The aim of this thesis was to investigate and to advance the assessment paradigms in machine-learning based myoelectric upper limb prostheses, with a focus on the assessment of functional prosthesis use and on the assessment of the user skill in controlling the prosthesis.

Over the last decades, the technological components of externally powered upper limb prostheses have undergone rapid advancements. The state-of-the-art in prosthetic hands moved from simple gripper-like devices to multi-articulated hands which can perform a variety of wrist- and hand movements and which are controlled by interpreting the user's muscle activation patterns in the remnant muscles with the help of machine-learning techniques. These developments clearly addressed the needs of the users, who expressed that they desired hand prostheses with more functionalities (i.e., more possibilities to move the wrist and the fingers) and with more intuitive control interfaces. Before the introduction of the first commercially available multi-articulated hand prosthesis (iLimb, 2007), the state-of-the-art was essentially limited to hand- or hook-like devices which could perform an "open-and-close" movement of the hand aperture. Similarly, before the introduction of the first machine-learning based control interface (2013, COAPT complete control), the fundamental control concept remained relatively unchanged for a long time: The users controlled the opening and closing of the hand by contracting two different groups of muscles in their remnant arm, usually the flexors and extensors of the wrist, respectively. Two small surface electrodes embedded in the inner wall of the prosthesis socket would measure the electric potential (i.e., the myoelectric activity) of these muscles and feed this signal to the motors of the prosthetic hand, which were powered by batteries. In contrast, in machine-learning myoelectric control (also often referred to as pattern recognition prosthesis control), usually there are a multitude of electrodes which measure the muscle activity around the entire forearm. Furthermore, instead of mapping the activity of one electrode to one prosthesis function, this control paradigm maps activation patterns of all electrodes to the hand functions, which should in theory facilitate more natural and intuitive control.

With these significant technological advancements becoming commercially available, the need emerged to assess these technologies and to essentially answer the question whether these devices also delivered a meaningful improvement in the daily life of the users. However, the drastic changes in the technological parts could require a change in the way of assessing upper limb prostheses, too. That is because multi-function prosthetic hands and intuitive machine-learning based control interfaces might affect the way in which these prostheses are used. The multitude of hand and wrist functions in state-of-the art prostheses combined with intuitive control interfaces should in theory allow the user to seamlessly orient and position the hand appropriately to different objects and tasks – and therefore facilitate more natural movements and a potential decrease in the pathological compensation strategies which are well-known to occur in the "conventional" prosthetic hands. However, it was questionable whether the clinical tests for (prosthetic) upper limb function would be able to cover these aspects of prosthesis use, since most of the popular tests were designed before the arrival of machine-learning based control. Moreover, machine-learning based myoelectric control required that the users generate consistent and distinguishable muscle activation patterns, but it remained largely unknown how the skill of the user in generating such patterns, and thus the skill in controlling the prosthesis, could be quantified. This question became of particular importance in light of the evidence that machine-learning based myoelectric control hardly worked right from the start in naïve users. Evidently, the users needed to train in order to improve their control ability, and without a proper metric to quantify their control ability and without insights into the underlying mechanisms, this process essentially remained a trial-and-error-paradigm.

In summary, the technological state-of-the-art in upper limb prosthetics has demonstrated impressive and rapid advancements. However, the introduction of these technologies has opened up new questions in two domains of the assessment paradigms – namely, whether the current clinical tests of prosthetic upper limb function are still adequate to fully assess these devices and how the user skill in controlling these devices can be quantified. The goal of this thesis was to investigate and advance the assessment procedures in these domains.

The first chapter of this thesis is a general introduction which outlines the state-of-the-art and the historical development of upper limb prostheses – ultimately arriving at high-tech multi-function hands and machine-learning driven control interfaces. It describes the way in which these devices changed the scope of functional assessments through their technological advancements, and furthermore it dives into the technical challenges of assessing the control skill of the users when interacting with these prostheses. Last, the introduction outlines the requirements which the prostheses' software imposes on the users in generating "appropriate" activation patterns for the prosthesis control and it introduces the basic concepts of gauging the quality of such patterns, and how this might be related to the control skill of the users.

The second chapter investigates how state-of-the-art hand prostheses are used (and perceived) by the users (and hand therapists). In an interview study, the users and therapists were asked about the benefits and drawbacks of the prostheses, for which activities the prostheses were used, and whether multi-function prosthetic hands and/or machine-learning based control interfaces had an impact on the benefits, the drawbacks, and on the activities. We found that the multitude of the functions in state-of-the-art hands are only of little benefit when they are controlled with the conventional 2-electrode control paradigm, simply because switching the functions of the hand is too mentally demanding and slow. In contrast, we found that under machine-learning based control, the users expressed that switching the function is much easier and so they used a wider variety of the different functions depending on the activity they want to perform. Moreover, the users explained that they were generally faster with machine-learning based control interfaces and that the control felt more natural. Irrelevant of the control interface, the users and the therapists expressed that the majority of the commercially available high-tech hands were too fragile, prone to failure and too expensive. With regard to the activities performed with the prosthesis, the most important finding was that despite the technological advancements, the prosthesis was pre-dominantly used as a supporting and stabilizing assistance for the sound hand with little prehensile action and object manipulation (in individuals with unilateral limb difference).

The third chapter investigates the current state-of-the-art in clinical tests for prosthetic upper limb function, it formalizes a set of requirements which should be met by future tests to address the technological state-of-the-art in prosthetics, and it investigates how the tests currently available could be developed to meet these requirements. The chapter concludes that in order to fully evaluate the benefits and drawbacks of modern prosthetic hands, functional tests need to objectively assess the quality of the users' movements, e.g., by assessing the inter-segmental coordinative patterns or the smoothness of limb motion. That is because modern prosthetic systems promise benefits which are primarily to be expected in these domains, hence such properties need to be evaluated. The chapter concludes that the tests currently available do not fully cover these aspects and that a future generation of clinical tests needs to be developed which advances the test paradigm. Furthermore, the third chapter presents a first suggestion of how such a future generation of tests could look like. For that purpose, ten able-bodied participants and six prosthesis users performed a newly designed task, resembling an activity of daily life, while kinematics were recorded with an off-the-shelf inertial-magnetic-measurement system, which is relatively easy and fast to set up. The results between the two participant groups were compared and it was found that several kinematics-related outcomes indeed reveal differences between the movements of the prosthesis users and the able-bodied. This indicated that the suggested test could be a useful tool in evaluating the qualities of the prosthesis users' movements with respect to natural, able-bodied movements.

The fourth chapter addresses the question of how the user skill in controlling a machine-learning based myoelectric prosthesis can be quantified. The chapter investigates whether a set of EMG metrics which quantify properties of the muscular activation patterns relates to the control ability. For that purpose, a large group of able-bodied participants was trained during 5 training sessions over 5 days in controlling the output of a myoelectric machine-learning based device. The participants needed to repeatedly perform a set of hand and wrist movements while the EMG activity around the forearm was recorded. These EMG data were used to train a linear-discriminant-analysis classifier in recognizing these movements. The participants subsequently needed to perform the same movements in a match-prompt test and the outcome of the test reflected how many times the classifier recognized the prompted movements from the EMG data. Subsequently, the association between three EMG metrics and the performance in the test was assessed. The chapter focuses on three metrics suggested in the literature, related to the EMG pattern variability, separability and repeatability. The main results showed that in accordance with previous literature, the participants improved their performance in the test, reflecting an improvement in the control performance. However, in contrast to previous publications, the results showed no apparent association between the suggested EMG metrics and the control performance during the tests, indicating that these metrics are likely ill-suited to characterize the control skill in myoelectric machine-learning based control.

The fifth chapter is a follow up to the fourth chapter, where a deeper investigation of the EMG metrics was performed by studying the metrics in individual EMG features, and by studying a potential association between the metrics themselves. From a machine-learning perspective, it is somewhat surprising and contradictory that neither the EMG pattern separability, nor the EMG pattern repeatability showed any strong association with the control performance in the match prompt test. Chapter five therefore aimed to investigate possible causes for the absence of such strong associations. For that purpose, EMG data from chapter four and from two other similar studies were pooled, including data of able-bodied and individuals with upper limb difference. The chapter set out to address two research questions: (1) Do different EMG features yield differences in the separability and the repeatability? And (2) Is there an association between the separability and the repeatability? The underlying rationale to investigate these questions was that both, (1) a potential difference in separability and repeatability between individual EMG features, as well as (2) a potential association between the separability and the repeatability could be the underlying mechanisms causing the absence of an association between the EMG metrics and the control performance. We found that there were indeed significant differences in the separability and the repeatability between four common time-domain EMG features. Two features related to the EMG amplitude showed better separability and better repeatability, compared to two features related to the EMG frequency content. Moreover, we found that there was a strong association between the separability and the repeatability, where better separability was associated with worse repeatability. In conclusion, it appeared that the amplitude-dependent features might be more suitable for myoelectric machine-learning based control. Moreover, the association between the separability and the repeatability suggests that both metrics have a somewhat detrimental effect on each other. In simple words, it appeared that the more the EMG patterns were separable from each other, the more difficult it became for the user to accurately repeat them, which might explain the lack of strong associations between any of the EMG metrics and the control performance.

The sixth and last chapter is a general discussion of this thesis. The chapter discusses the wider implications of the findings with regard to the assessment paradigm and the conceptual assumptions in machine-learning myoelectric prosthesis control. It is argued that functional assessments of modern prostheses need to entail objective evaluations of movement qualities in real-life prosthesis use conditions, followed by a discussion of the challenges and possible technological solutions to implement such evaluations into routine clinical tests. Building on chapter four and five, the general discussion scrutinizes the core assumptions and the methodology of machine-learning myoelectric control. The question is raised whether control strategies, which assume that the “user intent” manifests itself in separable and repeatable patterns of EMG activity, would ever yield motor behavior of high dexterity. The chapter ends with a brief outlook into conceptually different control strategies.