Multi-layered perceptron as a model for the pupillary pathway

Abstract Derived from models of neural networks, a model for the pupillary pathway is introduced that can easily be handled computationally and analytically. To model the binocular pupillary reactions we use a feed-forward network, namely, a multi-layered perceptron. It is possible to calculate the pupillary reactions analytically as a function of the light stimuli of the retinal hemifields, on the one hand, and the set of neural couplings between the neural layers, on the other. Several lesions, e.g., lesions of the intercalated neurons between the afferent and efferent pupillary pathways, can be simulated within the model and the corresponding pupillary reactions such as anisocoria or relative afferent pupillary defects (RAPD) can be calculated analytically. Due to its neural network structure the model described herein can easily be extended to much more complex pupillary pathways while remaining calculable computationally and analytically.

Key words Neural networks · Multi-layered perceptron · Anisocoria · RAPD · Pupil

Zusammenfassung Abgeleitet von Modellen neuronaler Netzwerke stellen wir ein Modell für die Pupillenbahn vor, welches sowohl numerisch (PC) als auch analytisch einfach zu handhaben ist. Um die binokularen Pupillenreaktionen zu modellieren, verwenden wir ein Feed-forward-Netzwerk (hier mehrschichtiges Perzepton). Die Pupillenreaktionen können analytisch als Funktion der Lichtreize auf die retinalen Halbfelder und als Funktion der synaptischen Kopplungen zwischen den neuronalen Schichten berechnet werden. Verschiedene Läsionen, z.B. Läsionen von Interneuronen zwischen afferenten und efferenten Pupillenbahnen, können simuliert und die resultierenden Pupillenreaktionen, wie z.B. Anisokorien oder relative afferente Pupillenstörungen (RAPD), analytisch berechnet werden. Aufgrund seiner Netzwerkstruktur kann das oben beschriebene Modell leicht auf komplexere Pupillenbahnssysteme erweitert werden, ohne seine numerische bzw. analytische Berechenbarkeit zu verlieren.

Schlüsselwörter Neuronale Netze · Mehrschichtiges Perzepton · Anisokorien · RAPD · Pupille
Introduction

The method of neural networks has developed into a promising feature to describe processes taking place in the central nervous system. An extensive introduction to neural network theory can be found elsewhere (e.g., [11, 15]). Therefore, only some basic elements concerning neural networks are explained herein.

A biological neuron, shown in Fig. 1a, consists of three important parts from a physical point of view. First, the dendrite tree collects all the postsynaptic potentials emitted by other neurons; thus, the dendrites act as input channels of the neuron to which they belong. Second, the soma sums up all the collected postsynaptic potentials and generates a new synaptic potential dependent on a certain threshold. Finally, the axon represents a channel along which the new synaptic potential is propagated toward other neurons, thereby acting as the output channel of the neuron.

Modeling such a biological neuron may result in an artificial neuron such as that shown in Fig. 1b, which was first proposed in 1943 by McCulloch and Pitts [14] and is thus referred to as the McCulloch-Pitts neuron. Like its biological counterpart, this neuron also consists of some input and output channels as well as a summing and comparison unit. The postsynaptic potentials emitted by other neurons are then replaced by arbitrary input data such as numbers or pixels of a picture. These input data are then multiplied by some weight factors that represent neural coupling strengths. The result is denoted as the local field. All the local fields are summed and compared with a certain threshold. If the local field exceeds the threshold, a new output is calculated by means of a certain transfer function and the neuron becomes active; otherwise, the neuron remains inactive. This output is then transferred to one or more neurons by means of neural couplings of different strengths.

With this kind of artificial neuron, neural networks of different architectures can be constructed. In Fig. 2, one of the most important types of networks can be seen: the so-called feed-forward net, better known in this special case as a multi-layered perceptron. It consists of three parts: one input layer consisting of input neurons often denoted as input units; one or more hidden layers containing hidden units; and, finally, an output layer formed by one or more output units. The neural layers are connected to each other by neural couplings of different coupling strengths. The direction of the information flow is indicated by arrows. Thus, it is obvious why this type of neural net is called the feed-forward net.

The aim of this study was to apply the analytical technique of neural networks to a model of the pupillary pathway that is more complex than the classic one (e.g., [2]). We did not intend to introduce a complete new model, but we wanted to show how such models can be handled by the method of neural networks. It has to be emphasized that the classic model of the pupillary pathway comprises only one hidden layer because a direct connection between the retina and the pretectum is assumed. We chose a more complex model so as to simulate clinical findings that are unexplainable in the classic model as shown below in Results.

---

Fig. 1  a Biological neuron (after Hertz et al. [11]). b Artificial neuron of the McCulloch-Pitts type (after Hertz et al. [11])

Fig. 2  Neural net of the feed-forward type: multi-layered perceptron
Fig. 3a, b The pupillary pathway.

a Biological description.
b Translation into a multi-layered perceptron model

**Materials and methods**

Multi-layered perceptron model

Fortunately, the pupillary pathway depicted in Fig. 3a allows for a treatment with neural network theory. A translation of this biological description of the pupillary pathway into a more schematic network-like form is presented in Fig. 3b. To model the binocular pupillary reaction we use a multi-layered perceptron consisting of four neural layers, which can be identified as follows. The four retinal hemiels fields $R_j$ act as input neurons of the input layer. The following two hidden layers denote the left and right lateral geniculate nucleus (LGN), or the thalamus complex $C_i$, and the left and right pretectum $P_j$. However, it must be mentioned that from an anatomical point of view it has not been proven that the pupillary pathway synapses in the LGN region. Finally, the output units of the output layer are represented by the nuclei Edinger-Westphal $N_i$, which are assumed to control the binocular pupillary reactions quasi-directly. The output $N_i$ gives the amount of pupillary constriction in arbitrary units. Therefore, $N_i = 0$ means no constriction of the pupil $i$. The neural couplings $x_{ij}$, $y_{ij}$, and $z_{ij}$ connecting the neural layers can be adjusted according to neuroanatomical conditions such as the asymmetry of the chiasmal crossing (53% crossed, 47% not crossed [18]).

**Assumptions of the model**

Each model is based on some assumptions. The assumptions of the model presented herein are as follows. We allow for positive values only with regard to the neural coupling strengths. Moreover, we assume an overall symmetry between left and right. In particular, we assume equal light stimuli of all retinal hemiels fields and a left/right symmetry among crossing and noncrossing couplings. Furthermore, pathways that have crossed once in the chiasm do not cross again. Therefore, the couplings $y_{ij}$ and $z_{ij}$ do not occur in our model. Finally, if there are nasotemporal differences, they should be subtle. Mathematically speaking, this means:

1. **Symmetry between left and right**
   - (a) $R_1 = R_2 = R_3 = R_4$
   - (b) $x_{11} = x_{44}, x_{23} = x_{32}$
   - (c) $y_{11} = y_{24}, y_{12} = y_{23}, y_{14} = y_{21}$
   - (d) $z_{11} = z_{22}, z_{12} = z_{21}$

2. **Pathways that have crossed once in the chiasm do not cross again**
   - $y_{13} = y_{22}$; and
   - $z_{13} = 0 = y_{22}$.

3. **There are only subtle nasotemporal differences**
   - (a) $N_{12}^{temp} = N_{12}^{nasal}$ and $N_{21}^{temp} = N_{21}^{nasal}$
   - (b) $N_{12}^{temp} = N_{21}^{nasal}$ and $N_{13}^{temp} = N_{23}^{nasal}$.

The indices $r$ (right) and $l$ (left) denote which eye is illuminated, e.g., during performance of the swinging flashlight test.

**Results**

Calculation of the local fields

The calculation of the local fields at the several neurons is performed as described above. The input data of the LGN are the light stimuli $R_i$ of the retinal hemiels fields. Consequently, the local fields $C_i$ are calculated by multiplying these inputs by the coupling strengths $x_{ij}$ of the chiasm.
The same procedure is performed on the pretectum and the nuclei Edinger-Westphal, respectively, except that the input data are provided by the neurons/local fields of the preceding neural layer. The mathematical translation reads:

L/R LGN:

\[ C_i = x_{1i} R_1 \]
\[ C_2 = x_{23} R_2 \]
\[ C_3 = x_{32} R_3 \]
\[ C_4 = x_{44} R_4 \]

L/R Pretectum:

\[ P_1 = y_{11} C_1 + y_{12} C_2 + y_{14} C_4 \]
\[ P_2 = y_{21} C_1 + y_{23} C_3 + y_{24} C_4 \]

and

L/R NEW:

\[ N_1 = z_{11} P_1 + z_{12} P_2 \]
\[ N_2 = z_{21} P_1 + z_{22} P_2 \]

After some substitutions it is possible to calculate the pupillary reactions analytically as a function of external and internal parameter sets. The external parameters are the retinal light stimuli \( R_n \) whereas the couplings \( x_{ij}, y_{ij}, \text{ and } z_{ij} \) connecting the neural layers represent the internal parameters. The local fields at the nuclei Edinger-Westphal are as follows:

\[ N_1 = N_1(R; x, y, z) = (x_{11} y_{11} + x_{12} y_{12} R_3 + x_{44} y_{14} R_4) z_{11} + (x_{12} y_{12} + x_{23} y_{23} R_2 + x_{44} y_{24} R_4) z_{12} \]
\[ N_2 = N_2(R; x, y, z) = (x_{11} y_{11} + x_{12} y_{12} R_3 + x_{44} y_{14} R_4) z_{21} + (x_{12} y_{12} + x_{23} y_{23} R_2 + x_{44} y_{24} R_4) z_{22} \]

Actually, the third assumption is mathematically treated in such a way that no nasotemporal difference should occur. This results in total symmetry interneurons between the area praeptectalis and the nuclei Edinger-Westphal within the model as shown below. The nasotemporal differences in the direct pupillary reactions are:

\[ N_{1\text{temp}} = R_1 = 0 \]
\[ N_{1\text{nasal}} = R_1 = 0 \]
\[ N_{2\text{temp}} = x_{32} y_{23} R_2 \]
\[ N_{2\text{nasal}} = x_{32} y_{23} R_2 \]

and

\[ N_{1\text{temp}} = R_3 = 0 \]
\[ N_{1\text{nasal}} = R_3 = 0 \]
\[ N_{2\text{temp}} = x_{44} y_{14} R_4 \]
\[ N_{2\text{nasal}} = x_{44} y_{14} R_4 \]

From a clinical point of view, the following observations are expected: no anisocoria should occur and there should be an amaurotic fixed left pupil. Performance of the swinging flashlight test results in:

Left eye illuminated: \( R_1 = R_2 = 0 \)
\[ R_3 = R_4 = 0 \]

\[ N_{1\text{temp}} = 0 \]
\[ N_{1\text{nasal}} = 0 \]
\[ N_{2\text{temp}} = 0 \]
\[ N_{2\text{nasal}} = 0 \]

Unilateral amaurosis on the left

A unilateral amaurosis on the left side in our model (UAL) is achieved by eliminating the couplings \( x_{11} \) and \( x_{32} \) as shown in Fig. 4:

UAL \( \Leftrightarrow \) \( x_{11} = x_{32} = 0 \).

Thus, Eqs. 1 and 2 reduce to:

\[ N_{1\text{UAL}} = x_{23} y_{12} R_3 + (x_{44} y_{14} + x_{44} y_{24}) R_4 \]
\[ N_{2\text{UAL}} = x_{23} y_{12} R_3 + (x_{44} y_{14} + x_{44} y_{24}) R_4 \]

Again, the outputs at the nuclei Edinger-Westphal are equal:

\[ N_{1\text{UAL}} = N_{2\text{UAL}} \]

From a clinical point of view, the following observations are expected: no anisocoria should occur and there should be an amaurotic fixed left pupil. Performance of the swinging flashlight test results in:

Left eye illuminated: \( R_1 = R_2 = 0 \)
\[ R_3 = R_4 = 0 \]

\[ N_{1\text{UAL}} = 0 \]
\[ N_{2\text{UAL}} = 0 \]

Comparison with clinical findings

Below we focus on three major applications among others possible: (1) unilateral amaurosis, (2) unilateral tractus lesion, and (3) unilateral pretectum lesion. Aspects of interest are the appearance of anisocoria and relative afferent pupillary defects (RAPD).
Right eye illuminated: \( R_1 = R_2 = 0 \quad R_3 = R_4 \neq 0 \)
\[ \Rightarrow N_{1R} = N_{2L} = N_{1L} = 0 \]
\[ \Rightarrow N_{1L} = N_{2L} = N_{1R} = N_{2R} \neq 0. \]

As a consequence, as expected, no anisocoria occurs.

A comparison between the direct and consensual pupillary reaction reveals:
\[ N_{1L}^{UAL} = N_{2L}^{UAL} = 0 \]
\[ N_{1R}^{UAL} = N_{2R}^{UAL} \neq 0. \]

Therefore, an amaurotic fixed left pupil exists as expected.

**Tractus lesion on the left**

A unilateral tractus lesion on the left side in our model (TLL) is evoked by eliminating the couplings \( x_{11} \) and \( x_{23} \) as shown in Fig. 5:

\[ \text{TLL} \Leftrightarrow x_{11} = x_{23} = 0. \]

Thus, Eqs. 1 and 2 reduce to:
\[ N_{1L}^{TLL} = x_{32} y_{23} R_2 + (x_{44} y_{14} + x_{44} y_{24}) R_4 \]
\[ N_{2L}^{TLL} = x_{32} y_{23} R_2 + (x_{44} y_{14} + x_{44} y_{24}) R_4. \]

Again, the outputs at the nuclei Edinger-Westphal are equal:
\[ N_{1L}^{TLL} = N_{2L}^{TLL}. \]

From a clinical point of view, the following observations are expected: No anisocoria should occur and there should be an RAPD contralaterally to the lesion. Performance of the swinging flashlight test results in:

Left eye illuminated: \( R_1 = R_2 \neq 0 \quad R_3 = R_4 = 0 \)
\[ \Rightarrow N_{1L}^{TLL} = x_{32} y_{23} R_2 \]
\[ \Rightarrow N_{2L}^{TLL} = x_{32} y_{23} R_2 \]
\[ \Rightarrow N_{1L}^{TLL} = N_{2L}^{TLL}. \]

Right eye illuminated: \( R_1 = R_2 = 0 \quad R_3 = R_4 \neq 0 \)
\[ \Rightarrow N_{1R}^{TLL} = x_{44} (y_{14} + y_{24}) R_4 \]
\[ \Rightarrow N_{2R}^{TLL} = x_{44} (y_{14} + y_{24}) R_4 \]
\[ \Rightarrow N_{1R}^{TLL} = N_{2R}^{TLL}. \]

As a consequence, as expected, no anisocoria occurs.

A comparison between the direct and consensual pupillary reaction reveals:
\[ N_{1L}^{TLL} = x_{32} y_{23} R_2 \]
\[ N_{2L}^{TLL} = x_{32} y_{23} R_2 \]
\[ \text{under the condition: } y_{23} = \frac{\lambda_x}{1 + \lambda_x}, \quad \lambda_x = \frac{x_{11}}{x_{32}}. \]

Surprisingly, the model allows for no RAPD under one special condition given above. This may support the speculation that a tractus lesion alone does not necessarily cause an RAPD. However, under all other mathematical conditions an RAPD exists.
To obtain the condition mentioned above we calculate the following:
\[ x_{44} (y_{14} + y_{24}) R_4 = x_{32} y_{23} R_2. \]

Given the first assumption and that \( x_{44} \) can be rewritten as \( x_{44} = \lambda_x x_{32} \), where \( \lambda_x > 0 \), the equation listed above reduces to:
\[ \lambda_x (y_{21} + y_{23}) = y_{23}. \]

Introducing the normalization \( y_{21} + y_{23} + y_{24} = 1 \) (or 100%), we finally obtain the condition given above:
\[ y_{23} = \frac{\lambda_x}{1 + \lambda_x}, \quad \lambda_x = \frac{x_{11}}{x_{32}}. \]

Also under this condition, no nasotemporal difference exists. It should be mentioned that under all other conditions the nasotemporal difference does not vanish completely but remains subtle. Furthermore, it must be emphasized that the RAPD in tractus lesions has by no means been explained [3, 17]. It has been suggested that an additional lesion of the connections between the afferent pathways and the pretectum causes the RAPD observed in tractus lesions (H. Wilhelm, unpublished data). Such a situation is simulated below. A unilateral lesion of the input to one pretectum would result in the same effect caused by a unilateral lesion of the pretectum itself.

**Pretectum lesion on the left**

A unilateral pretectum lesion on the left side in our model (PLL) is realized by eliminating the couplings \( y_{11}, y_{12}, \) and \( y_{14} \) as shown in Fig. 6:

\[ \text{PLL} \iff y_{11} = y_{12} = y_{14} = 0. \]

Thus, Eqs. 1 and 2 reduce to:
\[ N_{1\text{PLL}} = x_{11} y_{21} R_1 + x_{32} y_{23} R_2 + x_{44} y_{24} R_4, \]
\[ N_{2\text{PLL}} = x_{11} y_{21} R_1 + x_{32} y_{23} R_2 + x_{44} y_{24} R_4. \]

Again, the outputs at the nuclei Edinger-Westphal are equal:
\[ N_{1\text{PLL}} = N_{2\text{PLL}}. \]

From a clinical point of view, the following observations are expected: no anisocoria should occur and there should be an RAPD. Performance of the swinging flashlight test results in:

Left eye illuminated: \( R_1 = R_2 = R_3 = R_4 = 0 \)

\[ \Rightarrow N_{1\text{PLL}} = x_{11} y_{21} R_1 + x_{32} y_{23} R_2 \]
\[ \Rightarrow N_{2\text{PLL}} = x_{11} y_{21} R_1 + x_{32} y_{23} R_2 \]

\[ \Rightarrow N_{1\text{PLL}} = N_{2\text{PLL}}. \]
Fig. 6a, b A unilateral pretectum lesion on the left side (PLL).  
a Biological description.  
b Network-like description

\[ \begin{align*} 
\Rightarrow N^{PLL}_{1R} &= x_{44}y_{24}R_4 \\
\Rightarrow N^{PLL}_{2R} &= x_{44}y_{24}R_4 \\
\Rightarrow N^{PLL}_{1L} &= N^{PLL}_{2R} \\
\end{align*} \]

As a consequence, as expected, no anisocoria occurs.

A comparison between the direct and consensual pupillary reaction gives:

\[ \begin{align*} 
N^{PLL}_{1L} &= x_{11}y_{21}R_1 + x_{32}y_{23}R_2 \\
N^{PLL}_{2R} &= x_{44}y_{24}R_4 \\
\Rightarrow N^{PLL}_{1L} &= N^{PLL}_{2R} \\
\end{align*} \]

under the condition: \( y_{24} = \lambda_x y_{21} + y_{23}, \quad \lambda_x = \frac{x_{11}}{x_{32}} \).

Therefore, it is possible that no RAPD occurs under one special condition given above, which is derived in a calculation similar to that presented above for the tractus lesion. However, under all other mathematical conditions an RAPD exists as expected.

**Discussion**

Modeling of neural pathways is helpful in the testing of a hypothesis or the planning and evaluation of experiments. The pupillary pathway is an ideal object for mathematical modeling. The existing model is based on anatomical considerations from the last century [2]. Classically, retinal ganglia cells are connected directly to pretectal neurons. Interneurons between the pretectum and the Edinger-Westphal nucleus connect the afferent and the efferent pupillary pathways.

Many experiments and clinical observations have shown that the pupillary pathway must be more complex than a simple reflex arc. As an example, the absence of the pupillary light reflex in patients with pure occipital lobe lesions cannot be explained by classic models [1, 8–10, 13, 18]. Additionally, the finding of an RAPD in patients with optic tract lesions and lesions close to the LGN is unexplainable [3, 5–7, 12, 17]. Therefore, we chose a more complex model to handle these findings analytically. We added one more coupling, assuming an additional synapse in the pupillary pathway and interconnections between both LGN regions (uncrossed fibers from the ipsilateral temporal retina should lead to the contralateral pretectum). This is the simplest amplification of the classic model. Analytically we found the following:

1. The typical situation of a unilateral amaurosis can be simulated realistically.
2. An optic tract lesion itself does not cause an RAPD as long as we do not accept the existence of a major difference in pupillomotor sensitivity between the temporal and the nasal retinal hemifield. It can clearly be demonstrated that the RAPD in a tract lesion would be caused exclusively by the difference in pupillomotor sensitivity between the remaining nasal retinal hemifield ipsilateral
of the tract lesion and the contralateral temporal retinal hemifield. The RAPD in a tract lesion has been measured in patients as being 0.3–0.9 log units [3]. This is much greater than the nasotemporal difference that can be demonstrated in humans [4, 16]. Thus far, the extended model has been as insufficient as the classic one; however, it allows the introduction of an additional lesion.

3. If neural couplings between the LGN and the pretectum are involved differentially, an RAPD can occur independently of any nasotemporal difference [5–7, 12]. This is called a pretectal lesion in our model. In reality, it is not the pretectum itself that needs to be injured but the neurons leading to the pretectum.

It must be emphasized that this model is based solely on theoretical considerations and has not yet been proven anatomically. However, the assumed interconnections are at least possible. Our results show that the nasotemporal difference is crucial in the evaluation of models of the pupillary pathway. Therefore, we need more information about the amount and the variability of such differences both in normals and in patients with lesions of the visual pathways. Inserting the results of experimental and clinical studies in our model may show which interconnections have to be involved in a lesion and this, on the other hand, could provide information about which neural pathways travel close together and which are separated from each other. These hints might facilitate anatomy studies. It should be mentioned that there are other pupillary findings that have not been sufficiently explained, such as contraction anisocoria [4, 16], which may be demonstrated in such a model and evaluated practically. For purposes of clarity, this aspect was not included in the present study.

A great advantage of the model presented herein is its simplicity and that only a few reasonable assumptions are used. The neural couplings connecting the neural layers can easily be adjusted according to neuroanatomical conditions. Although this model is a simple approach, it reveals a wide variety of pupillary reactions and is capable of explaining many clinical findings. Several lesions, e.g., lesions of the intercalated neurons between the afferent and efferent pupillary pathways, can be simulated within the model and the corresponding pupillary reactions such as anisocoria or RAPD can be calculated analytically without the need for changes in the coupling set. Furthermore, predictions made by the model may give rise to new experiments (e.g., intensified consideration of nasotemporal differences), which on their part might verify the model again and, if need be, result in its adjustment. Due to its neural network structure, the model described above can easily be extended to much more complex pupillary pathways while retaining its computational and analytical calculability.

References

2. Bach L (1908) Pupillenlehre. Karger, Berlin

1 A computer simulation of this model called PupilQuest is available for both PC and Macintosh. Interested readers should send a short message to wolgang.fink@uni-tuebingen.de to receive a copy of this simulation utility for education and research purposes only.