Psychometric Functions Within the Framework of Binary Signal Detection Theory: Coding the Face Identity

Petro Gopych[1], Anna Kolot[2]

[1] Universal Power Systems USA-Ukraine LLC, 3 Kotsarskaya st., Kharkiv 61012, Ukraine
[2] Prof. L.L. Girshman Municipal Clinic no. 14, 5 Oles Gonchar st., Kharkiv 61023, Ukraine
pmg@kharkov.com

Abstract. One of standard methods in vision research is measuring the psychometric functions (PFs) that are further analyzed implying the validity of traditional signal detection theory (SDT). This research paradigm contains essential inherent contradiction: in contrast to most empirical PFs the ones predicted by the SDT do not satisfy the Neyman-Pearson objective. The problem may successfully be overcome within the framework of recent binary signal detection theory (BSDT) providing PFs for which the objective required is always achieved. Here, the original BSDT theory for vision is for the first time applied to quantitative description of specific empirical PFs measured in experiments where the coding of facial identity has been studied. By fitting the data, some parameters of BSDT face recognition algorithm were extracted and it was demonstrated that the BSDT supports popular prototype face identification model. Results can be used for developing new high-performance computational methods for face recognition.

Keywords: neural networks, generalization through memory, Neyman-Pearson objective, face recognition, prototype face identification model

1 Introduction

Psychometric functions (PFs, hit rates vs. stimulus magnitudes [1,2]) are widely used in vision research, e.g., [3-8]. Implying (often implicitly) the validity of traditional signal detection theory (SDT) [1,2], they are further quantitatively described by two parameters: empirical stimulus threshold and the slope of PF curve at the point of subjective equality (the stimulus for which responses "yes" occur on 50% of trials). Such a consideration, in spite of its almost ubiquitous use, is essentially inappropriate to most psychophysics (including vision) experiments because subjects are usually instructed, explicitly or implicitly, when evaluating stimuli of different magnitudes to keep constant a given false-alarm rate ― the goal (so-called Neyman-Pearson objective [2]) that for SDT PFs can never be achieved even in principle. This often-ignored fact hinders the complete theoretical description of observed PFs and the discovering of underlying processes that generate their specific shapes.

   This SDT fundamental drawback can be overcome using the recent binary signal detection theory (BSDT [9-14]) because its PFs satisfy the Neyman-Pearson objective

and the BSDT provides its own parameter (false-alarm probability, *F*), specifying the measured PFs and reflecting the fact that the objective required is always achieved.
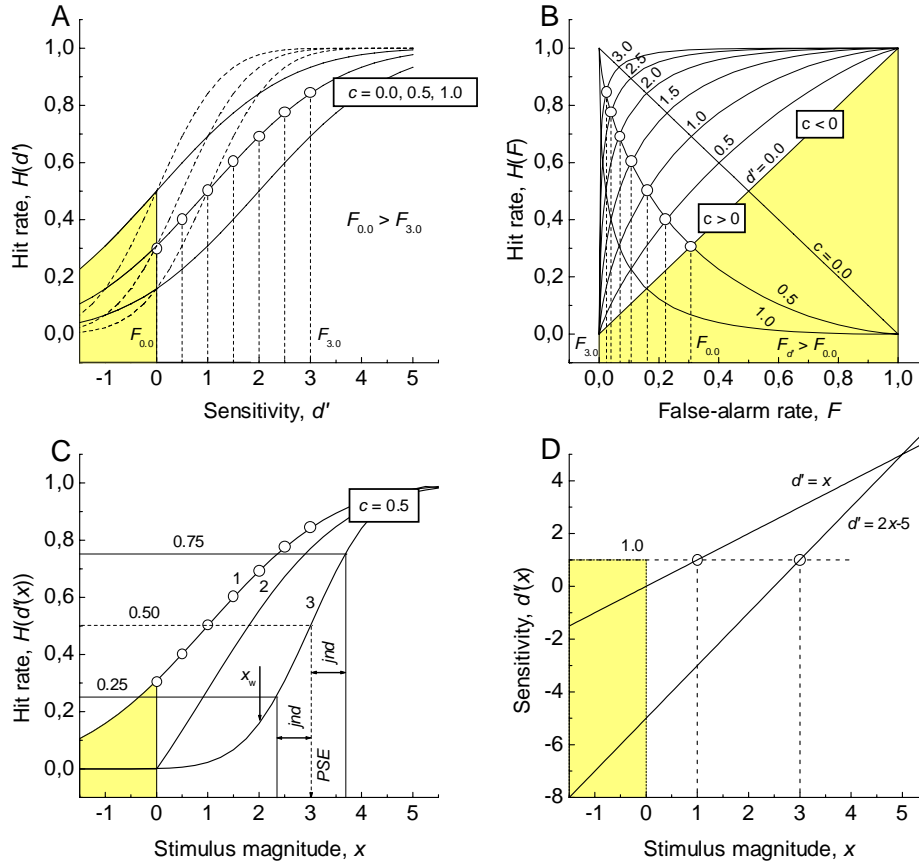
In this paper, as a development of (or an alternative to) the SDT, we introduce a BSDT description of empirical PFs and compare the both approaches' similarities and distinctions, advantages and disadvantages. The BSDT theory for vision [12] allowing generalization from the small number of examples or even from a single example is applied to the BSDT explanation of PFs measured in face identity experiments. It has been demonstrated that BSDT qualitative and quantitative predictions are consistent with empirical data [6,7] and a popular prototype face identification model [3-8].

2 Disadvantages of SDT Psychometric Functions

The SDT defines two quite separate spaces: the stimulus space where all an animal's input stimuli are being processed and a hypothetical decision space where stimulus magnitudes appear as locations (usually distributed as a Gaussian) on a decision parameter axis. On the one hand, the distinctness of these spaces is the SDT's advantage, because it allows to remain out of considering the problem of discovering specific mechanisms of input stimulus transduction into variables in decision space and to focus on solving, on the basis of mathematics and probability theory, decision-making problems in the space deliberately designed for this purpose. On the other hand, that is the SDT's disadvantage simultaneously because relations of input stimuli to their decision space representations (which are of essential interest) can be never discovered in principle as the nature of these representations is unknown and unspecified *by definition*. In spite of that, the SDT explicitly implies that stimulus values are monotonically (or even linearly) related to the values of their decision space representations [2]. As PFs are functions of stimulus magnitude, for their shape descriptions, that is a mandatory requirement (see Fig. 1C and D).

In order to explain the results of psychophysics experiments (hit rates measured, *H*, and false alarms, *F*), two basic parameters are introduced in SDT decision space [1,2]: sensitivity, d' (the ability to discriminate), and bias, *c* (a tendency to use the response irrespective to the stimulus presented). SDT performance functions are shown in Fig. 1: hit rates vs. sensitivity or basic decoding performance, BDP, curves (A); hit rates vs. false-alarm rate or receiver operating characteristic, ROC, curves (B); hit rates vs. stimulus magnitude or psychometric function, PF, curves (C); sensitivity vs. stimulus magnitude or a model for input stimulus signal transduction (D).

BDPs and ROCs are *universal* functions, not depending on the stimulus model. PFs are the BDPs rescaled under the assumption of validity of a stimulus model that produces particular functions relating the stimulus and its representation (the simplest examples are in Fig. 1D). As for each BDP the Neyman-Pearson objective is not achieved (Fig. 1A), the same holds for any PF based on it. Hence, SDT PFs (as in Fig. 1C) cannot be used for explaining most of empirical PFs even if their shapes would coincide completely. Simultaneously, SDT statistics for the parametrization of empirical PFs (the threshold, PSE, jnd or the slope of PF curve, Fig. 1C) retain their initial sense because they were defined only assuming the specific ogive shape of SDT (and empirical) PFs. The latter would perhaps excuse SDT fundamental inability of describing empirical PFs and explain why this fact is so often ignored.

Fig. 1. SDT performance functions. **A**, BDPs (isobias curves) for criteria $c = 0.0$, 0.5, and 1.0; dashed and solid curves, cumulative distribution functions for Gaussians with arguments $z = d'$ and $z = 2d'$, respectively (for drawing the left-most dashed curve, Table 5A.2 in [2] can be used); circles on the curve $c = 0.5$, $z = 2d'$ are the same as in **B**; for each circle, its specific false alarm, $F_{d'}$, may be readout from horizontal axis in **B** ($F_{0.0} > F_{0.5} > F_{1.0} > F_{1.5} > F_{2.0} > F_{2.5} > F_{3.0}$; $F_{d'} \to 0$ as $d' \to +\infty$ and $F_{d'} \to 1$ as $d' \to -\infty$; hence, here the Neyman-Pearson objective is not achieved). **B**, ROCs (isosensitivity curves) for sensitivities $d' = 0.0$, 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 (Eq. 1.8 in [2]) and isobias curves for $c = 0.0$, 0.5, and 1.0 (Eq. 2.7 in [2]); circles (the same as in **A**) indicate cross-points of all the ROCs shown with the curve $c = 0.5$ (dashed lines point to their false alarms, $F_{d'}$). **C**, PFs (isobias curves) for $c = 0.5$ and given linear stimulus transduction models shown in **D**; curves 1 and 2, $d'(x) = x$ (curve 2 is curve 1 corrected for guessing, Eq. 11.6 in [2]); curve 3, $d'(x) = 2x - 5$; jnd, just noticeable difference (in a similar way for curve 1, jnd cannot be found because negative x will be involved); PSE, the point of subjective equality or an empirical threshold; x_w, the weakest detectable stimulus for curve 3 (another though rarely used definition of this threshold); given $d'(x)$, in a broad range of c (if $c > c_0$ where c_0 is large enough), PFs produce the same jnd and different PSEs; different linear $d'(x)$ generate different jnd and c_0. **D**, Examples of physical stimulus transduction models (cf. Fig. 5.3a in [2]), $d'(x) = x$ (cf. curves 1 and 2 in **C**) and $d'(x) = 2x - 5$ (cf. curve 3 in **C**); circles and dashed lines indicate $x = 1$ and $x = 3$ corresponding to $d' = 1$ (in **C** on curves 1 and 3 they lead to $H(x) = 0.5$). In **A**, **B**, **C**, and **D**, nonphysical areas ($d' < 0$ or $x < 0$) are shaded.

3 Advantages of BSDT Psychometric Functions

The BSDT [9-14] operates with N-dimensional vectors x with their components $x^i = \pm 1$, a reference vector $x = x_0$ representing the information stored or that should be stored in a neural network (NN), binary noise $x = x_r$ (the signs of its ± 1 components are randomly chosen with uniform probability, ½), and vectors $x(d)$ with components

$$x_i(d) = \begin{cases} x_0^i, & if \quad u_i = 0, \\ x_r^i, & if \quad u_i = 1 \end{cases} \quad d = \sum u_i / N, \quad i = 1,...,N \tag{1}$$

where u_i is 0 or 1. If m is the number of marks $u_i = 1$ then $d = m/N$, $0 \le d \le 1$; d is a fraction of noise components in $x(d)$, $q = 1 - d$ is a fraction of intact components of x_0 in $x(d)$ or an intensity of the cue (cue index), $0 \le q \le 1$. If $d = m/N$, the number of different $x(d)$ is $2^m C^N_m$, $C^N_m = N!/(N-m)!/m!$; if $0 \le d \le 1$, this number is $\sum 2^m C^N_m = 3^N$ ($m = 0, 1, ..., N$). As the set of $x(d)$ is complete, always $x = x(d)$.
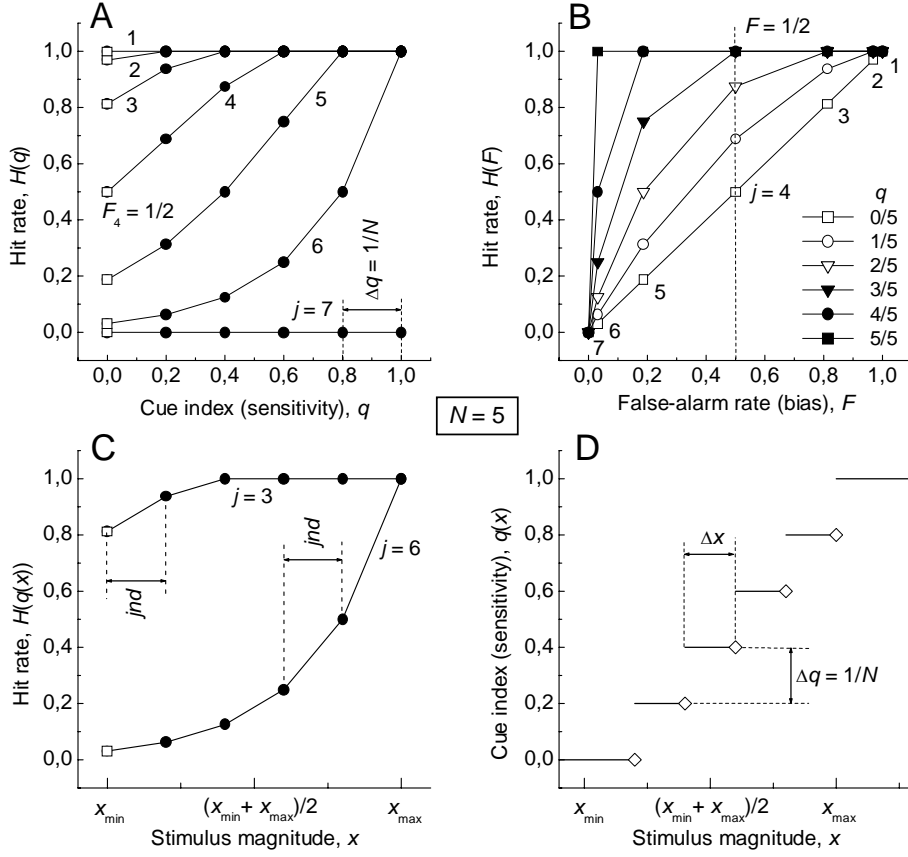
BSDT decoding algorithm exists in NN, convolutional, and Hamming forms that are equivalent and the best in the sense of pattern recognition quality [13,14]; all its predictions are essentially discrete. For intact perfectly learned NNs, the hit rate, H (or correct decoding probability of vectors x, P), can be calculated analytically [9,14]:

$$P(N,m,\Theta) = \sum_{k=0}^{K} C_k^m / 2^m, \quad K_0 = \begin{cases} (N-\Theta-1)/2, & if \quad N \quad is \quad odd \\ (N-\Theta)/2-1, & if \quad N \quad is \quad even \end{cases} \tag{2}$$

where Θ is an even integer θ, $-N \le \Theta < N$ (θ is the neuron's triggering threshold, $-\infty < \theta < +\infty$); if $K < K_0$ then $K = m$ else $K = K_0$ (k, the Hamming distance from $x = x(d)$ to x_0; K, its threshold value; K_0, the K for a given Θ). If $\Theta < -N$ then $P(N,m,\Theta) = 1$, if $\Theta \ge N$ then $P(N,m,\Theta) = 0$. For any $\theta \in \Delta\theta_j$, $P(N,m,\theta) = P(N,m,\Theta_j)$ where $\Theta_j \in \Delta\theta_j$ (here, $j = 0, 1, 2, ..., N+1$; $\Theta_j = 2j - N - 1$). If $0 < j < N + 1$ then $\Delta\theta_j = [\Theta_j - 1, \Theta_j + 1)$ and $\Delta\theta_j = [\Theta_j, \Theta_j + 2)$ for odd and even N, respectively; if $j = 0$ and $N + 1$ then $\Delta\theta_0 = (-\infty, -N)$, $\Delta\theta_{N+1} = [N, +\infty)$ and $P(N,m,\Theta_0) = 1$, $P(N,m,\Theta_{N+1}) = 0$ [9,10].

For a given N, $P(N,m,\Theta_j) = P(N,q,j) = P(q,F) = H(q,F)$ where j is the confidence level of decisions [13] and $F = F_j$ (F, false-alarm probability). If q (cue index) is fixed then $P(q,F) = P_q(F) = H(F)$, this function is called receiver operating characteristic (ROC) curve; if F is fixed then $P(q,F) = P_F(q) = H(q)$, this function is called basic decoding performance (BDP) curve. ROCs and BDPs are fundamentally discrete-valued although as $N \to \infty$ they tend to become continuous. ROCs at all q and BDPs at all F provide the same information [the lattice of $(N + 1) \times (N + 2)$ values of $P(q,F) = H(q,F)$]. Examples of BDPs and ROCs are in Fig. 2A and B, respectively.

As BSDT BDPs [or $P(q,F) = H(q)$] are defined at a given F, for them the Neyman-Pearson objective is achieved *by definition*. For this reason, theoretical BDPs (Fig. 2A) may naturally be used for explaining empirical PFs (Fig. 2C) and in this case, in contrast to the SDT, any contradiction between theoretical and empirical assumptions does not appear. The fruitfulness of this approach was earlier demonstrated by the comparison of theoretical BDPs found for a BSDT NN local feature discrimination algorithm (peak search code *PsNet* [15]) to appropriate empirical data in humans [16] (in fact, they coincide completely, ref. 15 and its Fig. 6).

Fig. 2. BSDT performance functions. **A**, Complete set of $N + 2$ BDPs given N (signs connected by strait-line segments); j, the confidence level of decisions (the BDP's series number) [13]; F_j (open square), false alarm or 'bias' specific to the jth BDP (hence, in this case the Neyman-Pearson objective is always achieved); Δq, minimal cue-index interval (a BSDT dimensionless counterpart to the jnd). **B**, Complete set of $N + 1$ ROCs given N (signs connected by strait-line segments); each ROC corresponds to its specific cue index, q ($q = 1 - d$ is the BSDT's sensitivity measure [10], ROCs and their cue indices are related in the insertion); dashed line indicates hit rates for $F = 1/2$ (the same as along curve 4 in **A**). **C**, Two ($j = 3$ and 6) of $N + 2$ possible PFs for the stimulus transduction model shown in **D** (designations as in **A**, other PFs are not shown in order to not overload the picture); $jnd = (x_{max} - x_{min})/N$, just noticeable difference (the PSE is not here required; cf. Fig. 1C). **D**, The simplest example of a possible model, $q(x)$, of the stimulus' physical signal transduction into the activity of decision-making neurons — a set of horizontal line segments defined by the rules: if $x \in (-\infty, x_1)$ then $q(x) = 0$, if $x \in [x_k, x_{k+1})$ then $q(x) = k/N$ ($k = 1, ..., N - 1$), if $x \in [x_N, +\infty)$ then $q(x) = 1$; $x_k = k\Delta x$ ($k = 1, ..., N$), $\Delta x = (x_{max} - x_{min})/N$; $q(x)$ is a piecewise constant single-valued function, in points of discontinuity open diamonds designate where $q(x)$ is not defined (cf. [11,14]); Δq, minimal cue-index interval (the same as in **A**). Hit rates were analytically calculated (Eq. 2) for a perfectly learned intact NN of the size $N = 5$ (it may be an apex NN for the learned NN hierarchy related to a neural subspace [12], in particular, to a face neural subspace discussed in Section 4).
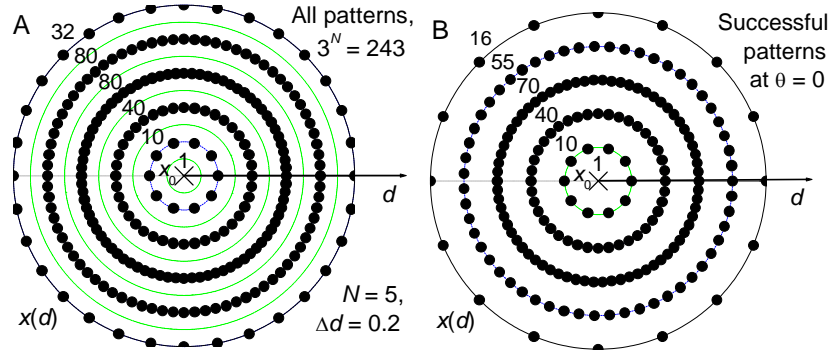
Another BSDT advantage is the existence of a tool (a *neural space* [10,12]) for the relating of particular stimulus signals to the activity of particular decision-making neurons or, in other words, for discovering in principle the function $q(x)$ (Fig. 2D) on the basis of neuroscience information. BSDT neural space, which is common for input stimuli and decisions concerning them, is a hierarchy of universal NN units of known structure each of which is learned to recognize optimally its own reference pattern x_0 [12,14] though, as final decisions concerning particular stimulus processing are made by the NN (NN memory unit) sitting at the apex of particular hierarchy, the circuits consuming processing results see the apex's NN output only (cf. Fig. 2) and have no direct information on the hierarchy itself. For this reason, for the description of a stimulus, it is often enough to use its apex NN and x_0 stored in it (in such cases, for external observers, the hierarchy's internal structure is not essential in general). Because each type of inputs is being processed by its specific hierarchy of learned NNs, we refer to this neuron space fraction as a *neural subspace* (or a *semi-representational* trace of memory [12] dealing with particular type of stimuli and exploring a kind of predictive coding [17])—it should be purposefully designed in machines or anatomically predefined in living organisms. Hence, BSDT neural space (the whole semi-representational memory) consists of a huge amount of overlapping neural subspaces (semi-representational memory traces) constituting a milieu where (brain) input codes (stimuli) are being processed and decisions are being made.

4   BSDT Face Neural Subspace, FNS

Generalization problem is traditionally considered in the context of the classic learning theory as a problem of the best approximate interpolation of some data points ('examples') by a rather smooth continuous function—generalization from examples [18]. The more the number of examples the more productive this approach is while if there is the only example then it does not work at all. It is implicitly assumed that the transition from one empirical point (or image) to any other one is being performed continuously, through an *infinite* number of other intermediate points. As memories of infinite capacity are impossible, such intermediate patterns are calculated by interpolating among examples. The BSDT offers another approach—generalization in discrete finite dimensional spaces (Fig. 3) or through NN memories (such as apex NNs mentioned above) each of which stores a single binary pattern x_0 only [12].

Because of Eq. 1 each $x(d)$ may be interpreted as specific x_0 damaged by noise to the damage degree $d = m/N$. For this reason, BSDT BDPs and ROCs (as in Fig. 2A and B) describe also *generalization ability* of an NN memory unit sitting at the apex of a NN hierarchy concerning particular neural subspace—the probability (Eq. 2) of interpreting $x(d)$ as x_0 or retrieving the x_0 if the retrieving process is initiated by $x = x(d)$. This very special role of $x_0 = x(0)$ provides the BSDT's possibility to generalize even from a single example, x_0 (because any $x(d)$ may be considered as a noised version of x_0, see also Eq. 7 in [12]). Generalization through memory also explores the idea of optimal recognition of a pattern of any complexity if it is presented in a semi-representational way**,** as its specific hierarchy of learned NN units (i.e., as a BSDT neural subspace [10,12]). The rules required for designing the hierarchy are so far explicitly not known but, as we can see, in many cases this fact is not essential.
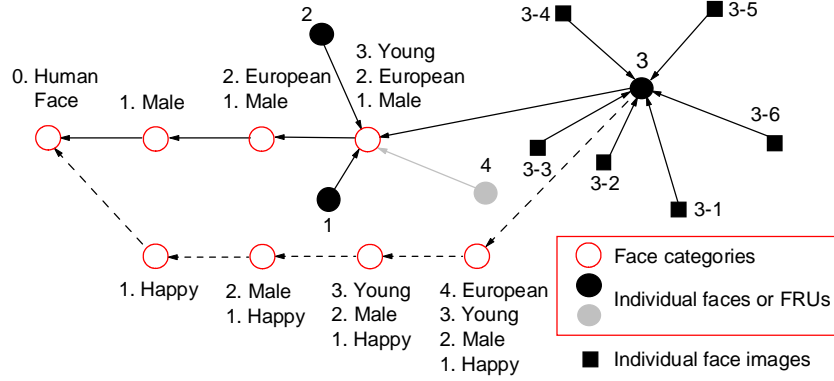
Without any loss of generality, here we say of one specific sort of patterns only, human faces (Fig. 3). If so, reference face (RF) x_0 of a BSDT face neural subspace (FNS) is a direct counterpart to a stored in memory norm, prototype, mean/average, or abstract face against which, as it is assumed in current vision theory, the face identity is coded and which is dynamically tuned by experience [3-8]. For details see Fig. 4.

Fig. 3. BSDT N-dimensional ($N = 5$) face neural subspace (FNS) corresponding to reference face (RF) x_0. **A**, Complete set of all the faces $x(d)$ (the versions of x_0) constituting the FNS; cross in the center, RF x_0 [$x(d)$ at $d = 0$, $x_0 = x(0)$]; filled points at the distance d of the cross, faces $x(d)$ given their damage degree $d = m/N$ ($m = 1, …, 5; \Delta d = 1/N$); numbers, the amount ($2^m C^N_m$) of different $x(d)$ given d. **B**, Complete set of all successful (successfully recognizable) faces at $\theta = 0$ (θ, neuron triggering threshold for an apex NN; other designations as in **A**); numbers, the amount of successful faces $x(d)$ given d and θ found by Eq. 2 (data are the same as in Fig. 2A, curve 4). Each face of the FNS in **A** is only designated as $x(d)$ stored in apex NN of its NN hierarchy while the hierarchy itself is not shown; at early stages of processing, different hierarchies share common neural substrates (this claim is empirically supported [4]); each $x(d) \neq x_0$ in **A** may serve as a RF for a similar FNS embedded into the initial one, in turn each $x(d) \neq x_0$ of previous FNS may be a center for a next FNS embedded into previous one, etc (such a discrete self-similar conserving the scale fractal-like hierarchy of embedded spaces may explain the existence of brain areas consisting entirely of face-selective cells [19]).

The dimensionality (N) of a BSDT FNS and the total number (3^N) of its separate faces (if embedded FNSs are not taken into account) are defined by the size of the FNS's apex NN, N (in Fig. 3 $N = 5$). Like the norm face of a traditional ('morphable') human face space [3-8], the RF of an FNS is also in its center (crosses in Fig. 3, circles in Fig. 4) and the more the distance from the center (RF damage degree, d) the more the face individuality is. A given $d = m/N$ may correspond to $2^m C^N_m$ different individual faces $x(d)$. By changing d, we change face individuality parameters and, then, the BSDT's HFRA (human face recognition algorithm) corresponding to a given FNS [12] may produce BDPs (Fig. 2A) or PFs (Fig. 2C) which are suitable for their comparison with empirical PFs [3-8] (if signal transduction model is as in Fig. 2D). The number of theoretical points along a BSDT PF is defined (fitted) by N, the size of the NN at the apex of particular FNS hierarchy. If an image $x(d)$ contains 0% of the target face ($d = 1$, $q = 0$) then BSDT FNS, running as a HFRA, produces false-alarm recognition probability, F. As target face percentage increases ($d \to 0$, $q \to 1$), the recognition probability also increases approaching 1 (Fig. 2 A and C). Comparing at a

given *F* the shapes of any two empirical PFs, we may reveal which one corresponds to better (more powerful) decoding/recognition/generalization algorithm [15].

Fig. 4. Examples of possible hierarchies of embedded BSDT FNSs. Circles, RFs for the 'main' FNS (left-most circle) and for some its embedded spaces (their NN hierarchies are not shown); arrows are directed from RFs of embedded FNSs to RFs of their embodied FNSs; open and filled circles correspond to categories of faces and individual faces, respectively (the latter are counterparts to face reference units, FRUs [20]); near each open circle, its meaning is written in the number of rows that is the number of the level of embedding (e.g., 'European Male' is the RF for European-male FNS—the subspace of all encountered European-male faces, here its embedding level is 2); near each filled circle is its individual number that has nothing common with the personal name of corresponding face and is only used in this figure (circles 1 to 3, familiar faces; gray circle 4, 'unfamiliar' face having no its developed FNS). Squares, individual (e.g., view-dependent) face images (3-1 to 3-6, six such images of face 3; for faces 1 and 2, their FNSs are not displayed; RF 4 is a single individual face image ensuring generalization from this single example). Solid and dashed arrows indicate two (of many) possible embedding hierarchies for the current recognition of face 3 (in the case of success, the face is perceived as a familiar one no matter either its personal name storing in semantic memory is remembered or not; the binding of visual and semantic memory traces is here not discussed).

5 BSDT Explanation of Face Identity Experiments

First, we consider PFs measured along face identity trajectories 'consisting of morphs, in which the contribution of the target face is varied from 0% (no target face) to 80%' [6] (in such a way 'the face identity aftereffect for computationally opposite and non-opposite adapt-test pairs' [6] has been studied). For the parametrization instead of SDT empirical thresholds ([6] and Fig. 1C), BSDT false alarms *F* (recognition rate of 0% target face, Fig. 2A and C) are used, taking into account the shape of PF curves.

For opposite adapt-test pairs [6, Fig. 5 top] all PFs have a common false alarm, $F_{\text{opposite}} \approx 0.05$, similar to that for skilled subjects, $F_{\text{skill}} = 0.012 \pm 0.004$ [16]. For %Target ≥ 0, PF(Opposite Adaptation) $\geq$ PF(Baseline-Pre) and PF(Baseline-Post) $\approx$ PF(Baseline-Pre). Consequently, after the adaptation, the genuine (so far unknown but sought-for) human face recognition algorithm (HFRA) becomes uniformly more powerful and the face identity aftereffect is simply a manifestation of this mathematical (decision theory) fact. For non-opposite adapt-test pairs [6, Fig. 5

bottom], all the PFs have probably different F similar to that typical for dilettantes, $F_{dilettants} \approx 0.3$ [16], and $F_{opposite} < F_{non-opposite}$. Very large variability of F for some individuals (e.g., MZ) says about high instability of decision criterion (or decision confidence [13]) they implicitly used. Because all F are different, corresponding PFs (quality/power functions of the HFRA) are strictly not comparable; in that case, aftereffects may be caused by the difference in F as well as the HFRA current quality. Power increasing of the HFRA, after its adaptation to an opposite face, means that opposite and test faces are on the same identity trajectory defined by a common (for this adapt-test pair) set of face identity parameters that were adjusted (tuned) in the course of the HFRA adaptation. For non-opposite adapt-test pairs, that is not the case.

For opposite pairs in three successive baseline sessions [6, Fig. 6 left], all the PFs have equal false alarms, $F \approx 0.12$, and from session to session the quality of the HFRA becomes slightly better (the small 'learning effect' [6]). For non-opposite pairs [6, Fig. 6 right], from session to session F increases from ≈ 0.12 to ≈ 0.41 (to the level of dilettantes [16]) and that is the so-called 'strong learning effect' [6].

For opposite and non-opposite trajectories in interleaved experiments [6, Fig. 7], before and after adaptation, values of F are probably the same, although $F_{opposite} < F_{non-opposite}$. After its adaptation, the HFRA becomes more powerful and that is the cause of aftereffects observed.

View-dependent face images may also be considered. In particular, poor HFRA performance (squares in Fig. 3 of ref. 7) is naturally explained assuming that, during the adaptation, the HFRA was tuned to one face while it was tested by another one (related to another FNS with another RF in its center, Figs. 3 and 4).

Of the finding that the magnitude of identity aftereffects increases as familiarity with faces increases [21] follows that familiarity enhances the HFRA's recognition power. The idea of adaptation score as a measure of familiarity effects [20] is consistent with the BSDT if PFs will be measured at a given F.

We see that empirical PFs are consistent with the notion of existence of an RF against which individual faces are coded and the BSDT HFRA [an FNS devoted to the recognition of an $x(d)$] is tuned. Adaptation to a particular face makes the HFRA more powerful, because of better matching of its parameters to individuality parameters of the adaptation face (for the lack of space, complete comparison of theoretical BSDT predictions with empirical PFs [6] is not shown). For recognizing individual faces [$x(d)$, $d > 0$, points in Fig. 3], the HFRA uses the more branched path of embedded FNSs (Figs. 3 and 4) the more the d is and, consequently, generates such neuron population responses as for empirical face tuning curves [5,8] having their minima at their RFs. Thus, the BSDT explains all the facts that traditional prototype face model explains [3-8] and even more (see legends to Figs. 3 and 4).

5 Conclusion

BSDT approach [9-14] to the description of PFs has for the first time been applied to account for PFs measured in face identity experiments. Prototype face tuning by experience and face identity aftereffects [6,7], face tuning curves [5,8], existence of brain areas consisting entirely of face-selective faces [19] agree with the BSDT which, then, might be considered as an NN substantiation of prototype face

identification model [3-8]. The results can be used for developing a new robust high-performance NN computational method for biologically plausible face analysis.

References

1. Green, D., Swets, J.: Signal Detection Theory and Psychophysics. Wiley, New York (1966)
2. Macmillan, N.A., Creelman, C.D.: Detection Theory: A User's Guide. 2nd edn. Lawrence Erlbaum Associates, Mahwah (2005)
3. Leopold, D.A., O'Toole, A.J., Vetter, T., Blanz, V.: Prototype-Referenced Shape Encoding Revealed by High-Level Aftereffects. Nature Neurosci. **4** (2001) 89-94
4. Anderson, N.A., Wilson, H.R.: The Nature of Synthetic Face Adaptation. Vision Res. **45** (2005) 1815-1828
5. Loffer, G., Yourganov, G., Wilkinson, F., Wilson, H.R.: fMRI Evidence for the Neural Representation of Faces. Nature Neurosci. **8** (2005) 1386-1390
6. Rhodes, G., Jeffery, L.: Adaptive Norm-Based Coding of Facial Identity. Vision Res. **46** (2006) 2977-2987
7. Ryu, J.-J., Chaudhuri, A.: Representations of Familiar and Unfamiliar Faces as Revealed by Viewpoint-Aftereffects. Vision Res. **46** (2006) 4059-4063
8. Leopold, D.A., Bondar, I.V., Giese, M.A.: Norm-Based Face Encoding by Single Neurons in the Monkey Inferotemporal Cortex. Nature **442** (2006) 572-575
9. Gopych, P.M.: ROC Curves within the Framework of Neural Network Assembly Memory Model: Some Analytic Results. Int. J. Inf. Theo. Appl. **10** (2003) 189-197
10. Gopych, P.M.: Sensitivity and Bias within the Binary Signal Detection Theory, BSDT. Int. J. Inf. Theo. Appl. **11** (2004) 318-328
11. Gopych, P.M.: Neural Network Computations with Negative Triggering Thresholds. In: Duch, W. et al. (eds): Artificial Neural Networks: Biological Inspirations. Lecture Notes in Computer Sciences, Vol. 3696. Springer-Verlag, Berlin-Heidelberg (2005) 223-228
12. Gopych, P.M.: Generalization by Computation Through Memory. Int. J. Inf. Theo. Appl. **13** (2006) 145-157
13. Gopych, P.M.: Performance of BSDT Decoding Algorithms Based on Locally Damaged Neural Networks. In: Corchado, E., Yin, H., Botti, V., Fyfe, C. (eds): Intelligent Data Engineering and Automated Learning. Lecture Notes in Computer Sciences, Vol. 4224. Springer-Verlag, Berlin-Heidelberg (2006) 199-206
14. Gopych, P.M.: Foundations of the Neural Network Assembly Memory Model. In: Shannon, S. (ed.): Leading-Edge Computer Sciences. Nova Science, New York (2006) 21-84
15. Gopych, P.M.: Identification of Peaks in Line Spectra Using the Algorithm Imitating the Neural Network Operation. Instr. Exp. Tech. **41** (1998) 341-346
16. Gopych, P.M., Sorokin, V.I., Sotnikov, V.V.: Human Operator Performance when Identifying Peaks in a Line Spectrum. Instr. Exp. Tech. **35** (1992) 446-449
17. Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., Hirsch, J.: Predictive Codes for Forthcoming Perception in the Frontal Cortex. Science **314** (2006) 1311-1314
18. Poggio, T., Bizzi, E.: Generalization in Vision and Motor Control. Nature **431** (2004) 768-774
19. Tsao, D.Y., Freiwald, W.A., Tootell, R.B.H., Livingstone, M.S.: A Cortical Region Consisting Entirely of Face-Selective Cells. Science **311** (2006) 670-674
20. Burton, A.M., Bruce, V., Hancock, P.J.B.: From Pixels to Peoples: A Model of Familiar Face Recognition. Cog. Science **23** (1999) 1-31
21. Jiang, F., Blanz, V., O'Toole, A.J.: The Role of Familiarity in Three-Dimensional View-Transferability of Face Identity Adaptation. Vision Res. **47** (2007) 525-531