMR) @ threshold of 6
match-3d-2d-face-portrait	1.71%	2.21%
match-3d-2d-3rdparty-idphoto	2.41%	5.62%
match-3d-2d-3rdparty-idphoto-low-quality	2.11%	-

Figures below provide the distribution of comparison scores for genuine (mated) and imposter (non-mated) comparisons for each individual matcher.

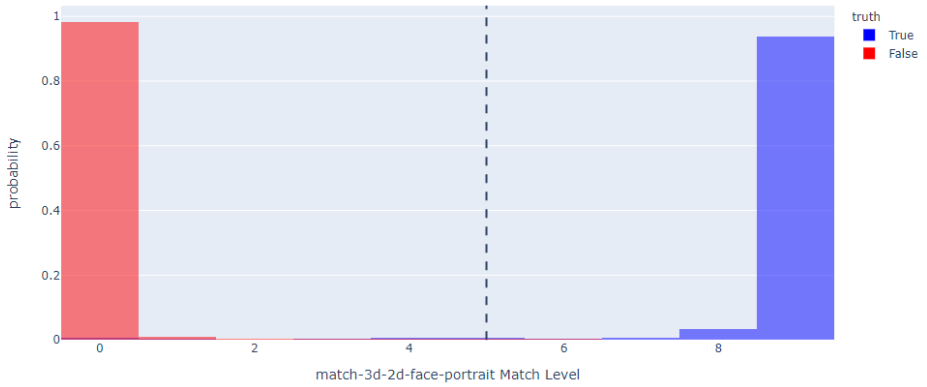


Figure 6 Histogram showing distribution of mated (blue) and non-mated (red) similarity scores for 3D-to-2D Face Portrait Matcher

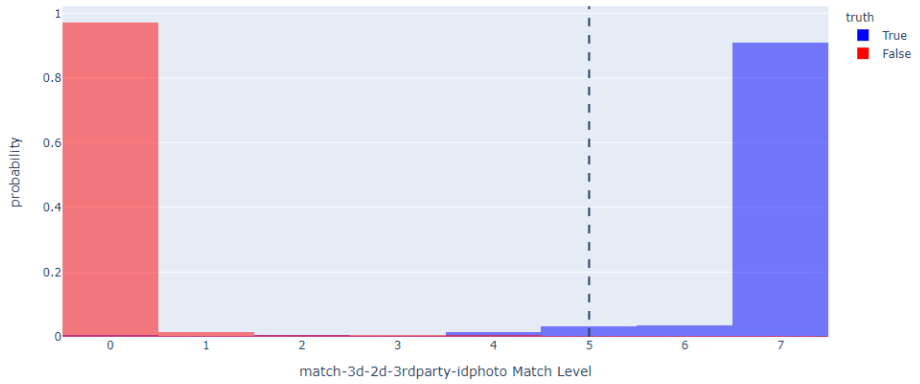


Figure 7 Histogram showing distribution of mated (blue) and non-mated (red) similarity scores for 3D-to-2D 3rd Party ID Photo Matcher

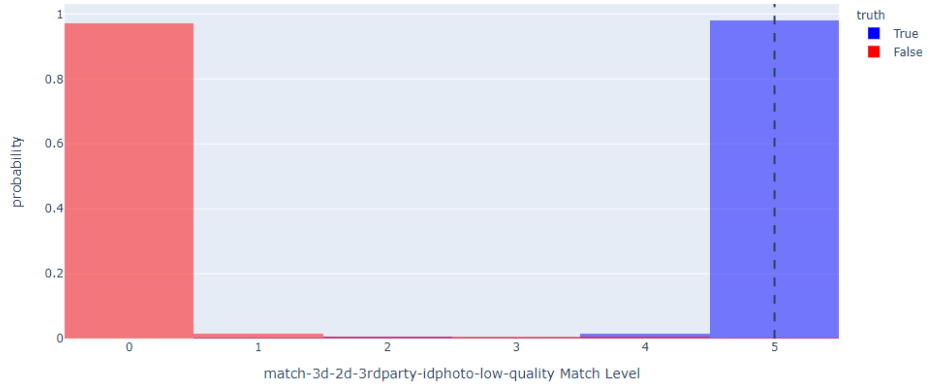


Figure 8 Histogram showing distribution of mated (blue) and non-mated (red) similarity scores for 3D-to-2D 3rd Party ID Photo Low Quality Matcher

7.0 Performance Results: Revised Analysis Set

A thorough review of false accepts and rejects across the three APIs was undertaken by reviewing the audit trail images against each user's 3D selfie reference from the dashboard and their 2D images. Based on the manual review, a small percentage of users were excluded from the analysis set due to quality related factors – "The environment that the ToE is intended to be used in any environment with good, even, and ambient lighting".

Instances of quality factors included users who took a selfie while wearing glasses with visible reflections, users who submitted a 3D selfie taken in a poorly lit environment, users who submitted a selfie with an unfavourable pose angle, or users whose 2D image did not meet the applicable quality standards for the evaluation.

After 43 non-compliant users were dropped, the analysis set contained 959 mated (genuine) comparisons and 957252 non-mated (imposter) comparisons. Based on the analysis of these comparisons, the following performance metrics were derived:

Table 9 Performance Metrics based on revised analysis set: at threshold 5

Error Rate @ threshold 5	match-3d-2d-face-portrait	match-3d-2d-3rdparty-idphoto	match-3d-2d-3rdparty-idphoto-low-quality
False Match Rate (FMR)	0.040%	0.065%	0.065%
False Non-Match Rate (FNMR)	0.83%	1.04%	1.043%

Table 10 10 11Performance Metrics based on revised analysis set: at threshold 6

Error Rate @ threshold 6	**match-3d-2d-face-portrait**	**match-3d-2d-3rdparty-idphoto**	**match-3d-2d-3rdparty-idphoto-low-quality**
False Match Rate (FMR)	0.005%	0.011%	-
False Non-Match Rate (FNMR)	1.15%	3.33%	-

8.0 Usability Performance

BixeLab provided all test subjects with a survey form to seek information relating to the number of transactions they performed, average number of re-attempts it took to complete each step in the application process flow, in addition to their general thoughts and comments on the application. 905 subjects returned surveys with a mixture of failures and successes as well as significant feedback for analysis.

Below is a table providing average number of re-attempts for users for each step in the application process flow:

Table 11 12 13Average number of re-attempts per stage in application process flow

Stage	Average number of re-attempts
Average Liveness	1
Average Document scan	2
Average NFC scan	2

9.0 Findings

BixeLab has completed an analysis of technology evaluation results for BRYK.ID Extension v9.6.30 (Android v9.6.20, iOS v9.6.21) – System Under Test (SUT). The target of Evaluation for this testing were three biometric matchers forming part of the SUT:

- match-3d-2d-face-portrait
- match-3d-2d-3rdparty-idphoto
- match-3d-2d-3rdparty-idphoto-low-quality

The purpose of this report is to provide the measured performance rates based on testing completed thus far. Refer to the section *Performance Results* of this report.

Constraints

BixeLab has evaluated what it believes to be a representative sample of the commercially available solution with the utilisation of appropriate testing methodology stemming from the specifications of ISO/IEC 19795-2 standards.

The results associated with the technology evaluation of the SUT have been reported in this document.

Refer to section *Test Constraints* for limitations associated with this evaluation.

Deviations

BixeLab has taken every measure to ensure that no deviations were made from the specifications of the standards.

Conclusions

There are no other comments or thoughts from BixeLab that are not addressed in this report. BixeLab is not liable for any claims made outside of the results contained in this report.

10.0 Appendix 1: Receiver Operating Characteristic (ROC) Curves

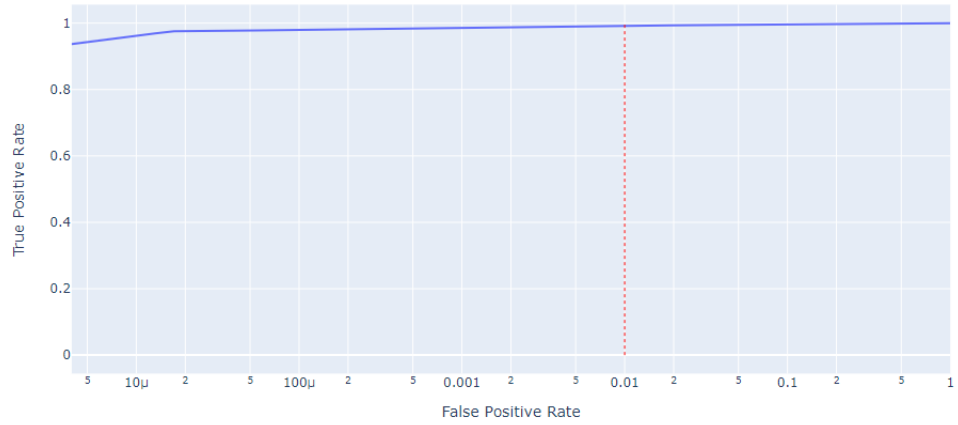


Figure 9 ROC for 3D-to-2D Face Portrait matcher

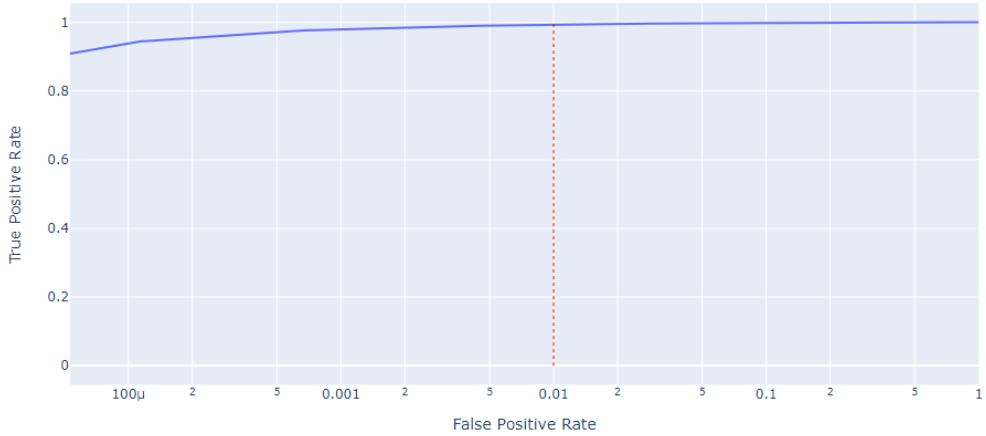


Figure 10 ROC for 3D-to-2D 3rd Party ID Photo matcher

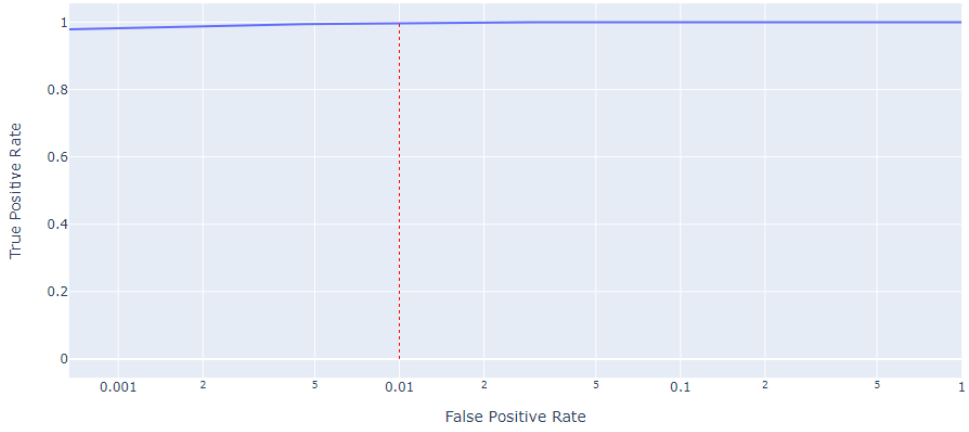


Figure 11 ROC for 3D-to-2D 3rd Party ID Photo Low Quality matcher

11.0 Appendix 2: Verification Metrics

Technology Evaluations performed to the ISO/IEC 19795-1 & ISO/IEC 19795-2 standards are classifiable into three categories – verification, closed set identification, or open set identification.

Following metrics have been reported in this test report:

$$\text{FMR} = \frac{\text{Number of imposter transactions for which decision is accept}}{\text{Number of imposter transactions}}$$

$$\text{FNMR} = \frac{\text{Number of genuine transactions for which decision is reject}}{\text{Number of genuine transactions}}$$

$$\text{FTAR} = \frac{\text{Number of attempts that failed to be processed}}{\text{Total number of acquisition attempts}}$$

$$\text{FTER} = \frac{\text{Number of test subjects for whom the enrolment was rejected (failed)}}{\text{Total number of enrollment attempts made}}$$

12.0 Appendix 3: Terms and Definitions

Term	Abbreviation	Definitions
Technology Evaluation		
biometric candidate		biometric reference identifier of a biometric reference in the biometric reference database determined to be sufficiently similar to the biometric probe to warrant further analysis
biometric probe		biometric sample or biometric feature set input to an algorithm for biometric comparison to a biometric reference(s)
comparison score		numerical value (or set of values) resulting from a comparison
mated		of or having to do with a paired biometric probe and biometric reference that are from the same biometric characteristic of the same biometric data subject
non-mated		of or having to do with a paired biometric probe and biometric reference that are not from the same biometric characteristic of the same biometric data subject
authentication		the act of proving or showing to be of undisputed origin or veracity
biometric mated comparison trial		Comparison of a biometric probe and a biometric reference from the same biometric capture subject and the same biometric characteristic as part of a performance test
biometric non-mated comparison trial		comparison of a biometric probe and a biometric reference from different biometric data subjects as part of a performance test
failure to acquire	FTA	failure to accept for subsequent comparison the output of a biometric capture process, a biometric sample of the biometric characteristic of interest
failure to enrol	FTE	failure to create and store a biometric enrolment data record for an eligible biometric capture subject, in accordance with a biometric enrolment policy
false match rate	FMR	proportion of the completed biometric non-mated comparison trials that result in a false match
false non-match rate	FNMR	proportion of the completed biometric mated comparison trials that result in a false non-match
Subject		The person from whom the biometric enrolment was taken. The target of the attack
Target of evaluation	TOE	Within Common Criteria, the IT product that is the subject of the evaluation. Note: The TOE in Common Criteria evaluations is the equivalent of IUT in biometric evaluations.
test approach		Totality of considerations and factors involved in technology evaluation

13.0 Appendix 4: Configuration Log

The Client application software specifications are noted below						
Provider	**Title**	**Client CODE (If Applicable)**	**Version**	**Build**	**Identified Anomalies**	**Biometric Modality**
BRYK	BRYK.ID Demo app Extension version (Android)	BXL018	9.6.30	9.6.20	None	Face
BRYK	BRYK.ID Demo app Extension version (iOS)	BXL018	9.6.30	9.6.21	None	Face
Specifications regarding the system under test are noted below						
Provider	**Title**	**Client CODE (If Applicable)**	**Version**	**API End Point Access**	**Build**	**Biometric Modality**
BRYK	BRYK.ID Extension v9.6.30 with three APIs integrated and working together: • highMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-face-portrait • moderateMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-3rdparty-idphoto • lowMatchLevel: https://dev.facetec.com/api-guide#match-3d-2d-3rdparty-idphoto-low-quality	BXL018	9.6.30	The 'Get Users' endpoint provides the list of enrolled users (i.e., users whose reference (template) has been created – https://api.brykid.brykgroup.ai/users/export The 'Comparison' endpoint provides with comparison results give a user ID and base64 encoded image – https://api.brykid.brykgroup.ai/users/export	N/A	Face
The BixeLab software and hardware system specifications are noted below						
Manufacturer & Model Name	**Software**	**BXL CODE (If Applicable)**	**Operating System**	**Identified Anomalies**	**Firmware**	

Custom built computer (No Manufacturer)	-	BXL138 and Performix version 6.5.1	Windows 11	Based on the review of software configuration log, no factors that could affect testing were identified	N/A
Custom built computer (No Manufacturer)	-	BXL 134	Ubuntu 20.04	Based on the review of software configuration log, no factors that could affect testing were identified	N/A
Custom built computer (No Manufacturer)	-	BXL 141	Windows 10 Pro	Based on the review of software configuration log, no factors that could affect testing were identified	N/A

14.0 Appendix 5: Test Application Installation Instructions

Please use these step-by-step instructions below to install the application you will test today. You may also refer to the instructional video provided with this document.

14.1. Android

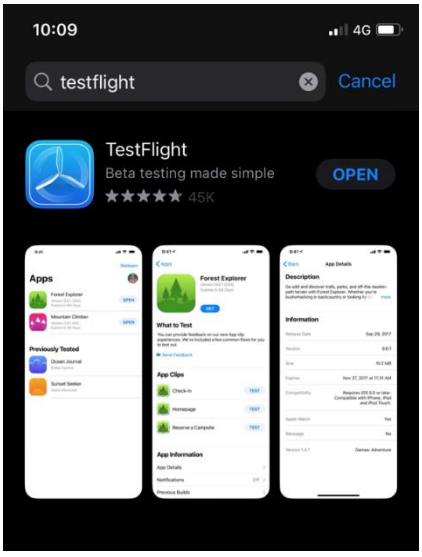
To install the app, follow these steps:

1. Please click the following link which will direct you to the page of the app on the Google Play Store: https://play.google.com/store/apps/details?id=com.brykgroup.brykid.extension
2. Select “Install” and once downloaded press “Open”.
3. You will be provided with your User ID in the email sent by *AGENCY* where you found these instruction documents.

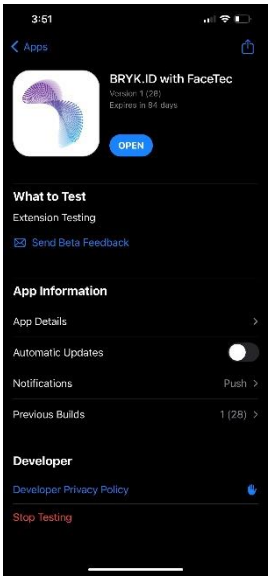
14.2. iPhone

To install the application, follow these steps:

1. On your iPhone, open the App Store.
2. On the App Store, using the search function (bottom right corner), search for “TestFlight” and press the “search” button.
3. You should see an application that looks like this.



4. Click on the “Get” button next to the application name. Wait for the application to install.
5. You should now see an “Open” button next to the name of the app (as seen in the image above). Click on this “Open” button to start the app.
6. Once you open this app, you should see a pop up for Notifications. Click “Allow” to proceed to the next step.
7. On the “Welcome to TestFlight” page, click on continue.
8. You should arrive at the “Ready to Test” page, with a “Redeem” button.
9. Once you get to this page, click on the following link from your iPhone: https://testflight.apple.com/join/E444PZj1. This should redirect you back to the TestFlight application.



10. You should now see the BRYK.ID application on your screen.
11. Click on the “Accept” button.
12. Click on the “Install” button. The app will now begin installing on your iPhone.
13. Click the “Open” button to open the app where you will be prompted to enter your User ID. You will be provided with your User ID in the email sent by *AGENCY* where you found these instruction documents.

15.0 Appendix 6: Test Application User Instructions

Follow these instructions when using the app for testing.

You may also refer to the instructional video provided with this document.

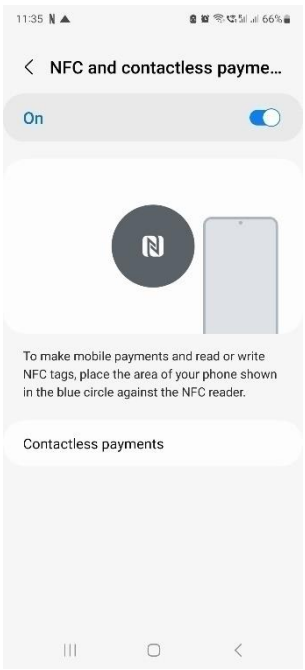
The differences between Android and iPhone will be shown within the screenshots provided after the steps.

DISCLAIMER FOR ANDROID USERS

Before initiating testing please ensure that you enable NFC if you are using an Android device.

Android:

1. Go to the Settings app on your Android phone.
2. Scroll down and look for "Wireless & networks" or "Connections."
3. Tap on it and look for "NFC" or "Near Field Communication."
4. Turn it on



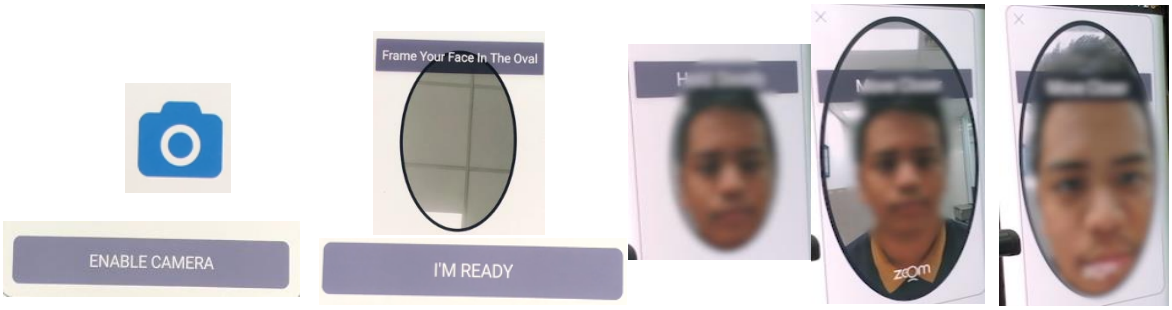
To complete the testing process, you are **REQUIRED** to perform the following procedure a total of 3 times.

1. Once the installation is complete, open the BYRK.ID app.

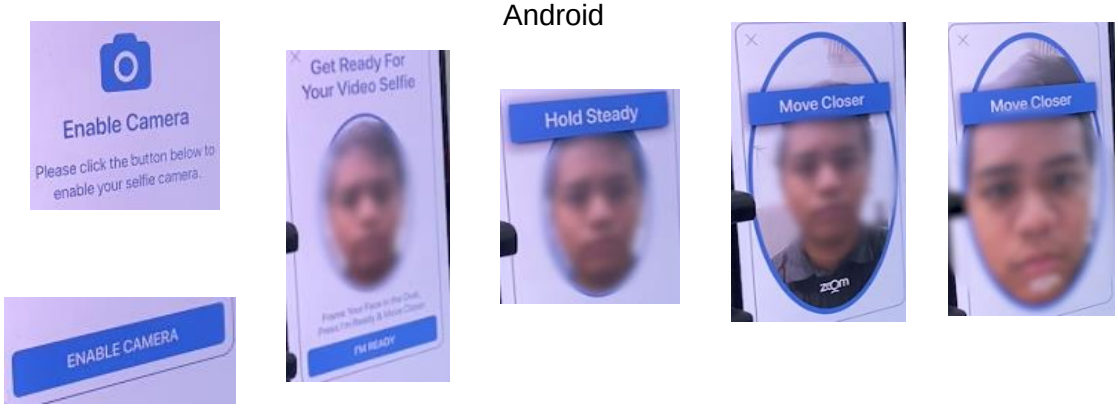
2. The initial screen will ask for you to enter your User ID. This will be supplied to you by your contact at *AGENCY*.

Android iPhone

3. Next, the app will ask you to “Enable Camera”. Once you allow this permission, the app will present an oval frame for you to place your face in. Press the "I'm Ready" button to proceed, then position your face within the frame while holding steady. The app will prompt you to "Move Closer" and "Hold Steady" again to complete this step.
   For this step, please ensure that nothing is obscuring your face i.e., remove glasses, masks etc.



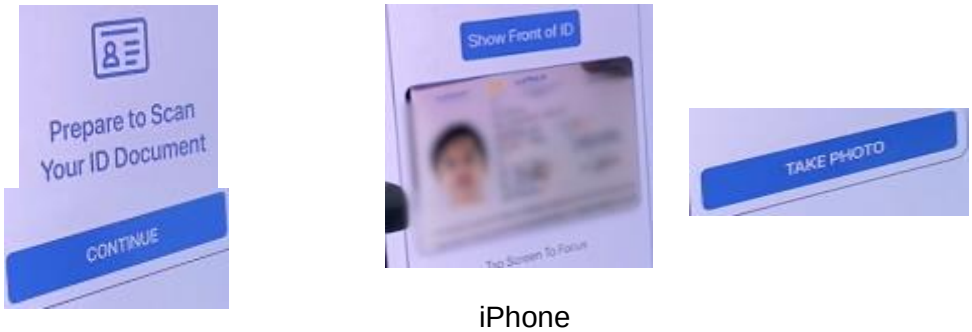
Android



iPhone

4. Following that, press “Continue”. The app will then proceed to the capturing of your passport. You will then be prompted to place your passport within a designated frame and capture an image of it.
   You may tap the screen to focus and assist with capturing an image of your passport.
   Also ensure there is no glare (or reflections) on your passport.

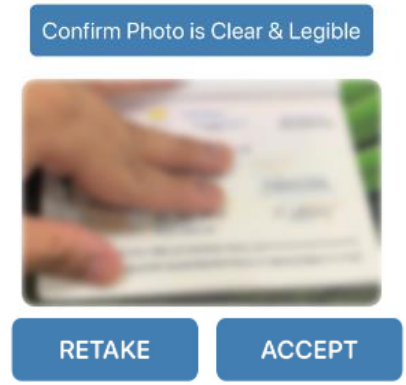
Android



iPhone

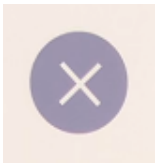
5. After you have captured the image, the app will ask you to confirm if the photo is clear & legible and you will be given two options “Retake” or “Accept”. If you are unsatisfied with the image you may press “Retake” to recapture the image, otherwise pressing “Accept” will make the app proceed to the next step.

Android

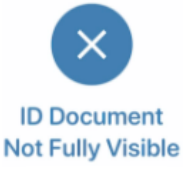


iPhone

6. The app will then check whether the image you have captured is adequate. If not, a cross icon like the one below will appear, and you will be asked to recapture the image of the passport.

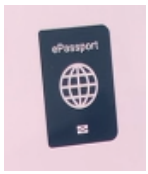


Android

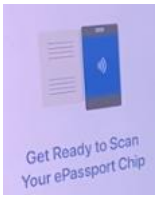


iPhone

7. After the app accepts the image of the passport, an animation will appear on the screen which indicates an NFC scan of the passport. To initiate the NFC scanning process, place the front of your passport on the back of your phone and the scanning will start automatically.
   Make sure to aim the back of your phone at the bottom of the passport, where the biometric logo is.
   If the app is unable to detect the NFC an icon will appear, and the app will ask you to rescan the NFC. Please reattempt until it reads the NFC completely and do NOT press the “Skip this step” button that will appear.



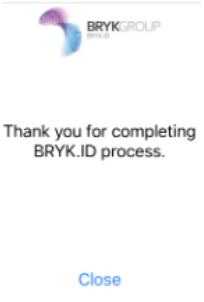
Android



iPhone

8. Finally, a screen with a "Start Over" or ‘Close’ button for Android and iPhone respectively, will appear. This indicates that you have completed the steps from start to finish.

Android



iPhone

9. Repeat step 1-9 two more times to ensure you have completed 3 total transactions.

Once you have completed the steps above 3 times, fill out the ‘*AGENCY* Post Testing Survey’ and send this form to *AGENCY*. This will conclude your testing session.